cortex 74 (2016) 305–314

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Discussion forum

Should the search for bilingual advantages in executive functioning continue?

Kenneth R. Paap [a,*], Hunter A. Johnson [b] and Oliver Sawi [c]

[a] Department of Psychology, San Francisco State University, San Francisco, CA, USA
[b] Department of Psychology, New Mexico State University, Las Cruces, NM, USA
[c] Department of Psychology, University of Connecticut, Storrs, CT, USA

1. Introduction

The commentators on our target article, (Paap, Johnson, & Sawi, 2015) nicely framed its importance and impact: Hartsuiker (2015) observed that early reports of bilingual advantages in executive functioning (EF) inspired a global research effort of unprecedented magnitude in the area of bilingualism research. In the wake of the first positive findings research articles often stated or inferred that the bilingual advantage was a well-documented and unquestionable finding. The societal importance of the research question led Treccani and Mulatti (2015) to observe that the bilingual advantage is doubtless one of the most newsworthy topics in cognitive psychology and is presented by media as a matter of fact. Thus, our conclusion that bilingual advantages in EF either do not exist or are restricted to very specific and undetermined circumstances (Paap et al., 2015) was likely to encounter varied reactions. We were pleased to observe that this view was widely shared across the 21 commentaries as our count of the hanging chads tallies: 15 endorsements, 3 challenges, and 3 mixed responses. But, the sample is biased. Not in terms of the breadth of perspectives of those invited to comment on our article, but rather with respect to those willing to engage in a reasoned and public debate about the research topic.

The sounds of silence emanating from the missing commentaries resonate with Morton's (2015) point that one reason for the lack of progress in resolving inconsistencies in the literature is that "dissenting opinions are simply dismissed" as proponents of the bilingual advantage hypothesis "march on ignoring all appeals for higher standards". Jared (2015) suggests that for the bilingual advantage hypothesis to remain viable, there needs to be a clearly articulated account of the large number of failures to replicate. Klein (2015) observes that challenges to established scientific ideologies are often met with colossal disregard of evidence. Morton further observes that some key stakeholders have been quite strident in their responses to criticism. In contrast, we hope that the spirited and interesting counterarguments raised by Bak (2015), Gold (2015), Li and Grant (2015), Linck (2015); Woumans and Duyck (2015), and others in this forum and our reply to those commentaries illustrate a back and forth that is constructive. From our perspective Bak's edgy, perhaps evanescent, but explicit disagreements promote a dialog that we believe will motivate readers to think more carefully about the evidence.

Many commentators stated or inferred that there is a crisis of confidence within this research area. Gathercole (2015) finds the scientific issues somewhat intractable because they require answers to difficult and high-level questions such as what "counts" as evidence, how evidence should be interpreted, and what we consider "publishable" research. She is also the only commentator to mention that the debate is also heated by the socio-political overtones regarding the cognitive consequences of bilingualism and the tension between melting pots and heritage rights. We had no initial skepticism.

* Corresponding author. Department of Psychology, Language Attention and Cognitive Engineering Laboratory, San Francisco State University, 1600 Holloway Avenue, San Francisco, CA 94132, USA.

E-mail addresses: kenp@sfsu.edu (K.R. Paap), haj@mail.sfsu.edu (H.A. Johnson), oliver.sawi@uconn.edu (O. Sawi).
http://dx.doi.org/10.1016/j.cortex.2015.09.010

When Paap and Greenberg (2013) first started to test the bilingual-advantage hypothesis we had strong expectations that we would replicate a strong advantage in the Simon task and wanted to explore how individual and group differences in interference control might predict the pattern of costs and benefits in a category priming task. Three studies, three additional tasks, and 273 participants later we reconsidered the hypothesis and, like Klein (2015), what changed our mind was simply the weight of the evidence. Given that the overall literature provides so little indication of a positive effect, Klein ponders why the field remains full of “fervent true believers”. He suggests that there are a set of factors that are delaying the eventual refutation of the bilingual-advantage hypothesis that he refers to as Type-I incompetence. Type-I incompetence is not simply the occurrence of a Type-I error, but rather confirmation biases and scholarly lapses that create more false positives and lead to the disregard of null effects.

2. Problems and some proposed solutions

2.1. Power and publication bias

It is certainly fair to say that studies reporting significant bilingual advantages are typically underpowered and that less than a handful of labs have invested in studies with large samples sizes. Several commentators explicitly endorsed the need for more powerful designs (Bakker, 2015; de Bruin & Della Sala, 2015; Duñabeitia & Carreiras, 2015; Klein, 2015; Marzecova, 2015; Wagenmakers, 2015). Based on the meta-analytic evidence that the effect size is no larger than .3 Bakker pointed out that in order to achieve desirable levels of power at least 139 participants are required in each language group. de Bruin and Della Sala (2015) link the selective reporting of small n studies not only to the potential for false positives, but also to the decline effect. The decline effect refers to the fact that initial studies of a psychological phenomenon tend to have small sample sizes coupled with large effect sizes and that subsequent larger studies tend to have smaller effect sizes. Some commentators (Bak, 2015; Gold, 2015; Woumans & Duyck, 2015) suggested that we selectively overlooked specific positive results with large samples sizes, but these were studies of the relative onset of dementia not tests for bilingual advantages on nonverbal measures of EF. We agree that there is an intriguing link between the two hypotheses and address these studies and those using prospective rather than retrospective methods in Section 4.

In addition to using larger sample sizes there are other ways of increasing power and these include the latent-variable approach advocated by Gade (2015), reducing variability between laboratories by sharing computer scripts (Jared, 2015) or sharing instruments for assessing language proficiency (Kousaie & Taler, 2015), using extreme group designs (Marzecova, 2015), using standard meta-analysis (Linck, 2015), and using longitudinal designs (Li & Grant, 2015).

Bakker (2015), Gathercole (2015), and de Bruin and Della Sala (2015) all echo our concern that the combination of the file-drawer effect and publication bias distorts the published database. No commentators expressed any doubts that a publication bias exists. The only unknown is the degree to which the bias distorts the meta-analytical results (Bakker, 2015). Jared (2015) emphasizes that publication biases against null results exacerbate the lack of incentive for attempting replications. With respect to solutions, bias in the form of questionable research practices (QRPs) should be reduced if more journals and authors take advantage of pre-registration (Bakker, 2015; de Bruin & Della Sala, 2015; Wagenmakers, 2015). These same commentators, together with Duñabeitia and Carreiras (2015), strongly emphasize the need to change publication practices so that all good research is published regardless of the direction or significance of the results. Wagenmakers also mentions the effectiveness of adversarial collaborations and we invite a collaboration with anyone who judges our work to be, well, adversarial.

Linck (2015) chides us for referring to our analysis of the relationship between sample size and significance (see Figure 4 of our target article) as an “updated meta-analysis”. The word choice is defensible if one takes a broad definition that meta-analysis is any method for contrasting and combining results from different studies in the hope of identifying patterns that come to light only in the context of multiple studies. Be that what it may, Linck is clearly right that a traditional meta-analysis would be the better choice if the primary intent was to hone in on the average effect size. That was not our goal for two reasons. First, more traditional meta-analyses focusing on effect size had already been conducted by Hilchey, Saint-Aubin, and Klein (in press) and de Bruin, Treccani, and Della Sala (2015). We were also aware of an on-going CUNY analysis of 73 published comparisons. The results of this meta-analysis reported at the meeting of the Cognitive Science Society (Donnelly, Brooks, & Homer, 2015) yielded an average interference cost of $d = .29$ which is very similar to the value $d = .30$ reported by de Bruin et al. (2015). Given that the published database is biased against null or negative results and that some of the positive results may have been due to failures in matching, the small (but statistically significant) positive effect size does not yield definitive evidence for or against the bilingual advantage hypothesis. Turning to our second reason for not doing a traditional meta-analysis, we wanted to quantify our argument that significant bilingual advantages tend to occur with small or medium sample sizes and not with large sample sizes. As we point out this pattern is inconsistent with the hypothesis that bilingual advantages exist, but are small in size. We thought the histograms might also correct the common false impression that bilingual advantages are easy to replicate and that null results are scattered events that can be dismissed as noise.

Woumans and Duyck (2015) state that a solid meta-analysis is the best synthesis of an effect and that de Bruin et al. (2015) reported a significant bilingual advantage across studies. Despite the fact that they explicitly acknowledge that the de Bruin et al. meta-analysis also showed evidence of a publication bias for positive findings, they suggest that the current yes/no discussion between believers and non-believers does not do justice to this meta-analysis result. If we understand their argument correctly and the intent is to say that a significant finding is “*solid*” even in the presence of strong evidence that the sample is biased, then we disagree.

2.2. Confounding variables

A majority of the commentators spoke to the conundrum of matching language groups on all other variables that are likely to affect measures of EF. One aspect of the problem is the sheer number of variables that likely affect the development and maintenance of EF: genetics, intelligence, gender, culture, education, immigrant status, SES (family income, parental education (PED), occupation), music training, video gaming, exercise, sport, social interaction. Confounds may be inevitable. From a dynamical systems perspective, different populations will form qualitatively different attractor basins in a space including not only languages spoken, but also variables like SES and education level. Reasonable linear renderings of these variables are possible in each attractor basin; however, the relationship between these characterizations can differ markedly across basins. For example, bilinguals are more likely to be better educated in Hyderabad, but less educated in Houston. As long as the language(s) one speaks matters in the social dynamics of a population there are likely to be other important factors that covary with bilingualism.

We agree that mixed effects modeling can enhance the generality of an experiment, but it is not a panacea if participant variables are confounded across groups. Thus, depending upon exactly what was intended by "*can account for*", we may disagree with this statement by van Heuven and Coderre (2015):

> "*The use of sophisticated approaches like mixed effects modelling can account for the fact that bilinguals are not, and never will be, completely matched either to monolinguals or to other bilinguals*".

Kousaie and Taler (2015) draw attention to the fact that most tests of the bilingual advantage use a single monolingual group. However, when two monolingual groups are included in the same study they often differ from one another (Kousaie et al., 2014; Carlson & Choi, 2009) and produce inconsistent patterns of bilingual advantages, disadvantages, or null effects. If the use of multiple baseline (monolingual) groups were more widespread, the literature would form an even more intricate mosaic, but not one that would fill in missing pieces to form a more coherent pattern.

Bak (2015) enthusiastically agrees that more attention should be paid to potential confounding factors and rightly criticizes us for not examining immigrant status in our work. He points out that confounding variables can operate both ways and could have masked significant bilingual advantages in our studies. We now include items that enable us to determine immigrant status, but have not in the past. Many of our college student participants undoubtedly come from immigrant families, but some are highly proficient in their native language, others are clearly English dominant, and some have little or no proficiency and would clearly be considered monolinguals. Nevertheless, let's grant the assumption that immigrant-family status is likely to be confounded in our samples. Bak suggests that an immigration advantage (presumably in combination with a bilingualism advantage) in our data were canceled out by SES disadvantages. Although our average bilingual has a lower level of PED, we have examined and reported on this extensively (Paap & Greenberg, 2013; Paap & Liu, 2014; Paap & Sawi, 2014). For example, across 13 different measures of EF the correlation between PED and EF is never significant and usually very near zero (Paap & Sawi, 2014). As one would predict from the lack of correlation, our null results remain null when we match bilinguals and monolinguals on PED. Our language groups also show no differences in the Ravens test of fluid intelligence despite the differences in PED. Apparently, those from lower SES families who are successful upper-division college students have either compensated for any early disadvantages or were not disadvantaged to begin with. With respect to the latter, education and income may constrain the resources that parents provide to their children, but some parents with less income or less formal education can, nonetheless, provide strong emotional support and cognitive stimulation. Through self-selection and academic selection this subset of low SES children may, several years down the road, be over-represented in our student samples. This is not to say that we should not bother to measure SES in young adult samples or consider those data carefully.

Woumans and Duyck (2015) suggested that our review was somewhat selective and failed to discuss key studies that paid specific attention to demographic factors and documented a bilingual advantage of "some sort". We agree that the Bialystok and Viswanathan (2009) study that uses the faces task provides some of the best evidence for bilingual advantages in inhibitory control. However, it does stand in marked contrast to the larger-scale studies by Antón et al. (2014); Duñabeitia et al. (2014), and Gathercole et al., (2014) which include children in the same age range. A large n study that includes the gaze conditions from the faces task and additional measures of inhibitory control would be very worthwhile.

Woumans and Duyck (2015) also cite a study by Yang, Yang, and Lust (2011) that suggests that bilingualism trumps culture, but each of the four language groups has only 13 or 15 participants. The study by Carlson and Choi (2009) that we discuss in our target article (Section 3.3.3 Cultural differences) has 53 to 69 children in each language group and used multiple measures of EF. The results are just the opposite showing that the effects of culture can overwhelm the language factor.

2.3. Direction of causality

Another complication in the study of differences between naturally occurring populations is that significant relationships need not be causal and, even when they are, the direction can be ambiguous. Thus, although the control required to learn a second language could enhance general EF, it is just as likely that individuals with better EF are better able to master a second language (Klein, 2015). This possibly is strongest when considering sequential bilinguals, especially when the L2 is acquired through formal instruction. For example, Kempe, Kirk, and Brooks (2015) refer to studies showing that adult second language learning consistently shows that nonverbal intelligence and working memory capacity are strong predictors of learners' success. The role of EF on language proficiency will be muted for bilinguals acquiring both languages very early, but could still influence the highest level attained, for example, on our scale a 6 (as fluent as a native speaker) versus a 7 (super fluency, more fluent that a native speaker).

Li and Grant (2015) also emphasize the importance of considering causal relationships in both directions and suggest that the current debate has focused too narrowly on viewing "type of language experience" as the independent variable (IV) and "level of EF" as the dependent variable (DV) and that the reverse IV–DV relationship should also be explored. Some traditionalist may be reluctant to characterize either bilingualism or level of EF as an IV, preferring to maintain the distinction between variables that can be manipulated through random assignment versus groups that differ with respect to a pre-existing condition. From this perspective it does not make sense to criticize us for being content to study the "relationship" between bilingualism and EF rather than trying to identify the causal link. The purpose of our target article was not only to examine the degree and consistency of the relationship between the two, but also to explore several alternative explanations for what may have caused the statistically significant outcomes. But quasi-experimental designs, by definition, do not manipulated IVs and do not allow inferences about causal relationships in either direction. However, if Li and Grant's point is simply that future research should be motivated by causal theories and that the relationship between language experience and EF ability might be bidirectional, then we agree.

A good example of the need to consider reverse causality is the study by Ladas, Carroll, and Vivas (2015) brought to our attention by Woumans and Duyck (2015). Two experiments recruit low-SES children who are either Greek-Albanian bilinguals or Greek monolinguals. Across multiple tasks and measures in both experiments there are no significant language group differences. These include the three attentional networks derived from the ANT task, global RT, and two tests of pragmatic ability. Thus, the preponderance of evidence favors no language-group differences. However, Woumans and Duyck see a silver lining because in one of the two experiments there was a significant correlation between switch-cost asymmetry and the interference effect **within** the bilingual group. Given this discussion of the causal direction of the relationship between language learning and EF it is entirely possible that in these budding bilinguals those most proficient in their L2 were those who had higher EF ability to begin with.

2.4. Baselines and interactions

Wagenmakers stresses the proper interpretation of interactions with a clarifying quartet of interactions all of which show a 200 ms congruency effect for the monolinguals and only a 100 ms effect for the bilinguals. Advancing arguments similar to those we presented in Section 3.4 of our target article Wagenmakers (2015) explains that only the interaction where performance for the congruent condition is the same for both groups warrants the conclusion of a bilingual advantage. Gathercole (2015) makes the very cogent point that given the impossibility of perfectly matching language groups baseline differences are important and must be adequately accounted for. A pessimist would point out that we do not have good tools for doing so.

Linck (2015) challenges our and Wagenmakers' assertion that a bilingual advantage in inhibitory control is inconsistent with a pattern of interaction that shows a monolingual advantage on the conflict-free congruent trials and comparable performance on the incongruent trials (see Figure 3 of our target article). He suggests that better inhibitory control may also influence performance through more efficient "disengagement" of inhibitory mechanisms, which, in turn, may manifest on non-conflict trials. If bilingual advantages in disengagement occur in nonverbal interference tasks, then they should appear when a congruent (C) trial follows an incongruent (I) trial forming an IC sequence. If bilinguals are better at this disengagement then they should be faster on IC trials compared to monolinguals. In fact, the congruency sequence effects (CSE) are identical for bilinguals and monolinguals in both the Simon and flanker tasks (Mendes, 2015; Paap & Greenberg, 2013). For example, Fig. 1 is the CSE for both bilinguals (left) and monolinguals (right) for the Simon task based on the means reported in Paap and Greenberg.

2.5. Lack of convergent validity

Wagenmakers (2015) evokes a feeling of depression that our collective research effort may have been a waste of time, effort, and resources. This waste will be compounded to the extent that the prevalent measures of EF (especially the inhibitory control and monitoring components) have no convergent validity. It appears that most of the research has not really been studying the relationship between bilingualism and EF, but rather the relationship between bilingualism and a host of task-specific mechanisms that have little interest or implication beyond the "psychology" of the Simon task or Stroop task or flanker task. Our concerns are summarized in Section 3.8 of our target article, but given that only four commentators addressed this problem we hope that bilingualism researchers take seriously the breadth and depth of the problem as presented in Paap and Sawi (2014).

van Heuven and Coderre (2015) suggest that it is time to move away from tasks like the flanker and Simon that lack convergent validity, but primarily because they may be too simplistic to elicit bilingual advantages. They suggest that bilingual advantages are more likely in higher-level cognitive functions such as working memory, metalinguistic awareness, and representational skills. This may be a worthwhile shift in research direction. If working memory is equated with updating, the proposal is a blend of standard components of EF and complex cognition. As Jared (2015) points out, aiming at new targets (her example was *coordination*) should be theoretically motivated and subject to the same demands for validity and reliability.

In contrast to van Heuven and Coderre's recommendation, Gade (2015) proposes that it is worthwhile to continue to test for bilingual advantages in EF, but that it requires large-scale test batteries, which allow the extraction of latent variables. As this is an energetic extension of our oft-expressed recommendation that studies should always include at least two tasks and at least two measures of any targeted EF component, we agree. To date, when multiple measures are combined with large sample sizes null results dominate and this is clearly the case in the von Bastian, Sousa, and Gade (2015) study.

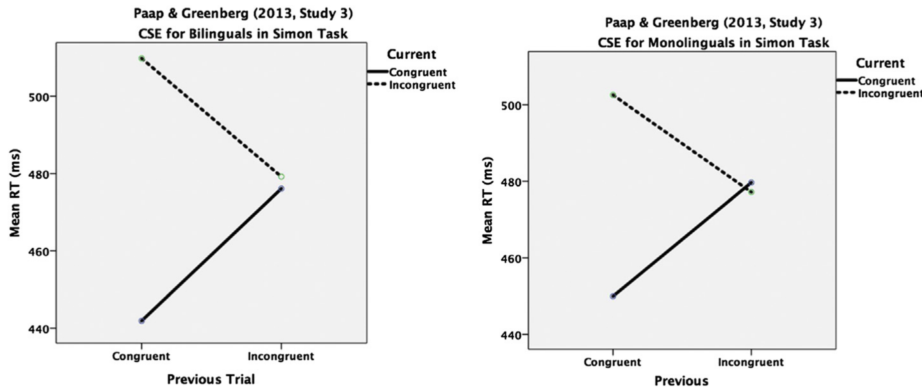

Fig. 1 – Congruency sequence effects (CSE) for bilinguals (left) and monolinguals (right) in a Simon task (Paap & Greenberg, Study 3).

3. Strategic recommendations

Many commentators offered strategic recommendations, including some that were very picturesque. Duñabeitia and Carreiras (2015) suggest that those who believe that the earth is flat should consult with those who have sailed beyond the land's view of the horizon. Hartsuiker (2015) recommends that we should stop hunting for treasure by randomly digging holes in uninhabited islands. Marzecova (2015) questions the utility of additional Loch Ness monster sightings. Morton (2015) suggests that we need to collaborate like the physicists at CERN.

3.1. *More theory-based research*

Beyond these metaphors is a majority opinion that progress requires more and better theories. Although Green's inhibitory control model (ICM) did receive kudos for launching early studies predicting bilingual advantage in inhibitory control (but see the discussion of explicit mechanisms in both Gade, 2015 and Hartsuiker, 2015) it has not proven to be helpful in predicting or explaining why advantages in inhibitory control appear infrequently or why advantages sometimes extend equally to both congruent and incongruent trials. Hartsuiker (2015) challenging commentary points out that we have no theory that predicts what circumstances engage language control, no theory of skill generalization, and no cognitive control theory that specifies what EFs can be improved by such generalization. Treccani and Mulatti (2015) are equally firm asserting that one of the most serious faults is the lack of a clear, well grounded, and broadly endorsed theory about how the management of two languages affects EF. Jared (2015) makes a similar appeal starting with her title: *What is the theory?* We agree that new research predicated on explicit models is more likely to be productive.

The hypothesis that there are other special experiences that may also influence the development or maintenance of EF further clouds the landscape. Viewing bilingualism as another case of specialized training or practice, gives rise to alternative expectations. One is that transfer is ubiquitous and, therefore, it would be surprising if bilingualism was an exception and did not enhance EF. The corollary to this proposition is that because transfer is ubiquitous it will be difficult for the specific effects of bilingualism to be detected (Valian, 2015; Gade, 2015). Paap (2015) expressed skepticism that this provides much cover for the plethora of null results. If bilinguals and monolinguals were equally likely to engage in activities that enhance EF then one would not expect a real bilingual advantage to be canceled out very often. Furthermore, monolinguals are not likely to engage in substantially more of these activities if their number is large and varied as suggested by Valian (music training, video gaming, sports, dance, acting, cooking, etc.) We are more sympathetic to the related argument that if many different "special" experiences enhance EF, then most young adults will be at asymptote and the contribution of bilingual experience will be masked.

It is noteworthy than other commentators expressed the opposite view, that the relevant research supports the conclusion that practice effects are proximal and that far transfer is rare (Hartsuiker, 2015; Klein, 2015; Wagenmakers, 2015). If this assessment is correct then bilingual advantages need to be predicated on the assumption that bilingual control is very special and an exception to the rule that specialized training transfers to near tasks, but not far ones.

3.2. *More neuroscience*

The commentators who actually do cognitive neuroscience expressed various degrees of optimism that neural data will contribute to the understanding of the relationship between bilingualism and EF. We would like to shape our response carefully by making some useful distinctions. First, we do not question that functional and structural imaging contributes to our understanding of neural processing and neuroanatomy. Second, we do not dispute that neural data can reveal differences in neural processing that are completely hidden in the behavioral measures (van Heuven & Coderre, 2015; Kousaie & Taler, 2015; Li & Grant, 2015). Third, we agree that bilingualism can lead to a substantial cortical reorganization that is interesting in its own right (Duñabeitia & Carreiras, 2015; Kousaie & Taler, 2015; Vaughn, Greene, Ramos Nuñez, & Hernandez, 2015). But, fourth (as initially asserted by Hilchey & Klein, 2011) we still maintain that the specific question of whether or not managing two languages produces a bilingual advantage in actual performance can only be adjudicated by behavioral data such as speed and accuracy-a position endorsed by Duñabeitia and Carreiras (2015) and Treccani and Mulatti (2015). Thus, when neural differences fail to align with behavioral differences we have evidence for brain plasticity, but no evidence for a bilingual advantage.

We are not taking a pedantic stance that alignment must occur every time. Rather alignment should be required when using measures that have a long track record of failing to replicate (e.g., interference effects in the Simon task). Hypothetically, if we had a task that consistently produced Loch Ness monster sightings (or bilingual advantages) when sample sizes were large (achieving say a desired power of .80), we would still expect null results 20% of the time and that should sound no alarm. However, that is not the current lay of the loch. A point of mild contention is Gold's (2015) view that we cavalierly dismissed the results of Gold, Kim, Johnson, Kryscio, & Smith, 2013 because the group differences in behavioral switch costs did not quite reach conventional statistical significance. Although it may sound like a quibble switching cost is another measure where there has been a preponderance of null effects (see Section 3.6 of our target article). Consequently, small differences obtained with small samples sizes should be interpreted cautiously. Gold (2015) also suggested that we might be missing the big picture by ignoring the impressive correlations between the BOLD responses in the regions of interest and better task-switching performance. As impressive as this relationship was across **all** participants, it does not change the fact that the group differences in the magnitude of that relationship were small and statistically marginal.

van Heuven and Coderre (2015) in referring to some of their work using both ERP and fMRI make the point that group differences that occur in specific early phases of processing and/or in certain cortical regions (in the absence of behavioral differences) can generate new hypotheses about bilingual language control and its relationship to cognitive control. These neural clues may spark new hypotheses and experiments, but if the correlation still does not emerge and there is still misalignment then we repeat our mantra that the neurophysiology demonstrates brain plasticity, but does not provide evidence for a bilingual advantage in performance.

We agree with Duñabeitia and Carreiras (2015) that the inconsistency of behavioral findings cannot be settled by structural or functional brain differences, and that simply changing the arena of the debate from cognition to the brain is not going to be helpful. Treccani and Mulatti (2015) note that the alignment and valence ambiguity problems we identified in the neurophysiological realm generalize to neuroanatomy. Their detailed analysis of two studies comparing monolinguals and bilinguals in terms of WM integrity illustrates how **opposite** patterns of differences in fractional anisotropy values are, nonetheless, both interpreted as bilingual advantages. It seems that no matter the outcome in the study of brain matter, differences are usually interpreted as accumulating evidence for the bilingual advantage hypothesis. If this is so, then it is fair to ask what conceivable pattern of neuroanatomical similarities or differences could falsify the hypothesis?

3.3. *More longitudinal studies*

Li and Grant (2015) suggest that neuroscience in combination with longitudinal designs may reduce uncertainty about causal mechanisms. Here is the ideal scenario. Suppose that participants are randomly assigned as either L2 learners or to some control activity and the two groups are equal at baseline in terms of both behavioral and neural measures of EF obtained in a nonlinguistic task, but that after intensive training the L2 learners are now superior to the control group on both sets of measures. This would enable the conclusion that the changes in the brain were caused by the L2 training and, in turn, caused the performance advantage. A casual reading of Li and Grant's commentary may lead to the impression that there are multiple studies that follow this script and provide compelling evidence that a fairly short period of L2 learning leads to enhanced EF. But, the studies cited by Li and Grant take liberties from the ideal script.

The study by Grant, Fang, and Li (2015) has no control group and no measures of cognitive control that are derived from nonlinguistic tasks. Showing that the overall activation in cognitive-control areas during lexical processing decreases over two semesters of Spanish classes has other interesting implications, but has no implications for the question of whether managing two languages enhances EF. The study by Yang, Gates, Molenaar, and Li (2015) compares a group of 23 English speakers who received six weeks of training on a novel tonal vocabulary to a group of 16 "non-learners". The two groups relied on different brain networks to process tonal and lexical information of target L2 words, but given the absence of data derived from nonverbal measures of EF there are no implications for the hypothesis examined in our target article beyond the mere fact that 18 training sessions on this task changes the way these stimuli are neurally processed.

The only study cited by Li and Grant (2015) that includes baseline measures of EF (a Go/Nogo task and a letter fluency task) is the study by Sullivan, Janus, Moreno, Astheimer, and Bialystok (2014) that followed two groups of participants, one studying Spanish and the other Introductory Psychology for six months. The results depart from the ideal described above. First, after six months of study there were differences between the two groups on some of the ERP measures, but not in performance. Thus, once again we confront the alignment problem. Proponents of the bilingual advantage hypothesis are attracted to the idea that the neural measures are more sensitive and that as L2 learning continues they will become sufficiently strong to produce bilingual advantages in behavior. Skeptics inquire about the strength and consistency of the neural measures and whether there is "valence ambiguity" regarding whether an increase in amplitude causes increases or decreases in performance.[1] For present purposes note that in the go/nogo task the predicted differences occurred for P3, but not for

[1] See Paap (2014) and Paap, Sawi, Dalibar, Darrow, & Johnson (2014, 2015) for a more in depth discussion of the valence ambiguity associated with the N2 in the go/nogo task and the N400 and P600 in the grammaticality task.

N2; and that the P3 differences were only marginal ($p = .056$). Likewise, in the grammaticality task the predicted differences occurred for the P600, but not for N400; and the P600 differences were only marginal ($p = .056$). In part, these inconsistencies are driven by completely inadequate power where after attrition and ERP signal failures the groups are down to n's of 21 and 22 participants. This is an intriguing first step, but more compelling evidence will require a substantially larger study that follows the training to a point where the neural differences are aligning with behavioral differences.

4. Bilingualism, aging, and cognitive differences

We had no intention of presenting a complete review of the research investigating the relationship between bilingualism and cognitive decline. Although the topic is relevant and important it does not directly address the main question of whether bilingualism enhances EF as most of the research explores cognitive ability in a much broader sense. Our target article (Section 3.3.2) mentioned that studies testing for bilingual advantages in EF with older adults have yielded advantages when immigrant status is confounded and null results when it is not. As an example we pointed out that Bialystok, Craik, and Freedman (2007) reported a bilingual advantage with respect to the onset of mild cognitive impairment (MCI) or dementia but that bilingualism and immigrant status were confounded. Both Bak (2015) and Gold (2015) criticized us for not mentioning several studies with large sample sizes and reasonable controls for immigrant status and other potential confounds. Here we make amends.

4.1. *Bilingual advantages in cognitive ability among aging populations*

The primary finding in the study by Bak, Nissan, Allerhand, and Deary (2014) was that individuals who learned a second language after the age of 11 years achieved better results in their 70's than would be predicted from their childhood performance, ergo, it is suggested that language learning and use can mitigate the effects of cognitive aging, independent of childhood intelligence. This study is intriguing, but perhaps less compelling when considered in full context. There were a total of six tests administered to this elderly cohort and the language learners performed better on just two of them: a measure of general intelligence and reading words with irregular spelling-sound correspondence. There were no effects on verbal fluency, a test purported by others (Luo, Luk, & Bialystok, 2010; Ljungberg, Hansson, Andrés, Josefsson, & Nilsson, 2013; Shao, Janse, Visser, & Meyer, 2014) to reflect EF. Only 35% of the language learners were actively bilingual. These effects, where they do occur, could be caused by second language learning or, alternatively, by rich and intense language processing in their L1. The effects are not very large, not very consistent, and were apparently achieved and maintained without the need to remain actively bilingual.

4.2. *Retrospective studies on the onset of decline and dementia*

In an earlier discussion forum Fuller-Thomson (2015) suggested that the retrospective studies that begin assessment when patients present themselves at a memory clinic and then estimate the onset of MCI retrospectively on the basis of patient and caretaker reports are open to biases in sampling, measurement, and publication. In addition to these general points, some remarkable anomalies in the retrospective study by Chertkow et al. (2010) have induced further skepticism on our part that these methods and procedures yield trustworthy results. Consider these two results: (1) French-speaking Canadian **monolinguals** were diagnosed with dementia an average of 5.3 years earlier than English-speaking Canadian **monolinguals**. This is a very large effect, as large as any reported difference between monolinguals and multilinguals, and no cogent explanation was found. This comparison involved, relatively speaking, very large groups: 290 English speaking monolinguals and 66 French-speaking monolinguals. If two large groups of monolinguals (all non-immigrants and all living in the same city, with educational levels matched and with any SES differences actually favoring the French speaking monolinguals) can produce large differences in onset of diagnosis, then what level of confidence should we have if, hypothetically, a new study reported a 5-year protective effect of bilingualism? (2) The same Canadian cohort of bilinguals was diagnosed with **earlier** (2.6 years) onset of diagnosed dementia compared to the Canadian cohort of monolinguals. This significant difference is the reverse of all previous significant differences between monolinguals and bilinguals in these retrospective studies.

Setting these concerns about the retrospective design aside, it is brought to our attention by Bak that one of the largest longitudinal studies ($n = 814$) of bilingualism and cognitive aging found a relationship between the number of languages spoken and performance on various cognitive tasks (Kavé, Eyal, Shorek, & Cohen-Mansfield, 2008). This study did not include monolinguals and, consequently, provides no evidence for a bilingual advantage in comparison to monolinguals. This is not to say that the differences between bilinguals, trilinguals, and multilinguals are not interesting, but even those differences are compromised by a substantial confounding with years of education. Kavé et al. make a common claim that they have "statistically controlled for" education, but as Miller and Chapman (2001) have warned "control" is altogether the wrong metaphor for understanding what ANCOVA accomplishes and that when the independence assumption is violated the regression adjustment may either obscure part of the group effect or produce spurious effects (see both our target article and Paap, Johnson, & Sawi, 2014). Another study that we failed to mention, Perquin et al. (2013), also used bilinguals as their baseline group (i.e., there was no monolingual group) and the results indicated that an appreciable effect of multilingualism may require three languages, practiced as early in life as possible.[2]

[2] This contrasts with results showing no advantages for trilinguals over bilinguals in young adults on nonverbal measures of EF (Paap, Johnson, & Sawi, 2014).

We agree with Bak (2015) that the Woumans, Santens, Sieben, Versijpt, Stevens, and Duyck (2015) retrospective study of mild cognitive decline and AD diagnosis is well controlled with respect to immigrant status and culture. Woumans et al. report a bilingual advantage of 4.6 years in symptom manifestation and 4.8 years in AD diagnosis. However, a deeper look at the results shown in their Table 2 is disconcerting. The unadjusted means show advantages of only 1.3 and 1.7 years, respectively. The language-group differences in the adjusted means are driven entirely by the 16 monolinguals with higher occupations who are younger when they first show up at the clinic and are better educated. Does bilingualism afford protection against cognitive decline? To paraphrase Morton's concern: this might well be true, but the more obvious interpretation is that there was little evidence for a bilingual advantage—it was mostly introduced by statistical means.

The study by Alladi et al. (2013) reporting that bilingual patients developed dementia 4.5 years later than the monolinguals deserves discussion because, as Bak points out, all participants were non-immigrants and similar with respect to culture. However, given the nature of studies that use samples of individuals who present themselves at clinics the language groups in these studies dramatically differ in other ways. The bilinguals were better educated, were from higher skill occupations, and included a higher proportion of men and a higher proportion from urban populations. However, the univariate regression analysis showed that number of languages was associated with age at onset of dementia **after** adjusting for the other variables. There was also a bilingual advantage among the subset of illiterates and this demonstrates that the education variable is not sufficient to account for all the group differences. We agree with Bak (2015) and Gold (2015) that this is an important study, although the marked differences on the four demographic variables should not be lightly dismissed.

A fair summary of the prospective studies is that most show that bilinguals develop dementia later than monolinguals, although sometimes-significant differences emerge only when other confounding variables are statistically taken into account. It is also the case that Gollan, Sandoval, & Salmon (2011) found bilingual advantages in some groups, but not others and that Clare et al. (2014) found no statistically significant differences at all. Furthermore, the question of whether bilingualism alone provides protection or if the threshold for protection is three languages remains open. Despite the caveats, the evidence from the retrospective studies makes the hypothesis that bilingualism builds some type of cognitive reserve attractive. But, this is not the only evidence and as briefly alluded to in our target article, the longitudinal studies of the onset of MCI and dementia call the working hypothesis into serious question.

4.3. *Longitudinal studies of cohorts studied over time*

Six studies have used a prospective cohort design following individuals without dementia at baseline. In his commentary Bak does not mention any of them. A study by Wilson, Boyle, Yang, James, and Bennett (2015) is the only one that resulted in a bilingual advantage. The other studies all produced nonsignificant results (Crane et al., 2009; Lawton, Gasquoine, & Weimer, 2015; Sanders, Hall, Katz, & Lipton, 2012; Yeung, St. John, Menec, & Tyas, 2014; and Zahodne, Schofield, Farrell, Stern, & Manly, 2014) with three trending in the direction of a monolingual advantage. Thus, the longitudinal studies show very little support for a bilingual advantage with the exception of the Wilson et al. study. One might note that the Wilson et al. study examined only the number of years of language instruction before age 18, did not measure current language use, only examined the onset of MCI, and found an advantage only on the non-amnestic tests. As we said in our target article, if the prospective studies are weighted more heavily, there is little evidence that bilingualism protects against cognitive decline.

5. Conclusion

Several commentators arrived at the conclusion that bilingual advantages in EF **probably** do not exist. No one expressed the view that bilingual advantages exist and that we know the necessary or sufficient conditions for producing them. Many advocate for a shift away from the yes-no question of does bilingualism enhance EF, but in somewhat different directions. The new direction for some is to forego comparisons of bilinguals to monolinguals in favor of examining the relationship between continuous measures of bilingualism and EF in which the inclusion of pure monolinguals would not be necessary. We would be sympathetic to such a shift if it was decoupled from strong inferences (or explicit conjectures) that the relationships observed within bilinguals support bilingual advantages (in comparison to monolinguals) in EF. If strong inferences of this type were allowed, then this would be a license to answer the yes-no question affirmatively while taking all the null results off the evidentiary table. Another popular direction is to rely far more on theory-based studies. We continue to be advocates for this approach, but it will not be easy; especially if the paucity of current theory is at the depths implied by Hartsuiker. Finally, several commentators expressed optimism that neuroscience will be the key that unlocks the elusive relationship between bilingual language control and general cognitive control. Given that good cognitive models are required in order to design the conditions in an fMRI or ERP experiment and are then used to interpret any neural differences observed, we think the priority should be for new and better cognitive models.

Acknowledgments

We thank Greg Stone for his guidance in using dynamical systems theory as a framework for understanding why bilingualism is likely to be confounded with other factors and how these patterns may dramatically vary across different cultures.

REFERENCES

Alladi, S., Bak, T. H., Duggirala, V., Surampudi, B., Shailaja, M., Shukla, A. K., et al. (2013). Bilingualism delays age at onset of

dementia, independent of education and immigration status. *Neurology, 8*, 1938–1944.
Antón, E., Duñabeitia, J. A., Estévez, A., Hernández, J. A., Castillo, A., Fuentes, L. J., et al. (2014). Is there a bilingual advantage in the ANT task? Evidence from children. *Frontiers in Psychology: Language Sciences, 5*, 1–12.
Bak, T. H. (2015). Beyond a simple "yes" or "no". *Cortex, 73*.
Bakker, M. (2015). Power problems: n > 138. *Cortex, 73*.
Bak, T. H., Nissan, J. J., Allerhand, M. M., & Deary, I. J. (2014). Does bilingualism influence cognitive aging? *Annals of Neurology*, 1–5.
von Bastian, C. C., Sousa, A. S., & Gade, M. (2015). No evidence for bilingual cognitive advantages: A test of four hypotheses. *Journal of Experimental Psychology: General*. in press.
Bialystok, E., Craik, I. M. F., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia, 45*, 459–464.
Bialystok, E., & Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition, 112*(3), 494–500.
de Bruin, A., & Della Sala, S. (2015). The decline effect: how initially strong results tend to decrease over time. *Cortex, 73*.
de Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive advantage in bilingualism: an example of publication bias. *Psychological Science, 26*(1), 99–107.
Carlson, S. M., & Choi, H. P. (2009, April). *Bilingual and bicultural: Executive function in Korean and American children*. Paper presented at the 2009 biennial meeting of the society for research in child development. Denver, Colorado.
Chertkow, H., Whitehead, V., Phillips, N., Wolfson, C., Atherton, J., & Bergman, H. (2010). Multilingualism (but not always bilingualism) delays the onset of Alzheimer disease: evidence from a bilingual community. *Alzheimer Disease and Associated Disorders, 24*(2), 118–125.
Clare, L., Whitaker, C. J., Craik, F., Bialystok, E., Martyr, A., Martin-Forbes, et al. (2014). Bilingualism, executive control and age at diagnosis among people with early-stage Alzheimer's disease in Wales. *Journal of Neuropsychology*.
Crane, P. K., Gibbons, L. E., Arani, K., Nguyen, V., Rhoads, K., McCurry, S. M., et al. (2009). Midlife use of written Japanese and protection from late life dementia. *Epidemiology, 20*(5), 766–774.
Donnelly, S., Brooks, P. J., & Homer, B. (2015, July). *Examining the bilingual advantage on conflict resolution tasks: A meta-analysis*. Paper presented at the 37th Annual Meeting of the Cognitive Science Society. Pasadena, CA.
Duñabeitia, J. A., & Carreiras, M. (2015). The bilingual advantage: acta est fabula? *Cortex, 73*.
Duñabeitia, J. A., Hernández, J. A., Antón, E., Macizo, P., Estévez, A., Fuentes, L. J., et al. (2014). The inhibitory advantage in bilingual children revisited. *Experimental Psychology, 61*(3), 234–251.
Fuller-Thomson, E. (2015). Emerging evidence contradicts the hypothesis that bilingualism delays dementia onset. *Cortex*, 1–3.
Gade, M. (2015). On tasks and cognitive constructs for the bilingual (non-) advantage. *Cortex, 73*.
Gathercole, V. C. M. (2015). Are we at a socio-political and scientific crisis? *Cortex, 73*.
Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Vinas-Guasch, R., et al. (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*, 11.
Gold, B. T. (2015). Executive control, brain aging and bilingualism. *Cortex, 73*.
Gold, B. T., Kim, C., Johnson, N. F., Kryscio, R. J., & Smith, C. D. (2013). Lifelong bilingualism maintains neural efficiency for cognitive control in aging. *The Journal of Neuroscience, 33*(2), 387–396.
Gollan, T. H., Sandoval, T., & Salmon, D. P. (2011). Cross-language intrusion errors in aging bilinguals reveal the link between executive control and language selection. *Psychological Science, 22*, 1155–1164.
Grant, A. M., Fang, S.-Y., & Li, P. (2015). Second language lexical development and cognitive control: a longitudinal fMRI study. *Brain and Language, 144*, 35–47.
Hartsuiker, R. J. (2015). Why it is pointless to ask under which specific circumstances the bilingual advantage occurs. *Cortex, 73*.
van Heuven, W. J. B., & Coderre, E. L. (2015). A call for sophisticated statistical approaches and neuroimaging techniques to study the bilingual advantage. *Cortex, 73*.
Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for plasticity of executive control processes. *Psychonomic Bulletin & Review, 18*, 625–658.
Hilchey, M. D., Saint-Aubin, J., & Klein, R. M. (2015). Does bilingual exercise enhance cognitive fitness in non-linguistic executive processing tasks. In J. W. Schwieter (Ed.), *Cambridge handbook of bilingual processing*. Cambridge University Press (in press).
Jared, D. (2015). What is the theory? *Cortex, 73*.
Kavé, G., Eyal, N., Shorek, A., & Cohen-Mansfield, J. (2008). Multilingualism and cognitive state in the oldest old. *Psychology and Aging, 23*(1), 70.
Kempe, V., Kirk, N. W., & Brooks, P. J. (2015). Revisiting theoretical and causal explanations for the bilingual advantage in executive functioning. *Cortex, 73*.
Klein, R. M. (2015). On the belief that the cognitive exercise associated with the acquisition of a second language enhances extra-linguistic cognitive functions: is type I incompetence at work here? *Cortex, 73*.
Kousaie, S., Sheppard, C., Lemieux, M., Monetta, L., & Taler, V. (2014). Executive function and bilingualism in young and older adults. *Frontiers in Behavioral Neuroscience, 8*.
Kousaie, S., & Taler, V. (2015). The bilingual advantage: elusive but worth the effect? *Cortex, 73*.
Ladas, A. I., Carroll, D. J., & Vivas, A. B. (2015). Attentional processes in low socioeconomic status bilingual children: are they modulated by the amount of bilingual experience? *Child Development, 86*(2), 557–578.
Lawton, D. M., Gasquoine, P. G., & Weimer, A. A. (2015). Age of dementia diagnosis in community dwelling bilingual and monolingual Hispanic Americans. *Cortex, 66*, 141–145.
Li, P., & Grant, A. (2015). Identifying the causal link: two approaches toward understanding the relationship between bilingualism and cognitive control. *Cortex, 73*.
Linck, J. (2015). Methods matter for critical reviews too. *Cortex, 73*.
Ljungberg, J. K., Hansson, P., Andrés, P., Josefsson, M., & Nilsson, L.-G. (2013). A longitudinal study of memory advantages in bilinguals. *PLoS One, 8*(9), e73029.
Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition, 114*, 29–41.
Marzecova, A. (2015). Bilingual advantages in executive control – loch Ness Monster case or an instance of neural plasticity? *Cortex, 73*.
Mendes, C. G. (2015). *The impact of bilingual on conflict control* (Unpublished doctoral dissertation). Dunedin, New Zealand: University of Otago.
Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology, 110*(1), 40–48.
Morton, J. B. (2015). Still waiting for real answers. *Cortex, 73*.
Paap, K. R. (2014). The role of componential analysis, categorical hypothesizing, replicability and confirmation bias in testing

for bilingual advantages in executive functioning. *Journal of Cognitive Psychology, 26*(3), 242–255.
Paap, K. R. (2015). Do many hones dull the bilingual whetstone? *Bilingualism: Language & Cognition, 18*(1), 41–42.
Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology, 66*, 232–258.
Paap, K. R., Johnson, H. A., & Sawi, O. (2014). Are bilingual advantages dependent upon specific tasks or specific bilingual experiences? *Journal of Cognitive Psychology, 26*(6), 615–639.
Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex, 73*.
Paap, K. R., & Liu, Y. (2014). Conflict resolution in sentence processing is the same for bilinguals and monolinguals: the role of confirmation bias in testing for bilingual advantages. *Journal of Neurolinguistics, 27*(1), 50–74.
Paap, K. R., & Sawi, O. (2014). Bilingual advantages in executive functioning: problems in convergent validity, divergent validity, and the identification of the theoretical constructs. *Frontiers in Psychology*, 5(962), 1–15.
Paap, K. R., Sawi, O., Dalibar, C., Darrow, J., & Johnson, H. A. (2014). The brain mechanisms underlying the cognitive benefits of bilingualism may be very difficult to discover. *AIMS Neuroscience, 1*(3), 245–256.
Paap, K. R., Sawi, O., Dalibar, C., Darrow, J., & Johnson, H. A. (2015). Beyond panglossian optimism: larger N2 amplitudes probably signal a bilingual disadvantage in conflict monitoring. *AIMS Neuroscience, 2*(1), 12–17.
Perquin, M., Vaillant, M., Schuller, A.-M., Pastore, J., Dartigues, J.-F., Lair, M.-L., et al. (2013). Lifelong exposure to multilingualism: new evidence to support cognitive reserve hypothesis. *PLoS One, 8*(4), e62030.
Sanders, A. E., Hall, C. B., Katz, M. J., & Lipton, R. B. (2012). Non-native language use and risk of incident dementia in the elderly. *Journal of Alzheimer's Disease*, 99–108.
Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology*, 5(772), 1–10.
Sullivan, M. D., Janus, M., Moreno, S., Astheimer, L., & Bialystok, E. (2014). Early stage second-language learning improves executive control: evidence from ERP. *Brain and Language, 139*, 84–98.
Treccani, B., & Mulatti, C. (2015). No matter who, no matter how… and no matter whether the white matter matters. Why theories of bilingual advantage in executive functioning are so difficult to falsify. *Cortex, 73*.
Vaughn, K. A., Greene, M. R., Ramos Nuñez, A. I., & Hernandez, A. E. (2015). The importance of neuroscience in understanding bilingual cognitive control. *Cortex, 73*.
Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition, 18*(1), 3–24.
Wagenmakers, E.-J. (2015). A quartet of interactions. *Cortex, 73*.
Wilson, R. S., Boyle, P. A., Yang, J., James, B. D., & Bennett, D. A. (2015). Early life instruction in foreign language and music and incidence of mild cognitive impairment. *Neuropsychology, 29*(2), 292–302.
Woumans, E., & Duyck, W. (2015). The bilingual advantage debate: moving toward different methods for verifying its existence. *Cortex, 73*.
Woumans, E., Santens, P., Sieben, A., Versijpt, J., Stevens, M., & Duyck, W. (2015). Bilingualism delays clinical manifestation of Alzheimer's disease. *Bilingualism: Language and Cognition, 18*(3), 568–574.
Yang, J., Gates, K. M., Molenaar, P., & Li, P. (2015). Neural changes underlying successful second language word learning: an fMRI study. *Journal of Neurolinguistics, 33*, 29–49.
Yang, S., Yang, H., & Lust, B. (2011). Early childhood bilingualism leads to advances in executive attention: dissociating culture and language. *Bilingualism: Language and Cognition, 14*(3), 412–422.
Yeung, C. M., St. John, P. D., Menec, V., & Tyas, S. L. (2014). Is bilingualism associated with a lower risk of dementia in community-living old adults? Cross-sectional and prospective analyses. *Alzheimer Disease and Associated Disorders, 28*(4), 326–332.
Zahodne, L. B., Schofield, P. W., Farrell, M. T., Stern, Y., & Manly, J. J. (2014). Bilingualism does not alter cognitive decline or dementia risk among Spanish-speaking immigrants. *Neuropsychology, 28*, 238–246.

Received 7 September 2015
Reviewed 9 September 2015
Revised 14 September 2015
Accepted 28 September 2015