CORE
Metadata, citation and similar papers at core.ac.uk
Provided by University of Birmingham Research Portal

UNIVERSITY OF BIRMINGHAM

Research at Birmingham

Data trimming procedure can eliminate bilingual cognitive advantage

Zhou, Beinan; Krott, Andrea

DOI:
10.3758/s13423-015-0981-6

License:
None: All rights reserved

Document Version
Peer reviewed version

Citation for published version (Harvard):
Zhou, B & Krott, A 2015, 'Data trimming procedure can eliminate bilingual cognitive advantage', Psychonomic Bulletin & Review. https://doi.org/10.3758/s13423-015-0981-6

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

• Users may freely distribute the URL that is used to identify this publication.
• Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
• User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
• Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 01. Feb. 2019

Psychonomic Bulletin & Review

Data Trimming Procedure Can Eliminate Bilingual Cognitive Advantage

Journal:	*Psychonomic Bulletin & Review*
Manuscript ID	PBR-BR-15-199.R1
Manuscript Type:	Brief Report
Date Submitted by the Author:	n/a
Complete List of Authors:	Zhou, Beinan; University of Birmingham, Psychology Krott, Andrea; University of Birmingham, Psychology
Keywords:	Bilingual advantage, non-verbal interference tasks, data trimming, response time distribution

SCHOLARONE™
Manuscripts

For Peer Review Only

For Review Only

Data trimming procedure can eliminate bilingual cognitive advantage

Beinan Zhou, Andrea Krott

Beinan Zhou, School of Psychology, University of Birmingham, UK; Andrea Krott, School of Psychology, University of Birmingham, UK.

Correspondence concerning this article should be addressed to Andrea Krott, School of Psychology, University of Birmingham, Birmingham, B15 2TT, UK. Tele: (+44) 121 414 4903. E-mail address: a.krott@bham.ac.uk

Abstract

Bilingualism and its cognitive impacts have drawn increasing interest. Recently, inconsistencies in the findings have raised discussions on what might have caused such discrepancies and how evidence should be evaluated. This review tries to shed new light onto the reasons for the inconsistencies by taking a novel perspective. Motivated by the finding that bilingualism affects response time distribution profiles, particularly findings that suggest bilinguals have fewer long responses, we investigated the relation between maximum response times allowed/included in the analysis of an experiment and the finding of a bilingual advantage. We reviewed 68 experiments from 33 articles that compared monolingual and bilingual speakers' performance in three commonly used non-verbal interference tasks (Simon, Spatial Stroop and Flanker). We found that studies that included longer responses in their analysis were more likely to report a bilingualism effect. We conclude that seemingly insignificant details such as the data trimming procedure can have a potential impact on whether an effect is observed. We also discuss the implication of our findings and suggest the usefulness of more fine-grid analytical procedures.

For Review Only

Recent years have seen an ever-increasing interest in bilingualism, especially in how the bilingual experience leads to cognitive changes (Kroll & Bialystok, 2013). One of the key discoveries was the finding that the two languages of a bilingual speaker are always active, even if the speaker is using only one language in a particular situation (BijeljacBabic, Biardeau, & Grainger, 1997; Colome, 2001; Spivey & Marian, 1999; Wu & Thierry, 2010). Therefore, bilingual experience requires that the speaker constantly monitors and controls language choice. But the soaring interest in bilingualism rather stems from reports that the constant demand on control processes appears to lead to a bilingual cognitive control advantage (for a review see Bialystok, 2009). Possibly the most exciting finding was that the onset of dementia appears to be later for bilingual speakers than for monolingual speakers (Alladi et al., 2013; Bialystok, Craik, & Freedman, 2007), suggesting that bilingualism provides a cognitive reserve. Another reason for the interest in bilingualism is the fact that evidence regarding the bilingual cognitive advantage is inconclusive. Quite a large number of studies have reported a bilingual advantage in cognitive control. But recent developments in the field have suggested that the evidence concerning the bilingual advantage, especially in non-verbal inhibition tasks, is far from conclusive (Hilchey & Klein, 2011; Paap, 2014; Paap & Greenberg, 2013).

For Review Only

Evidence for and against bilingual advantage

Early studies investigating speakers' cognitive control abilities have reported a bilingual advantage. These studies utilized the Simon task (Simon & Rudell, 1967), a task in which participants have to respond to a stimulus feature (e.g. respond according to the colour of a stimulus, left hand for red, and right hand for blue). The position of the stimulus could be either compatible or incompatible with the response hand. In this paradigm, responses are typically slower when stimulus position and response hand are incompatible. Bilingual speakers outperformed monolingual speakers in this task: they showed smaller stimulus congruency effects and/or overall faster response times (Bialystok, Craik, et al., 2005; Bialystok, Craik, Klein, & Viswanathan, 2004; Bialystok, Martin, & Viswanathan, 2005). Using other variants of the Simon task (e.g., Spatial Stroop task, Bialystok, 2006) or paradigms that also require interference inhibition/conflict resolution (e.g., Flanker task, Eriksen & Eriksen, 1974), such bilingual advantage over monolingual performance has been reported many times over the years since the first report (e.g. Costa, Hernandez, Costa-Faidella, & Sebastian-Galles, 2009; Costa, Hernandez, & Sebastian-Galles, 2008; Kapa & Colombo, 2013; Tao, Marzecova, Taft, Asanowicz, & Wodniecka, 2011).

But the evidence is not consistent. While a bilingual advantage is often found, this is not always the case (for a review see Hilchey & Klein, 2011). Most strikingly, in a recent comprehensive study, Paap and Greenberg (2013) did not find any evidence for the bilingual advantage. They conducted a series of non-verbal conflict tasks that had commonly been used in previous studies, including the Simon task, with monolingual and bilingual college students. Bilingual speakers were neither less vulnerable in the conflict condition nor faster overall. On the contrary, the only group difference pointed to a bilingual disadvantage.

For Review Only

These inconsistent findings have raised serious discussions about the nature of the bilingualism effect and thought-provoking debates on how the evidence should be evaluated (e.g. Kroll & Bialystok, 2013; Paap, 2014; Valian, 2015). Several reviews have drawn attention to the published and non-published null results regarding the bilingualism effect, arguing that one should not over-evaluate the significance of positive findings, and should not under-evaluate the meaning of null results. Others discussed methodological issues (Paap, Johnson, & Sawi, 2015), such as the appropriateness of using covariates to control for additional factors, and the convergent validity of tasks used to measure executive control, i.e. the lack of significant correlations between standard measures of inhibition. The series of discussions provided a great chance to reflect on the current status in this research field, and more importantly on where the field is heading. Bearing that in mind, one constructive way to enhance our understanding of the consequences of bilingualism is to understand what factor(s) drives the divergence of results. The focus of this article is to contribute to the discussion from a novel perspective, i.e. the impact of seemingly trivial data trimming procedures.

Factors that potentially drive the inconsistency

In order to shed light onto the reasons for the inconsistencies in the literature, it is important to understand how other factors might interact with speakers' cognitive control ability. Quite a number of factors have been pointed out. It was evident from the beginning that bilingual research is challenging due to the diversity of speakers' linguistic profiles and experiences (Bialystok, 2001; Grosjean, 1998). Depending on their life experience, one bilingual speaker can differ from another one in many ways. Such

For Review Only

heterogeneity in linguistic experiences has been shown to have led to diverse cognitive consequences, such as level of language proficiency (e.g. Mishra, Hilchey, Singh, & Klein, 2012), stage of second language acquisition (early bilingual VS late bilingual, Kalia, Wilbourn, & Ghio, 2014), the degree of bilingualism (dominant VS balanced bilingual, Goral, Campanelli, & Spiro, 2015), pattern of language use, varying experience with frequent language switch (Soveri, Rodriguez-Fornells, & Laine, 2011), the similarity between a bilingual speakers' two languages (Coderre & van Heuven, 2014) and multilingualism (Poarch & van Hell, 2012). In addition, there are factors that are closely related to bilingualism or factors that drive the different language experiences, which at the same time are related to general cognitive performances. These include social and economic status (Morton & Harper, 2007), different cultural backgrounds (Yang, Yang, & Lust, 2011) and immigration status. Last but not least there are factors that affect one's general executive functioning and that probably affect monolingual and bilingual speakers in the same way, such as age, education, exercise, music training, active video game experience and others (for an overview see Valian, 2015). These latter factors emphasize that cognitive control can be trained in other ways than by being bilingual and that the populations of monolinguals and bilinguals can substantially overlap with regards to their performance in cognitive tasks. Such an overlap would also explain why the bilingualism effect has not been found in every study.

For Review Only

A new proposal

Another reason for the inconclusiveness of the literature might be the nature of the bilingual cognitive effect, which is better described as a mixture of effects rather than a single one. For instance, Hilchey and Klein (2011) differentiated between two patterns of bilingual advantage, namely an inhibitory control advantage (i.e. bilinguals showing a reduced conflict effect) and an overall response speed advantage. This distinction suggests two routes through which bilingual experience could affect cognitive control: inhibitory control and attentional control. While enhanced inhibitory control ability should help to resolve conflict, resulting in a reduced conflict effect, enhanced attentional control should help to maintain task goals, leading to an overall speed advantage. Due to the general impurity of cognitive control tasks, a specific task does not provide a pure measure of a single control ability, but draws on many aspects of cognitive control. Some tasks might be more sensitive to participants' inhibitory control ability and some to their attentional control ability. Therefore depending on the task, one might observe a result pattern rather consistent with a bilingual inhibitory control advantage and/or bilingual attentional control advantage.

Tse and Altarriba (2012) utilized a novel analytical approach to the bilingual advantage effect. They performed an ex-Gaussian analysis to investigate response time distributions of bilingual speakers' performance in a Colour Stroop task. Response times in cognitive experiments typically present themselves in a positively skewed distribution, which can be approximated by an ex-Gaussian distribution (Heathcote, Popiel, & Mewhort, 1991), i.e. a convolution of a Gaussian distribution (the mean of which is captured by the parameter μ) and an exponential distribution (the mean and variance of which are captured by the parameter τ). While the Gaussian component (the parameter μ) can be understood as the main body of the distribution, the exponential component (the parameter τ) captures the tail of the distribution, i.e. extremely slow responses. Tse and

For Review Only

Altarriba (2012) suggested that the parameters of ex-Gaussian models of response time distributions in a Colour Stroop experiment are differentially sensitive to inhibitory control and attentional control. They argued that inhibitory control ability modulates the Gaussian component (μ) because differences in inhibitory control would affect the ease one resolves the competition between the two conflicting responses, leading to an overall shift of the response distribution in the conflict condition as compared to the no-conflict condition. In contrast, attentional control ability modulates the tail of the distribution (τ) because lapses of attention should lead to extreme long responses independent of condition. They found that more proficient speakers in L1/L2 showed a smaller interference effect in μ, and also smaller τ independent of condition. This result is important, first because it proposes a way to disentangle the contribution of inhibitory and attentional control, and second because it suggests that what is usually treated as unwanted responses (τ) conveys important information.

One other study that has utilized the ex-Gaussian approach to examine response time distributions is Calabria, Hernandez, Martin, and Costa (2011). They re-analyzed results from an Attentional Network Task (ANT) originally reported in Costa et al. (2008) and Costa et al. (2009). In the original studies, they tested participants' attentional networks using the ANT, which generates measurements for three attentional networks: an alerting network, an orienting network and an executive network. In their re-analysis, they focused on the executive network, or more specifically, on response times for trials with conflict stimuli versus trials with non-conflict stimuli (regardless of cue type). Results revealed an overall speed advantage for bilinguals in both the Gaussian (μ) and exponential (τ) components of the response distributions. Also, for an experiment that

For Review Only

contained only 25% inconsistent trials, i.e. under high monitoring demands, monolinguals had significant longer distribution tails (larger τ) in the incongruent compared to the congruent condition. This congruency effect was absent for bilinguals. These results suggest that the conflict effect in interference tasks is at least partially located in the tail of response distributions.

These observations lead to a new proposal that we investigated in the present study, namely that one reason for observing or not observing the bilingual advantage might be how the data was handled with regards to slow responses. In a traditional central tendency analysis, the typical procedure is to trim extreme responses, treating them as outliers. This is problematic if long responses are most sensitive to the experimental manipulation and/or group differences, as in studies where group/condition differences only emerged in the tail of response distributions (e.g. Epstein et al., 2011; Hervey et al., 2006). In other words, if a difference resides in the tail of the response distribution, by trimming the tail one also trims the potential to observe an effect. For instance, Leth-Steensen, Elbaz, and Douglas (2000) investigated response time distributions in a four-choice reaction time task for a group of children diagnosed with ADHD and a matched group with typical developing children. Using an ex-Gaussian analysis, they found that the two groups' performances differed only in the τ parameter (the distribution tails), not in the main part of the response time distribution. The authors concluded that data trimming in this situation is equivalent to an artificial elimination of effects.

Meta-analysis

For Review Only

In what follows, we present a new perspective on previous studies that compared monolingual and bilingual performance in non-verbal inhibition tasks, investigating their data trimming procedures. If bilingual advantages are at least partly located in the tails of response distributions, i.e. in the slow responses as in Calabria et al. (2011), then one would expect that cutting off slow responses would reduce the chance of finding such advantages. We focused on three non-linguistic inhibition tasks that have been most intensely used to investigate the bilingual advantage: the Simon task, the Spatial Stroop task and the Flanker task (the latter sometimes embedded in an ANT). To ensure comparability, some variations of the tasks were excluded (e.g. the Simon task with a delay component in Martin-Rhee & Bialystok, 2008). We found 68 experiments taken from 33 articles (see Table 1). Within these 68 experiments, 23 (34%) reported data trimming procedures, 4 reported excluding very short responses but did not report trimming of response distribution tails, 1 stated explicitly that long responses were not trimmed and the remaining 40 did not mention whether or not they trimmed the data. For the purpose of our analyses, we treated the latter studies as ones that did not trim the data.

[Table 1 About Here]

Two issues need to be pointed out. First, for studies that did trim the data, the practices differed. While some studies rejected long responses using standard deviations (e.g. 2.5 *SD* in Paap & Greenberg, 2013, which was approximately 700 ms after stimulus onset), others used a specific time cut-off (e.g. response times above 1700 ms). For comparison purposes, we translated cut-offs based on standard deviations into time cut-offs (using reported means and standard deviations). Second, for studies that did not trim the data, there was a big variation in terms of the maximum time allowed for making a

For Review Only

response. For example, in Kousaie and Phillips (2012) 750 ms were allowed for making a response, meaning that in this particular design it was not possible to observe responses slower than 750 ms. This is equivalent to trimming the data at 750 ms. For these reasons, we focused on maximum response times being included in analyses (Figure 1). For studies that did report a data trimming procedure, this is either the explicitly stated cut-off time or the RT calculated by the mean and *SD*. For studies that did not report a data trimming procedure, this was the maximum time allowed for making a response. For simplicity reasons, we will refer to both types of trimming as the maximum response time allowance.

We acknowledge that data trimming and varying maximum time allowances are not the same. Different maximum time allowances, but not data trimming, can lead to different response strategies. For example, participants might not monitor their response accuracy as thoroughly in an experiment with a response allowance of 1000 ms compared to an experiment with a response allowance of 5000 ms and data trimming at 1000 ms. However, in none of the 16 studies in the long allowance group data trimming was applied, which means that this cannot have affected our results.

Some of the studies reviewed were not included into further analysis because there was not adequate information about either the maximum time allowed or how the data was treated (the 3rd study in Bialystok, Martin, et al., 2005, all studies in Gathercole et al., 2014, Yang et al., 2011, and Mohades et al., 2014). Carlson and Meltzoff (2008) was not included either because they did not report RTs. This led to 58 studies being included in our statistical analysis.

[Figure 1 About Here]

For Review Only

Figure 1 shows the number of experiments that did and did not find a bilingual advantage for the various maximum times allowed. A clear pattern emerges: the shorter the maximum response time allowance, the less likely that a bilingualism effect was found. In order to statistically test whether observing a bilingualism effect depends on the maximum time allowed, we grouped the experiments into three types of maximum time allowance: short allowance (below 1000ms), medium allowance (1001ms – 3000ms) and long allowance (above 3001ms; see Table 2). A chi-square test of independence confirmed that the result patterns differed for the three allowance groups, χ^2 (2, N = 58) = 21.99, p <.001. Thus, consistent with our hypothesis, studies with short response time allowance were more likely to report no group difference whereas studies with longer allowance were more likely to report a bilingualism effect.

[Table 2 About Here]

In a meta-analysis of the bilingual advantage literature, Donnelly, Brooks, & Homer (2015) reported an effect of research lab. They suggested that this effect might be due to lab differences in, for instance, access to bilingual populations. In our long allowance group (>3001 ms), many studies are from the same research group. In fact, 14 out of 16 data points are from a research group around one particular author. In order to rule out that the current result is driven by a potential effect of lab, we excluded all data point from this lab. This led to two data points in the long allowance group. We therefore focussed on the short and medium allowance group, which both constituted a mixed contribution from different labs, meaning that the new result could not have been driven by any lab effect. This analysis confirmed our original result. The likelihood of observing a bilingualism effect depended on RT allowance, $\chi 2$ (1, N = 42) = 12.14, p < .001, with

For Review Only

the medium allowance group being more likely to observe a bilingualism effect than the short allowance group.

As introduced, it has been pointed out that the age of the participants might play a role in whether a bilingualism effect can be detected or not. More specifically, bilingual elderly have been found to show the cognitive control advantage more consistently than other age groups. In addition, elderly and children respond on average much slower than adults. It might therefore be that our result arose because the group with long response allowances might have consisted of studies with very young and very old populations. We therefore tested the relationship between maximum response time allowance and the likelihood of finding a bilingualism effect in children, adults and elderly participants separately (see Table 2). A chi-square test of independence showed that while the relationship was not significant for children, χ^2 (2, N = 10) = 2.86, p =.24, it was significant for adults, χ^2 (2, N = 37) = 13.40, p =.001, and was marginally significant for elderly participants, χ^2 (2, N = 11) = 5.29, p =.071. The non-significant result for the children was likely driven by limited power due to small numbers, especially in the short allowance category. Nevertheless, descriptively 83% (5 out of 6) of the studies in the long allowance group versus 66% (2 out of 3) in the medium allowance group showed a bilingualism effect. This result pattern is consistent with the conclusion that studies with longer response allowances are more likely to show a bilingualism advantage. Similarly in the elderly group, studies with longer allowances are more likely to report a bilingualism effect descriptively. The marginal effect seems again be due to a small sample size. In summary, it appears that data trimming reduces the likelihood of observing a bilingualism effect regardless of age group. Also, the very robust finding for

For Review Only

adults showed that our overall result is not driven by the results for children or older participants who tend to respond much slower than adult participants.

Discussion

Based on the assumption that the effect of bilingualism on cognitive control resides at least partly in the tail of response distributions, we investigated a potential relationship between data trimming procedures adopted and the likelihood of observing a bilingualism effect by reviewing 68 experiments reported in 33 articles that compared monolingual and bilingual speakers using non-verbal interference tasks. We found that studies that included longer responses in their analysis were more likely to report a bilingualism effect, either in the form of overall response speed advantage or in the form of reduced interference effect. And this was also the case when the potential effect of lab was eliminated. This is consistent with earlier findings that the bilingualism effect emerges partially in the tail of response distributions (Abutalebi et al., 2015; Calabria et al., 2011; Tse & Altarriba, 2012). It appears that, when these prolonged responses were trimmed or not recorded, group differences might have also been eliminated.

A further analysis showed that the general result pattern was true for studies testing children, adults and elderly alike, even though significantly so only for adults, suggesting that data trimming might be problematic independent of the age of the participants and therefore independent of the average response times or the cognitive abilities of the participants. It also showed that the overall result pattern was not caused by studies with participant groups with very long responses times. There is unfortunately an insufficient number of studies and/or information about the participants in studies

For Review Only

published to date to test whether data trimming procedure affects results independent of other factors that have been suggested to interact with the bilingual advantage (such as SES, immigrant status, language dominance, age of language acquisition, language usage etc.) Future studies will need to disentangle the relevance of and the potential interplay of all these factors.

This review provides some practical implications for future endeavours. Employing a more fine-grid investigation approach might be useful, particularly in situations where effects are subtle. As pointed out, one fruitful alternative to traditional approaches of analysis is the ex-Gaussian analysis of response time distributions (Abutalebi et al., 2015; Calabria et al., 2011). For instance, Abutalebi et al. (2015) reported that a group of bilingual elderly showed advantage in the τ component in the incongruent condition and the μ component in the congruent condition in a Flanker task, supporting the notion that bilingual speakers have enhanced attentional control. It has to be pointed out, though, that one needs to be cautious when interpreting the results from ex-Gaussian parameters in terms of underlying cognitive processes, because there is no one-to-one mapping between cognitive processes and those parameters (Matzke & Wagenmakers, 2009).

An alternative approach to the ex-Gaussian analysis is a delta plot analysis, which examines condition effects as a function of RT and which has been used in analysing conflict tasks such as the Simon task and the Colour Stroop task (Ridderinkhof, 2002a, 2002b; Ridderinkhof, Wildenberg, Wijnen, & Burle, 2004). In delta plot analyses, responses for each condition are grouped into bins according to their response times. Condition differences are then calculated and plotted for these bins. Delta plots

For Review Only

prototypically have a positive slope due to effect sizes being larger for slower responses (Roelofs, Piai, & Garrido Rodriguez, 2011). If one condition requires more inhibition than the other, the difference between conditions does not linearly increase with RT, but is reduced instead, resulting in 'levelling-off' of the delta plot in longer RTs. The levelling-off has been explained by inhibition building up slowly (Ridderinkhof, 2002a). The extent to which the plots level off can effectively reflect the amount of inhibition involved. The stronger the inhibition is applied, the smaller the slope of the plot is. For example, Ridderinkhof (2002b) observed that in a Simon task, delta plots for participants with smaller Simon effect, who are believed to have more efficient inhibitory control, levelled off more than those with larger Simon effect. Because a delta plot analysis is most useful within the discussion of inhibitory control, this approach could be used to test whether inhibition was applied equally fast and to the same degree in monolinguals and bilinguals.

In conclusion, the current review adds to the discussion about the reality of the bilingualism cognitive advantage in that seemingly insignificant details such as data trimming and maximum time allowed for response might have a significant influence on the findings. Therefore, it is important to take these into account in order to fully judge the evidence for and against a bilingual cognitive effect, next to other factor already pointed out in the literature. In addition, future studies are encouraged to report in detail how data were handled and possibly use more fine-grid analyses of RT data to shed light onto the effect of bilingualism on speakers' cognitive control abilities.

For Review Only

Figure 1. Reported bilingual advantage for maximum response times included in the analyses in 58 non-verbal conflict experiments. * indicates that the cut-off time was estimated by the mean and *SD.*

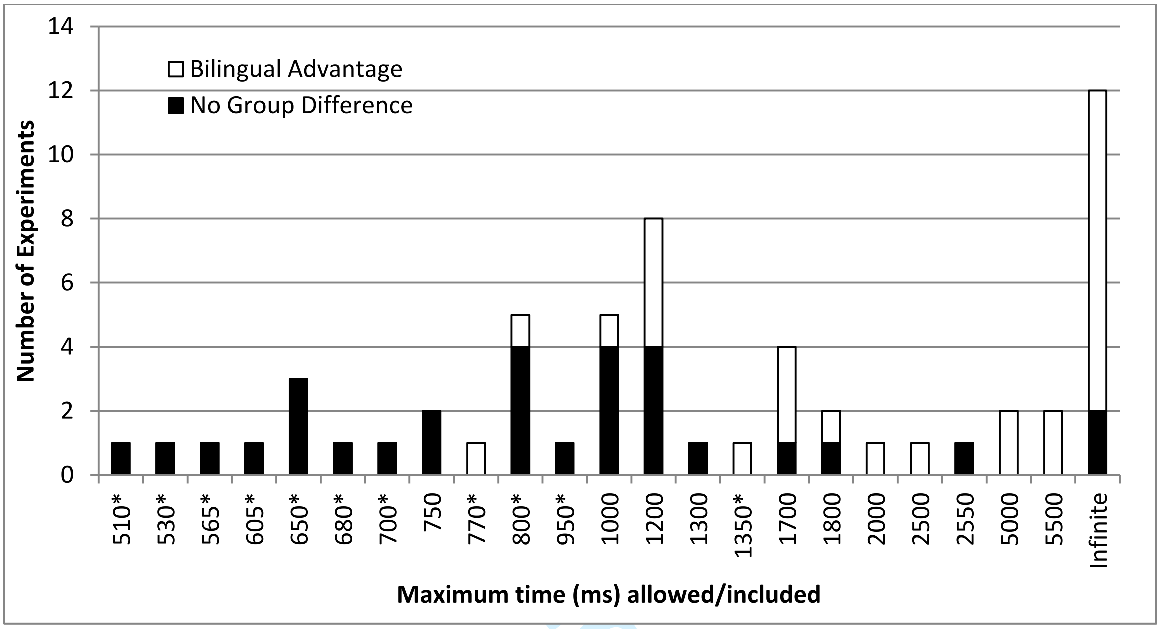

Table 1

Data Trimming Procedures and Results of Studies Using Non-verbal Interference Tasks in Bilingual Research

Study	Age Group (mean age)	Task	Trimming Procedure	Maximum Time Allowed	Results
Bialystok et al. (2004)	Adults (43)	Simon (study1)	-	Until Response	Overall/Effect[1]
	Elderly(71.9)	Simon (study 1)	-	Until Response	Overall/Effect
	Adults (42.6)	Simon (study 2)	-	Until Response	Overall/Effect
		Simon (4 colours)	-	Until Response	Overall/Effect
	Elderly(70.3)	Simon (study 2)	-	Until Response	Overall/Effect
		Simon (4 colours)	-	Until Response	Overall/Effect
	Adults (~39)	Simon	-	Until Response	Overall/Effect
Bialystok, Martin, et al. (2005)[2]	Children (5)	Simon (study 1)	-	5500	Overall
	Children (5)	Simon (study 2)	-	5500	Overall
	Young Adults (20-30)	Simon	-	Not mentioned	No group difference
Bialystok, Craik, et al. (2005)	Young Adults (29)	Simon	-	1800	No group difference
	Young Adults (29)	Simon	-	1800	Overall[3]
Bialystok (2006)	Young Adults (~22)	Simon (low switch)	-	1000	No group difference
		Simon (high switch)	-	1000	No group difference

For Review Only

		Spatial Stroop (low switch)	-	1000	No group difference
		Spatial Stroop (high switch)	-	1000	Overall
Morton and Harper (2007)	Children (6.8)	Simon	None	Until Response[4]	No group difference
Linck, Hoshino, and Kroll (2008)	Young Adults (~20)	Simon	-	Until Response[4]	Effect
Martin-Rhee and Bialystok (2008)[5]	Children (4.6)	Simon (study 1)	-	5000	Overall
	Children (4.2)	Simon (study 2)	-	5000	Overall
	Children (8)	Spatial Stroop	-	Until Response	Overall
Bialystok, Craik, and Luk (2008)	Young Adults (~20)	Spatial Stroop (high switch)	-	Until Response	No group difference
	Elderly(~68)	Spatial Stroop (high switch)	-	Until Response	Effect
Carlson and Meltzoff (2008)	Children (~6)	ANT	> 1700 ms	Not mentioned	N/A[6]
Costa et al. (2008)	Young Adults (22)	ANT	-	1700	Overall/Effect
Costa et al. (2009)	Young Adults (~20)	ANT (92% cong)	>1200 ms	1700	No group difference
	Young Adults (~20)	ANT (8% cong)	>1200 ms	1700	No group difference
	Young Adults (~20)	ANT (50% cong)	>1200 ms	1700	Overall
	Young Adults (~20)	ANT (75% cong)	>1200 ms	1700	Overall/Effect (Block1)
Emmorey, Luk, Pyers, and Bialystok (2008)	Adults (46-50)	Flanker	none	2000	Overall (for unimodal bilingual, not bimodal)
Bialystok and DePape (2009)	Young Adults (~23)	Spatial Stroop	>2500 ms	Until Response	Overall (for no music

For Review Only

					experience group)
Luk, Anderson, Craik, Grady, and Bialystok (2010)	Young Adutls (~21)	Flanker (with Go/No-Go element)	-	1300[7]	No group difference
Tao et al. (2011)	Young Adults (~20)	Lateralized ANT	>1200 ms	1820	Effect
Yang et al. (2011)	Children (~6.5)	ANT	-	Not mentioned	Overall
	Children (~6.5)	ANT	-	Not mentioned	Overall
	Children (~6.5)	ANT	-	Not mentioned	Overall
Salvatierra and Rosselli (2011)	Young Adults (~26)	Simon	-	800	No group difference
	Young Adults (~26)	Simon (4 colours)	-	800	No group difference
	Elderly (~64)	Simon	-	800	Effect
	Elderly (~64)	Simon (4 colours)	-	800	No group difference
Yudes et al. (2011)	Young Adults (21-25)	Simon	>1200 ms	2000	No group difference[8]
Engel de Abreu, Cruz-Santos, Tourinho, Martin, and Bialystok (2012)	Children (~8.5)	Flanker	>3 *SD*	5000	Overall/Effect
Kousaie and Phillips (2012)	Young Adults (~24)	Simon	-	750	No group difference
	Young Adults (~24)	Flanker	-	750	No group difference
Poarch and van Hell (2012)	Children (~7)	Simon (study1)	>2.5 *SD*	1000	No group difference[9]

For Review Only

Kapa and Colombo (2013)	Children (~9)	ANT	None	1700	Overall (Monolingual = Late bilingual > Early bilingual)
Nicolay and Poncelet (2013)	Children (~8.5)	Flanker (50% cong)	-	1700[10]	No group difference
Paap and Greensberg (2013)	Young Adults 1 (students)	Simon1	>2.5 *SD*	Not mentioned	No group difference
	Young Adults 2 (students)	Simon2	>2.5 *SD*	Not mentioned	No group difference
	Young Adults 3 (students)	Simon3	>2.5 *SD*	Not mentioned	No group difference
	Young Adults 3 (students)	Flanker	>2.5 *SD*	1700	No group difference
Kirk, Scott-Brown, and Kempe (2014)	Elderly (~70.3)	Simon	>2.5 *SD*	1000	No group difference
			None[11]	1000	No group difference
Marzecova et al. (2013)	Young Adults (~21)	LANT	>1200 ms	2000	Effect
Gathercole et al. (2014)	Children (~4)	Simon	-	Not mentioned	Monolingual advantage
	Grade Schoolers (~8)	Simon	-	Not mentioned	No group difference

For Review Only

	Young Adults (~25)	Simon	-	Not mentioned	Monolingual advantage
	Elderly (~67)	Simon	-	Not mentioned	No group difference
Mohades et al. (2014)	Children (8~11)	Simon	-	Not mentioned	Monolingual advantage
Blumenfeld and Marian (2014)[12]	Young Adults (~22)	Spatial Stroop (study1)	>2.5 *SD*	700	No group difference
	Young Adults (~22)	Simon (up/down arrow in study1)	>2.5 *SD*	700	No group difference
	Young Adults (~22)	Spatial Stroop (study2)	>2.5 *SD*	700	No group difference
	Young Adults (~22)	Simon (up/down arrow in study2)	>2.5 *SD*	700	No group difference
Pelham and Abrams (2014)	Young Adults (~20)	ANT	>2.5 *SD*	1700	Effect (Monolingual > Late bilingual = Early bilingual)
Kousaie, Sheppard, Lemieux, Monetta, and Taler (2014)	Young Adults (~21.5)	Spatial Stroop	>2.5 *SD*	Not mentioned	No group difference
	Elderly (~72.5)	Spatial Stroop	>2.5 *SD*	Not mentioned	No group difference
Grady, Luk, Craik, and Bialystok (2015)	Elderly (~70.5)	Simon	-	2550	No group difference
Abutalebi et al. (2015)	Elderly (~62)	Flanker	None	1700	Overall

For Review Only

Note: 1. Overall = Bilingual advantage in overall response speed (in both congruent and incongruent condition). Effect = Smaller congruency effect for bilinguals. 2. Study 4 and 5 in this article were the same as in Bialystok (2004); therefore we do not duplicate report here. 3. Bilingual overall advantage was found in the Cantonese/English bilingual group but not the French/English bilingual group; this is represented as two separate comparisons here. 4. Not stated directly, but followed the procedure by Bialystok et al. (2004). 5. Another two variations of the Simon task were conducted, with a short or a long delay. These were not included here. 6. A bilingual advantage was reported for response accuracies. Information regarding response speed is not available. 7. Assuming that responses can be made during the blank interval of 300 ms after the stimulus presentation for 1000 ms. 8. Yudes et al. (2011) also tested a group of simultaneous interpreters. We only report the monolingual/bilingual comparison. 9. A group of second language learners and a group of trilingual children were also tested. We only report the monolingual/bilingual comparison. 10. Not stated directly, but followed the procedure by Costa et al. (2008). 11. Kirk (2013) tested three monolingual control groups. We only report results regarding Angalo-English monolingual group which does not speak another dialect. Kirk (2013) also reported analysis without removing outlying responses and found no group difference. 12. Blumenfeld & Viorica (2014) also reported reverse efficiency scores which were based on both response latencies and accuracy rates. This analysis yielded slightly different result patterns. We focus on response latencies for reasons of comparability.

For Review Only

Table 2

Number of Experiments with Short, Medium or Long Cut-offs / Maximum Allowed Times and Findings of Bilingual Advantage

Cut-off Group	Children		Adults		Elderly		Overall	
	No Difference	Bilingual advantage	No Difference	Bilingual advantage	No Difference	Bilingual advantage	No Difference	Bilingual advantage
Short Allowance	1	0	16	2	3	1	20	3
Medium Allowance	1	2	5	8	2	1	8	11
Long Allowance	1	5	1	5	0	4	2	14
Column Total	3	7	22	15	5	5	30	28

For Review Only

References

Abutalebi, J., Guidi, L., Borsa, V., Canini, M., Della Rosa, P. A., Parris, B. A., & Weekes, B. S. (2015). Bilingualism provides a neural reserve for aging populations. *Neuropsychologia, 69*, 201-210. doi: 10.1016/j.neuropsychologia.2015.01.040

Alladi, S., Bak, T. H., Duggirala, V., Surampudi, B., Shailaja, M., Shukla, A. K., . . . Kaul, S. (2013). Bilingualism delays age at onset of dementia, independent of education and immigration status. *Neurology, 81*(22), 1938-1944.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition.* : Cambridge, UK: Cambridge University Press.

Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Canadian Journal of Experimental Psychology-Revue Canadienne de Psychologie Experimentale, 60*(1), 68-79.

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology-Learning Memory and Cognition, 34*(4), 859-873.

Bialystok, E., Craik, F. I. M., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia, 45*(2), 459-464. doi: 10.1016/j.neuropsychologia.2006.10.009

Bialystok, E., Craik, F. I. M., Grady, C., Chau, W., Ishii, R., Gunji, A., & Pantev, C. (2005). Effect of bilingualism on cognitive control in the Simon task: evidence from MEG. *NeuroImage, 24*(1), 40-49.

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

For Review Only

Bialystok, E., & DePape, A.-M. (2009). Musical Expertise, Bilingualism, and Executive Functioning. *Journal of Experimental Psychology-Human Perception and Performance, 35*(2), 565-574. doi: 10.1037/a0012735

Bialystok, E., Martin, M. M., & Viswanathan, M. (2005). Bilingualism across the lifespan: The rise and fall of inhibitory control. *International Journal of Bilingualism, 9*(1), 103-119. doi: 10.1177/13670069050090010701

BijeljacBabic, R., Biardeau, A., & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory & Cognition, 25*(4), 447-457.

Blumenfeld, H. K., & Marian, V. (2014). Cognitive control in bilinguals: Advantages in Stimulus-Stimulus inhibition. *Bilingualism-Language and Cognition, 17*(3), 610-629. doi: 10.1017/s1366728913000564

Calabria, M., Hernandez, M., Martin, C. D., & Costa, A. (2011). When the tail counts: the advantage of bilingualism through the ex-Gaussian distribution analysis. *Frontiers in Psychology, 2*, 8. doi: 10.3389/fpsyg.2011.00250

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science, 11*(2), 282-298. doi: 10.1111/j.1467-7687.2008.00675.x

Coderre, E. L., & van Heuven, W. J. B. (2014). The effect of script similarity on executive control in bilinguals. *Frontiers in Psychology, 5*. doi: 10.3389/fpsyg.2014.01070

Colome, À. (2001). Lexical Activation in Bilinguals' Speech Production: Language-Specific or Language-Independent? *Journal of Memory and Language, 45*(4), 721-736.

Costa, A., Hernandez, M., Costa-Faidella, J., & Sebastian-Galles, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition, 113*(2), 135-149.

Costa, A., Hernandez, M., & Sebastian-Galles, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition, 106*(1), 59-86.

For Review Only

de Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive Advantage in Bilingualism: An Example of Publication Bias? *Psychological Science, 26*(1), 99-107. doi: 10.1177/0956797614557866

Donnelly, S., Brooks. P.J., & Homer, B.D. (2015). Examining the Bilingual Advantage on Conflict Resolution Tasks: A Meta-Analysis. *The Annual Meeting of the Cognitive Science Society*.

Emmorey, K., Luk, G., Pyers, J. E., & Bialystok, E. (2008). The Source of Enhanced Cognitive Control in Bilinguals: Evidence From Bimodal Bilinguals. *Psychological Science, 19*(12), 1201-1206. doi: 10.1111/j.1467-9280.2008.02224.x

Engel de Abreu, P., Cruz-Santos, A., Tourinho, C. J., Martin, R., & Bialystok, E. (2012). Bilingualism Enriches the Poor: Enhanced Cognitive Control in Low-Income Minority Children. *Psychological Science, 23*(11), 1364-1371. doi: 10.1177/0956797612443836

Epstein, J. N., Langberg, J. M., Rosen, P. J., Graham, A., Narad, M. E., Antonini, T. N., . . . Altaye, M. (2011). Evidence for Higher Reaction Time Variability for Children With ADHD on a Range of Cognitive Tasks Including Reward and Event Rate Manipulations. *Neuropsychology, 25*(4), 427-441. doi: 10.1037/a0022155

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of Noise Letters Upon Identification of A Target Letter in A Nonsearch Task. *Perception & Psychophysics, 16*(1), 143-149.

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

Goral, M., Campanelli, L., & Spiro, A. (2015). Language dominance and inhibition abilities in bilingual older adults. *Bilingualism-Language and Cognition, 18*(1), 79-89. doi: 10.1017/s1366728913000126

For Review Only

Grady, C. L., Luk, G., Craik, F. I. M., & Bialystok, E. (2015). Brain network activity in monolingual and bilingual older adults. *Neuropsychologia, 66*, 170-181. doi: 10.1016/j.neuropsychologia.2014.10.042

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition, 1*(02), 131-149. doi: doi:10.1017/S136672899800025X

Heathcote, A., Popiel, S. J., & Mewhort, D. J. K. (1991). Analysis of Response - Time Distributions - An Example Using the Stroop Task. *Psychological Bulletin, 109*(2), 340-347. doi: 10.1037/0033-2909.109.2.340

Hervey, A. S., Epstein, J. N., Curry, J. F., Tonev, S., Arnold, L. E., Conners, C. K., . . . Hechtman, L. (2006). Reaction time distribution analysis of neuropsychological performance in an ADHD sample. *Child Neuropsychology, 12*(2), 125-140. doi: 10.1080/09297040500499081

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review, 18*(4), 625-658.

Kalia, V., Wilbourn, M. P., & Ghio, K. (2014). Better early or late? Examining the influence of age of exposure and language proficiency on executive function in early and late bilinguals. *Journal of Cognitive Psychology, 26*(7), 699-713. doi: 10.1080/20445911.2014.956748

Kapa, L. L., & Colombo, J. (2013). Attentional control in early and later bilingual children. *Cognitive Development, 28*(3), 233-246. doi: 10.1016/j.cogdev.2013.01.011

Kirk, N. W., Fiala, L., Scott-Brown, K. C., & Kempe, V. (2014). No evidence for reduced Simon cost in elderly bilinguals and bidialectals. *Journal of Cognitive Psychology (Hove, England), 26*(6), 640–648. http://doi.org/10.1080/20445911.2014.929580

For Review Only

Kousaie, S., & Phillips, N. A. (2012). Conflict monitoring and resolution: Are two languages better than one? Evidence from reaction time and event-related brain potentials. *Brain Research, 1446*, 71-90. doi: 10.1016/j.brainres.2012.01.052

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*. doi: 10.3389/fnbeh.2014.00250

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology, 25*(5), 497-514. doi: 10.1080/20445911.2013.799170

Leth-Steensen, C., Elbaz, Z. K., & Douglas, V. I. (2000). Mean response times, variability, and skew in the responding of ADHD children: a response time distributional approach. *Acta Psychologica, 104*(2), 167-190. doi: 10.1016/s0001-6918(00)00019-6

Linck, J. A., Hoshino, N., & Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *Ment Lex, 3*(3), 349-374. doi: 10.1075/ml.3.3.06lin

Luk, G., Anderson, J. A. E., Craik, F. I. M., Grady, C., & Bialystok, E. (2010). Distinct neural correlates for two types of inhibition in bilinguals: Response inhibition versus interference suppression. *Brain and Cognition, 74*(3), 347-357. doi: 10.1016/j.bandc.2010.09.004

Martin-Rhee, M. M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism-Language and Cognition, 11*(1), 81-93.

Marzecova, A., Asanowicz, D., Kriva, L., & Wodniecka, Z. (2013). The effects of bilingualism on efficiency and lateralization of attentional networks. *Bilingualism-Language and Cognition, 16*(3), 608-623. doi: 10.1017/s1366728912000569

For Review Only

Matzke, D., & Wagenmakers, E. J. (2009). Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review, 16*(5), 798-817. doi: 10.3758/pbr.16.5.798

Mezzacappa, E. (2004). Alerting, orienting, and executive attention: Developmental properties and sociodemographic correlates in an epidemiological sample of young, urban children. *Child Development, 75*(5), 1373-1386. doi: 10.1111/j.1467-8624.2004.00746.x

Mishra, R. K., Hilchey, M. D., Singh, N., & Klein, R. M. (2012). On the time course of exogenous cueing effects in bilinguals: Higher proficiency in a second language is associated with more rapid endogenous disengagement. *Quarterly Journal of Experimental Psychology, 65*(8), 1502-1510. doi: 10.1080/17470218.2012.657656

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J. C. (2004). Inner speech as a retrieval aid for task goals: the effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica, 115*(2-3), 123-142. doi: 10.1016/j.actpsy.2003.12.004

Mohades, S. G., Struys, E., Van Schuerbeek, P., Baeken, C., Van De Craen, P., & Luypaert, R. (2014). Age of second language acquisition affects nonverbal conflict processing in children: an fMRI study. *Brain and Behavior, 4*(5), 626-642. doi: 10.1002/brb3.246

Morton, J. B., & Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental Science, 10*(6), 719-726.

Nicolay, A. C., & Poncelet, M. (2013). Cognitive advantage in children enrolled in a second-language immersion elementary school program for three years. *Bilingualism-Language and Cognition, 16*(3), 597-607. doi: 10.1017/s1366728912000375

Paap, K. R. (2014). The role of componential analysis, categorical hypothesising, replicability and confirmation bias in testing for bilingual advantages in executive functioning. *Journal of Cognitive Psychology, 26*(3), 242-255. doi: 10.1080/20445911.2014.891597

For Review Only

Paap, K. R., Darrow, J., Dalibar, C., & Johnson, H. A. (2015). Effects of script similarity on bilingual advantages in executive control are likely to be negligible or null. *Frontiers in Psychology, 5*. doi: 10.3389/fpsyg.2014.01539

Paap, K.P., 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, 69, 265-278

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology, 66*(2), 232-258. doi: 10.1016/j.cogpsych.2012.12.002

Pelham, S. D., & Abrams, L. (2014). Cognitive Advantages and Disadvantages in Early and Late Bilinguals. *Journal of Experimental Psychology-Learning Memory and Cognition, 40*(2), 313-325. doi: 10.1037/a0035224

Poarch, G. J., & van Hell, J. G. (2012). Executive functions and inhibitory control in multilingual children: Evidence from second-language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology, 113*(4), 535-551. doi: 10.1016/j.jecp.2012.06.013

Ridderinkhof, K. R. (2002a). *Activation and suppression in conflict tasks: empirical clarification through distributional analyses*.

Ridderinkhof, K. R. (2002b). Micro- and macro-adjustments of task set: activation and suppression in conflict tasks. *Psychological Research-Psychologische Forschung, 66*(4), 312-323.

Ridderinkhof, K. R., Wildenberg, v., Wijnen, J., & Burle, B. s. (2004). *Response Inhibition in Conflict Tasks Is Revealed in Delta Plots*. Paper presented at the British Journal of Mathematical & Statistical Psychology.

For Review Only

Roelofs, A., Piai, V. r., & Garrido Rodriguez, G. (2011). Attentional inhibition in bilingual naming performance: Evidence from delta-plot analyses. *Frontiers in Psychology, 2*. doi: 10.3389/fpsyg.2011.00184

Salvatierra, J. L., & Rosselli, M. (2011). The effect of bilingualism and age on inhibitory control. *International Journal of Bilingualism, 15*(1), 26-37. doi: 10.1177/1367006910371021

Simon, J. R., & Rudell, A. P. (1967). Auditory S-R Compatibility - Effect of An Irrelevant Cue on Information Processing. *Journal of Applied Psychology, 51*(3), 300-304.

Soveri, A., Rodriguez-Fornells, A., & Laine, M. (2011). Is there a relationship between language switching and executive functions in bilingualism? Introducing a within-group analysis approach. *Frontiers in Psychology, 2*, 8. doi: 10.3389/fpsyg.2011.00183

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science, 10*(3), 281-284.

Tao, L., Marzecova, A., Taft, M., Asanowicz, D., & Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: the role of age of acquisition. *Frontiers in Psychology, 2*, 19. doi: 10.3389/fpsyg.2011.00123

Tse, C. S., & Altarriba, J. (2012). The effects of first- and second-language proficiency on conflict resolution and goal maintenance in bilinguals: Evidence from reaction time distributional analyses in a Stroop task. *Bilingualism-Language and Cognition, 15*(3), 663-676. doi: 10.1017/s1366728912000077

Valian, V. (2015). Bilingualism and cognition. *Bilingualism-Language and Cognition, 18*(1), 3-24. doi: 10.1017/s1366728914000522

Wu, Y.J.,& Thierry, G. (2010). Chinese-English bilinguals reading English hear Chinese. *Journal of Neuroscience.* **30** (22), 7646-7651.

Yang, S. J., Yang, H. J., & Lust, B. (2011). Early childhood bilingualism leads to advances in executive attention: Dissociating culture and language. *Bilingualism-Language and Cognition, 14*(3), 412-422. doi: 10.1017/s1366728910000611

For Review Only

Yudes, C., Macizo, P., & Bajo, T. (2011). The influence of expertise in simultaneous interpreting on non-verbal executive processes. *Frontiers in Psychology*, *2*, 9. doi: 10.3389/fpsyg.2011.00309

For Review Only