Sophia Czapka | Christiane Wotschack | Annegret Klassert |
Julia Festman

A path to the bilingual advantage

Pairwise matching of individuals

Research Article

CAMBRIDGE
UNIVERSITY PRESS

A path to the bilingual advantage: Pairwise matching of individuals

Sophia Czapka[1], Christiane Wotschack[1], Annegret Klassert[1] and Julia Festman[1,2]

[1]University of Potsdam, Human Sciences Faculty, Research group: Diversity and Inclusion, Potsdam, Germany and [2]Pedagogical University Tyrol, Innsbruck, Austria

Abstract

Matching participants (as suggested by Hope, 2015) may be one promising option for research on a potential bilingual advantage in executive functions (EF). In this study we first compared performances in three EF-tasks of a naturally heterogeneous sample of monolingual (n = 69, age = 9.0 y) and multilingual children (n = 57, age = 9.3 y). Secondly, we meticulously matched participants pairwise to obtain two highly homogeneous groups to rerun our analysis and investigate a potential bilingual advantage. The initally disadvantaged multilinguals (regarding socioeconomic status and German lexicon size) performed worse in updating and response inhibition, but similarly in interference inhibition. This indicates that superior EF compensate for the detrimental effects of the background variables. After matching children pairwise on age, gender, intelligence, socioeconomic status and German lexicon size, performances became similar except for interference inhibition. Here, an advantage for multilinguals in the form of globally reduced reaction times emerged, indicating a bilingual executive processing advantage.

Introduction

The probably most highly debated issue in bilingualism research to date is the possible enhancement of executive functions (EF) through bilingualism (Bialystok, 2015; Paap, Johnson & Sawi, 2015). From the late 1990s onwards, many studies reported an advantage of bilinguals over monolinguals in EF tasks (for meta-analyses, see Adesope, Lavin, Thompson & Ungerleider, 2010; de Bruin, Treccani & Della Sala, 2014), but recent findings invoke doubt about the existence of the BILINGUAL ADVANTAGE (for more detail, see Paap et al., 2015). With this study we pursue two goals. First, we compare the EF performance of mono- and multilinguals taken from a large, naturally heterogeneous sample of primary school children. By accepting the multilinguals' disadvantageous position in terms of socioeconomic status (SES) and lexicon size, we can explore the children's abilities as they can be observed in the German school context. Secondly, we try to illustrate the importance of meticulous pairwise matching of participants when comparing multilinguals with monolinguals. Only by pairwise matching can we control for the influence of many potentially influential or confounding variables by creating experimental groups that are comparable on a group and on an individual level.

The bilingual advantage in EF

One explanation for the EF advantage of bilinguals is the continuous, parallel activation of their languages: The inactive language needs to be inhibited to avoid intrusion errors, while the language currently in use needs to be activated by directing attention towards it (Green, 1998; Kroll, Dussias, Bogulski & Kroff, 2012). This constant language control puts high demands on the EF system and therefore constitutes a training, which is suggested to yield the bilingual advantage in EF (Bialystok, 2015). Since the development of EF starts in early childhood, investigating the bilingual advantage in primary school children is particularly interesting, as their EF still undergo refinement in terms of processing speed and accuracy (Best & Miller, 2010).

Some reviews and meta-analyses support the existence of the bilingual advantage in EF, despite significant variability in the data (Adesope et al., 2010; Barac & Bialystok, 2011; de Bruin et al., 2014; Hilchey & Klein, 2011). EF comprise a family of cognitive control mechanisms with three core functions: inhibition, working memory or updating, and shifting (Diamond, 2013; Miyake & Friedman, 2012). The strongest evidence for the bilingual advantage exists for one subcomponent of inhibition, namely interference inhibition, but not for response inhibition. Tasks assessing response inhibition require control to react to a target stimulus but to withhold a reaction upon presentation of a non-target stimulus. There is hardly

any evidence for a bilingual advantage for this EF component (Esposito, Baker-Ward & Mueller, 2013; Martin-Rhee & Bialystok, 2008; but Bialystok, Barac, Blaye & Poulin-Dubois, 2010).

Bilinguals outperform monolinguals in interference inhibition tasks, since these probably involve the same control mechanism as in bilingual language control (as described above). Besides simple, congruent conditions, they comprise incongruent conditions with additional inhibitory demands. In these conditions, conflicting stimulus information requires focusing on the relevant dimension while inhibiting interference from the irrelevant dimension. Experimental paradigms supporting the bilingual advantage are, for example, the Simon task (Martin-Rhee & Bialystok, 2008; Poarch & van Hell, 2012), Stroop task (Costa, Fuentes & Vivas, 2010) and Flanker task (Costa, Hernandez, Costa-Faidella & Sebastián-Gallés, 2009).

Depending on the result pattern in interference inhibition tasks, Hilchey, Saint-Aubin and Klein (2015) either assume a BILINGUAL INHIBITORY CONTROL ADVANTAGE (BICA) or a BILINGUAL EXECUTIVE PROCESSING ADVANTAGE (BEPA). A superior performance in conditions with additional inhibitory demands (measured by the interference effect, the difference between incongruent and congruent conditions) indicates a BICA. However, since many studies report globally faster reaction times (RT), Hilchey et al. (2015) consider BEPA the predominant finding in research on the bilingual advantage. On this account, also Bialystok (2017) supports the change of an advantage solely in inhibition to a more general advantage in EXECUTIVE ATTENTION. Here, monitoring of attention is additionally involved to detect potential conflicts, which plays an important role in the afore-mentioned interference inhibition tasks.

There is also evidence that bilingualism improves updating (Blom, Küntay, Messer, Verhagen & Leseman, 2014; Luo, Craik, Moreno & Bialystok, 2013), but results are equivocal (Engel de Abreu, 2011; Namazi & Thordardottir, 2010).

Doubt about the bilingual advantage

Some researchers argue that the bilingual advantage is "restricted to very specific and undetermined circumstances" (Paap et al., 2015, p. 1). Doubt stems from an initial publication bias for positive results (de Bruin et al., 2014) and the fact that after 2011, more than 80% of the published studies reported null results (Paap et al., 2015). Moreover, studies with large-scale designs and carefully matched groups could not replicate seminal findings (Duñabeitia, Hernández, Antón, Macizo, Estévez, Fuentes & Carreiras, 2014; Paap, Johnson & Sawi, 2014) and the bilingual advantage was detected especially in studies with small sample sizes (n < 30) what increases chances of false positive results (Paap et al., 2015). However, Hope (2015) suggested, "by reducing samples to ensure that groups are more comparable, (…) smaller studies can be more informative than larger studies, if they are better controlled" (p. 59).

The above-mentioned null-results can either indicate the non-existence of a bilingual advantage or result from experimental manipulations. For example, participants' age seems to affect the outcome, since behavioral effects are predominantly found in children and older adults but not in young adults (Antón, Duñabeitia, Estévez, Hernández, Castillo, Fuentes, Davidson & Carreiras, 2014). The reason is probably that young adults are at the height of their cognitive performance and bilingualism does not constitute an additional boost to EF (Bialystok, Martin & Viswanathan, 2005). Moreover, task demands affect behavioral results, because only conditions with high cognitive demands reveal a bilingual advantage (Costa et al., 2009; Martin-Rhee & Bialystok, 2008). This fact applies also to EF training studies in general, since intervention and control groups usually differ in the cognitively most demanding conditions (Diamond, 2014).

Another major concern is the uncontrolled influence of confounding factors. SES is in the center of this discussion as it correlates with EF (Carlson, Zelazo & Faja, 2013). Consequently, SES needs to be controlled for when comparing multilinguals with monolinguals to avoid skewed results (Calvo & Bialystok, 2014; Morton & Harper, 2007). In certain contexts, multilingualism is even accompanied with lower SES. This is the case in Germany, where multilingual children stem more often from families with a lower SES, referring to both income and educational level (Bildungsberichterstattung, 2014; for other settings, see Engel de Abreu, Cruz-Santos, Tourinho, Martin & Bialystok, 2012; Paap & Greenberg, 2013). In Germany, multilingualism also frequently coincides with migration background (Leerhoff, Rehkämper, Rockmann, Brunner, Gärtner & Wendt, 2013), but if and how migration status interacts with EF remains unclear, as empirical evidence is scarce. Another factor often differing between mono- and multilinguals is the size of the mental lexicon, which is smaller for bilinguals when comparing each language separately (e.g., Bialystok & Luk, 2012). Moreover, multilingual children residing in Germany (especially those with a migration status) show on average lower language proficiency in German (Niklas, Schmiedeler, Pröstler & Schneider, 2011). Therefore, test conditions including long and complex verbal instructions and tasks based on verbal stimuli or culture-specific knowledge may put multilingual children at a disadvantage (Hagmann-Von Arx, Petermann & Grob, 2013).

Pairwise matching

Friedman (2016, p. 2) stated as follows: "This bilingual advantage hypothesis seems straightforward to test: Obtain a sample of bilinguals and appropriately matched controls and test whether they differ on a measure of EF". However, finding the ´right` bilingual sample for a study poses a problem, since multilingualism is a complex and continuous variable itself comprising aspects like language dominance and proficiency, age of acquisition (AoA) etc. (Luk & Bialystok, 2013). Matching participants poses a great challenge, too, because numerous characteristics of an individual potentially influence EF (Diamond, 2013), and multilingualism is just one of them. In many studies participants were MATCHED ON A GROUP-LEVEL. This was done either by controlling that mono- and multilingual groups do not differ significantly on a certain number of background variables (e.g., Bialystok & Viswanathan, 2009; Kousaie, Laliberté, López Zunini, Taler, Zunini & Taler, 2015; Scaltritti, Peressotti & Miozzo, 2017) or by excluding individuals of a larger sample to reach comparable groups (Bak, Long, Vega-Mendoza & Sorace, 2016; Stocco & Prat, 2014). This matching method at the group-level is used in studies examining small samples sizes with n < 30 (e.g., Kousaie et al., 2015; Morton & Harper, 2007), as well as large samples with n > 100 (Antón et al., 2014; Duñabeitia et al., 2014; Paap & Sawi, 2014). This approach is always applied at a group or grade level, but it ignores the complex set of characteristics on an individual level. In this paper, we demonstrate how to overcome this problem by MATCHING INDIVIDUAL PAIRS of participants on a number of background variables.

Research aim

The discussion on the existence of the bilingual advantage in EF is still vivid and we approached this problem by studying a natural, unselected group of mono- and multilingual children from Germany. Our goal was twofold. First, we compared the children's performance in EF to investigate their abilities in a natural setting. Secondly, we demonstrated how to use pairwise matching on an individual level to investigate the existence of the bilingual advantage in EF. For that purpose, we applied a two-step analysis, i.e., without and with pairwise matching of participants.

We collected data from a larger cohort of third graders in Germany, who were categorized post-hoc as mono- or multilingual (according to reported home language use; for details, see Methods section). In our first analysis without pairwise matching, we compared performance in EF tasks between these large, unmatched groups (further referred to as HETEROGENEOUS GROUPS) accepting all natural group differences in background variables. We hypothesized that if multilingualism does not influence EF, multilingual children should underperform throughout all tasks due to their anticipated disadvantage in the background variables. In a second step, we matched participants pairwise on important influential variables known from the literature (i.e., age, gender, intelligence, lexicon size and SES), creating two HOMOGENEOUS GROUPS and compared their EF-performance. We hypothesized that if a bilingual advantage exists, it should arise in this step, because detrimental effects from confounding variables are eliminated.

Since we have specific hypotheses about the nature of the bilingual advantage, our study comprises three tasks testing the following EF components: interference inhibition, response inhibition and updating. We predict that multilinguals outperform monolinguals in interference but not in response inhibition (Esposito et al., 2013; Martin-Rhee & Bialystok, 2008). This advantage would manifest itself either in lower global RT – indicating a BEPA – or in a lower interference effect – indicating a BICA. Furthermore, we expect an advantage in the more demanding conditions (e.g., of the N-back task measuring updating, or of the interference inhibition task when having to switch between conditions in the mixed block), since only high cognitive demands allow benefits of bilinguals to emerge.

Method

Participants

In total, 168 children who attended third grade in different schools in Germany took part in the study. They were categorized as monolinguals or multilinguals based on their home language use in the following way: the monolingual group (n = 69; 36 female) included children who spoke only German at home and had no further contact with another language. The multilingual group consisted of bilingual (n = 50; 24 female) and trilingual children (n = 7; 3 female). To be included in the multilingual group, the child's family needed to use at least one other language at home besides German, and the child's verbal proficiency in this language had to be at least good (for more detail see section Questionnaire below and Table 1). We will refer to the home language as L1, because for successive multilinguals it is their L1 (their acquisition of German starts only in kindergarten and German is therefore their L2/L3); simultaneous multilinguals have two/three first languages, hence their home language is one of these L1s. We suggest that these selection criteria ensure

Table 1. Descriptive statistics for heterogeneous groups

	Monolinguals		Multilinguals		Sig.
n / females	69 / 48%		57 / 53%		
Age (in months)	107.6	(4.8)	111.0	(9.0)	*
Intelligence	35.3	(5.5)	35.3	(5.0)	
Short-term memory	15.4	(3.4)	15.2	(3.5)	
SES (mother's ISCED)	3.8	(1.0)	3.1	(1.5)	***
SES (father's ISCED)	3.8	(1.0)	3.2	(1.6)	**
German expressive lexicon	25.1	(6.1)	16.3	(8.4)	***
German receptive lexicon	38.1	(2.2)	33.8	(4.9)	***
Self-rated German proficiency	3.8	(0.4)	3.6	(0.6)	*
Self-rated L1 proficiency			2.9	(0.8)	
AoA German (in months)			27.6	(30.8)	

Note. Means, standard deviations (in brackets) and group differences between monolinguals and multilinguals in background variables (Sig. - Significance *** $p < .001$, ** $p < .01$, * $p < .05$)

an active and continuous use of L1, which is probably necessary for the bilingual advantage to develop in this age group. Consequently, we excluded 42 children from further analyses. These comprise three children with an IQ < 70 indicating an intellectual disability (DIMDI, n.d.) and 39 children who could not be assigned to either language group. Exclusion criteria were (a) contact with another language besides German did not occur at home (n = 33), and (b) questionnaires were incomplete (n = 3) or contained conflicting information on the language background (n = 3).

We recruited participants through class teachers who distributed written information about the study among the children's parents. The parents gave written informed consent before the start of the study. All participants had normal or corrected-to-normal vision, were naive with regard to the hypotheses of the study and received small gifts for their participation.

Questionnaires

The parents completed a paper-pencil questionnaire at home (a German and a Turkish version were available) and reported the country of birth of father, mother and child (to determine the migration background). The family's SES was specified by the ISCED (International Standard Classification of Education, UNESCO Institute for Statistics, 2012) for both parents, which reflects the highest school and professional qualification. Parents reported also their children's language background including the languages spoken at home and in kindergarten, the AoA of German, language proficiency and frequency of use at home.

The participating children filled in a child-friendly version of largely the same questionnaire in the first experimental session (for more detail, see below). The children's proficiency in German and in L1 was rated both by themselves and their parents for comprehension and production on a scale of 1 (*none*) to 4 (*very good*). The overlap between these ratings was high for German (comprehension: 74% / production: 67%), but lower for L1 (39% / 30%). Faced with this discrepancy between a rather high convergence for ratings of German proficiency but lower

convergence for rating L1 verbal abilities, we decided to rely only on the parents' ratings. Regarding the ratings in German, both parents and children might include grades and feedback from teachers influencing their ratings; the ratings in the L1, however, are based on parents' judgements, most likely in comparison to their own language proficiency as speakers of that language and compared to the children's abilities in German. The children, though, may lack this point of reference for comparing their own language abilities in their L1, as they are growing up in a multilingual community rather than with peers of the same L1 only.

EF tasks

Participants sat at a desk in front of a tablet computer in a quiet classroom. All three EF tasks were presented on a Microsoft Surface Pro 2 Tablet with a display size of 25.5 cm x 17 cm and a resolution of 2160 px x 1440 px.

Response inhibition

We administered the Go/Nogo task to measure response inhibition. In this task, we presented two black-and-white line drawings of animals (a goat and a deer from Snodgrass & Vanderwart, 1980), which were visually highly similar, to increase task difficulty. Children were instructed to touch a response bar on the screen with the index finger of their dominant hand when they saw the target picture (Go-condition, e.g., the goat) but to withhold the response in the Nogo-condition (e.g., the deer). Targets were randomly assigned to group sessions and were counterbalanced over participants. We explained the task orally to the group and pictograms on the tablet illustrated the procedure to ensure all children understood independent of their German skills or processing speed. Children completed 10 practice trials with feedback (happy or sad smiley); if less than 60% of Go-trials were answered correctly, practice was repeated up to two times. Afterwards participants completed two experimental blocks consisting of 20 trials each (10 Go- and 10 Nogo-trials) without feedback. Items were pseudo-randomized so that no more than two equal items followed each other. Each item (7 cm x 7 cm) appeared in the center of the screen for 2 s, with an inter-stimulus interval of 2 s, allowing a maximum RT of 4 s to meet the requirements of the heterogeneous sample.

Interference inhibition

To test interference inhibition, we administered an adapted version of the Bivalent Shape Task (BST, Mueller & Esposito, 2014). Children were instructed to sort circles and squares (presented in the center of the screen, Ø 4.2 cm) according to their shape by pressing corresponding buttons on the tablet screen (lower left and right side, Ø 2.1 cm) with their index fingers. The difficulty lies in ignoring the more salient color of the shapes and concentrating on their shape. Depending on the color, the stimuli belonged to three congruency conditions: (a) black outlines in the neutral condition, (b) a red and a blue shape matching the response buttons in color and shape in the congruent condition, and (c) interchanged colors in the incongruent condition, so that the shape but not the color matched the response buttons. Position (right, left) and color (blue, red) of the response buttons were counterbalanced over participants. The experiment consisted of three uniform blocks in a fixed order (neutral, followed by congruent and incongruent) with 20 randomized trials (each item appeared 10 times). Afterwards children completed one mixed block including all congruency conditions that comprised 30

randomized trials with 10 items of each condition. Items appeared immediately after button press or after 3 s, if no answer was given. Between each block, there was a break, but children were encouraged to complete the experiment as fast as they could. Instructions and examples were given orally to the whole group. Additionally, children completed a practice block on the tablet including 12 randomized trials (four of each condition) with visual feedback in German ("Richtig!", engl. correct, in green or "Falsch", engl. wrong, in red letters).

Updating

As a measure of updating, the children completed two versions of the N-back task: a 1-back followed by a 2-back task. The stimuli, five letters (A, B, O, R and S; 3.7 cm high), were presented one-by-one in the center of the screen. We asked children to press a response bar with their index finger of the dominant hand when the displayed letter was the same as one trial (1-back) or two trials ago (2-back). Letters were presented for 2 s with an inter-stimulus interval of 2 s without feedback. Both tasks comprised two pseudorandomized blocks including 20 items of which 40% were critical items and no more than two succeeded one another. The procedure was the same as for the Go/Nogo task regarding instructions and practice (10 trials with visual feedback).

Control variables

Intelligence

We used the second part of the Culture Fair Intelligence Test Scale 1 (CFT 1-R, Weiß & Osterland, 2013), which assesses figural reasoning skills. The sum of correct answers from the short version of the subtests Matrix, Series and Classification served as measure of intelligence. The CFT, a culture-independent and language-free test, should not disadvantage children from other cultural backgrounds or with lower German skills.

Phonological short-term memory

The number of correctly repeated non-words from the ZLT-II (Petermann & Daseking, 2012) served as a measure of short-term memory. The test comprises 30 items, which consist of 2 to 6 syllables presented with increasing length (6 items per length). Items were pre-recorded to standardize testing conditions and consisted of meaningless CV-syllables presented with a neutral prosody (equal stress on each syllable) to avoid German-specific structures (e.g., consonant clusters or prosody).

German lexicon

We used the WWT (short version test 2, Glück, 2011) to assess expressive and receptive lexicon size in German. First, a picture-naming task with 40 items (including nouns, verbs and naming opposites of adjectives) served as measure of expressive lexicon size. Second, the test for receptive lexicon comprised all incorrect items from the preceding picture-naming task. Items were presented via loudspeaker and the child was instructed to touch the corresponding out of four pictures on the screen. The number of correct answers in each task served as measure of expressive or receptive lexicon.

Procedure

This study is part of a larger project conducted in three inclusive primary schools (see for example Czapka, Klassert & Festman,

2019); we report here only the tasks relevant for this study. After receiving the completed parents' questionnaire, three experimental sessions proceeded in the following way: in two group sessions, children filled in the questionnaire and completed the EF tasks and the intelligence test; in a third, individual session, we assessed German lexicon size and short-term memory.

In group sessions (with max. 20 children per session), tasks were explained by one trained experimenter. Depending on group size, each additional experimenter (up to nine per session) supervised between one child to maximally four children to ensure a quiet atmosphere, correct administration of tests, immediate help in case of technical problems (with the tablet computers) and understanding of instructions despite language, comprehension or processing speed difficulties.

Data analysis: statistical analysis and matching

In the EF tasks, we analyzed only RTs, since performance in terms of accuracy was at ceiling and error rates (calculated with Mann-Whitney-U test due to non-normal distributions) did not differ significantly between groups in any tasks or conditions. For the analysis, RTs of correct responses were log-transformed to normalize distributions. We removed outliers in two steps: first, exclusion of participants when their performance clearly deviated from the group; and secondly, removal of single data points by visual inspection of the RT-distribution for each task (see below).

We calculated linear mixed effect models for each task separately with the lme4 package (Douglas et al., 2015) in R version 3.2.2. (R Core Team, 2015). The models included RTs for each trial (level-one units) nested in the different subjects who were level-two units. To calculate group differences, the models included fixed effects for GROUP (monolinguals, multilinguals), the experimental variables and interactions between them. We added random intercepts for participant and random by-subject slopes for the experimental manipulations (block for the Go/Nogo task, congruency and type for the BST and difficulty for the N-back). This procedure was applied to the heterogeneous and the homogeneous sample.

To obtain a homogenous sample for the second analysis, we matched participants of both groups pairwise on variables known from the literature to possibly influence executive functions. These important covariates were age, gender, intelligence, lexicon size and SES (ISCED mother; cf. Table 2). To avoid multicollinearity due to high correlations between some background variables (cf. Table 3), we chose the mother's ISCED as measure of SES and expressive lexicon size as measure of German lexicon size. To match participants, we selected pairs of mono- and multilinguals from the data, 1) who did not differ in gender, 2) whose ISCED differed not more than one point (on a 4-point scale), 3) whose age, lexicon and intelligence scores were as close as possible (group differences between matched pairs amount for age in months: M = 5.1, SD = 5.2, for lexicon on a scale between 0 to 40: M = 3.0, SD = 2.4, and for intelligence on a scale between 0 and 45: M = 2.4, SD = 1.5). A summary of the demographic variables (group means with standard deviations) with group comparisons calculated with unpaired t-tests can be found in Table 1 for the heterogeneous and in Table 2 for the homogenous groups. Pairwise matching (instead of adding all these variables as covariates to the models) was chosen to make the process of group matching transparent (what is rarely done in the literature), to ensure that children are comparable on an individual level (in contrast to matching on a group-level), to

Table 2. Descriptive statistics for homogeneous groups

	Monolinguals		Multilinguals	
n / females	21 / 57%		21 / 57%	
Age (in months)	106.38	(4.5)	107.7	(6.5)
Intelligence	37.0	(3.3)	38.0	(3.9)
Short-term memory	15.5	(3.1)	16.1	(3.8)
SES (mother's ISCED)	3.8	(1.0)	3.8	(1.2)
SES (father's ISCED)	3.7	(1.1)	3.6	(1.5)
German expressive lexicon	23.2	(6.1)	23.4	(5.8)
German receptive lexicon	37.5	(2.4)	37.5	(2.1)
Self-rated German proficiency	3.8	(0.4)	3.9	(0.4)
Self-rated L1 proficiency			2.8	(0.4)
AoA German (in months)			11.2	(13.5)

Note. Means and standard deviations (in brackets) for pairwise matched monolinguals and multilinguals.

simultaneously control for a set of background variables that impact on EF development and to avoid overfitting the regression models with too many variables.

Results

Heterogeneous groups

Background variables

In Table 1, we show means and standard deviations for the heterogeneous sample (monolinguals vs. multilinguals) on all collected background variables. Multilinguals were significantly older, showed poorer German skills and their parents' ISCED was significantly lower. The heterogeneity of our sample was also reflected in the multilinguals' proficiency in L1, which was lower and more heterogeneous than their proficiency in German, as well as in the wide range of AoA (see Table 1). The frequency with which the multilingual children used their languages at home was also very diverse: 30% reported that they used *predominantly one or more languages other than German,* 17% spoke *mainly German* and 32% used *both German and another language similarly often* (for 21% of the cases no information was given). Most bilinguals spoke Turkish (37%) or Arabic (17%), the others Albanian, Bosnian, Chinese, Edo, English, Greek, Hungarian, Kurdish, Persian, Polish, Punjabi, Russian, Serbian or Spanish. In the trilingual group, children spoke Albanian, Edo, English, Hebrew, Mandinka, Polish, Rumanian, Serbian, Turkish or Zaza. Overall, 83% of the multilingual group and 7% of the monolingual group indicated a migration background, meaning at least one parent or the child was born outside of Germany.

EF-Tasks

An overview of the raw values for the three EF tasks can be found in Figure 1.

Response inhibition (Go/Nogo)

The regression model for the response inhibition task consisted of 2370 observations from 125 participants and is presented in Table 4. Outliers comprised the complete data set of one

Table 3. Correlations between background variables

	1	2	3	4	5	6
1. Age	–					
2. Intelligence	−.10	–				
3. German expressive lexicon	−.26***	.35***	–			
4. German receptive lexicon	−.26***	.34***	.87***	–		
5. SES (mother's ISCED)	−.41***	.17	.49***	.46***	–	
6. SES (father's ISCED)	−.37***	.18	.37***	.37***	.58***	–
7. Short-term memory	−.17	.09	.31***	.28***	.23**	.08

Note. Spearman-correlation coefficients (*** $p < .001$, ** $p < .01$, * $p < .05$)
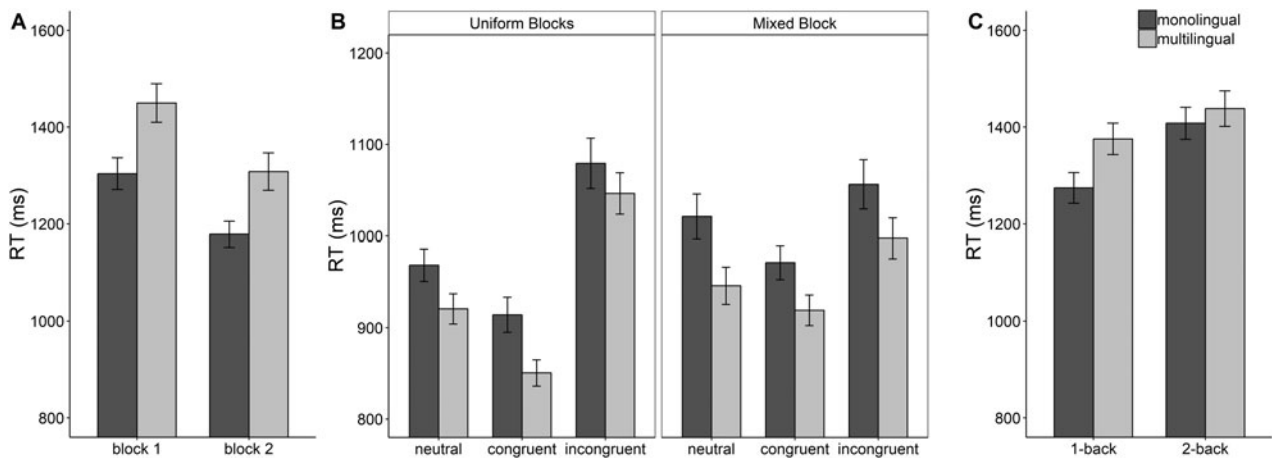

Fig. 1. Mean response latencies (with error bars indicating standard errors) of the A) go/nogo task, B) BST and C) N-back for the heterogeneous groups (monolinguals - dark bars and multilinguals - light bars).

bilingual who performed below chance and RTs with log(RT) > 7.8 or log(RT) < 6.4 (1.2% of data points, corresponds to about M ± 2 SD). There was significant variance in intercepts across participants. RTs significantly decreased from the first to the second block and multilinguals answered significantly slower than monolinguals.

Interference inhibition (BST)

Table 4 shows the results for the interference inhibition task. 10,435 observations of 124 participants were included in this model. We excluded two multilingual children from the analysis due to below chance performance in one of the uniform blocks. Additionally, single blocks with less than four correct responses and RTs with log(RT) < 5.8 were removed (0.1% of all data points, corresponds to approximately M ± 3 SD). There was significant variance in intercepts across participants. Significant experimental manipulations were block type (RTs increased in the mixed block) and congruency (in comparison to neutral trials, performance on congruent trials was faster and on incongruent trials slower); additionally, incongruency had a smaller effect in mixed compared to uniform blocks. No significant interaction of multilingualism with any variable was found, but there was a marginal effect that multilinguals respond overall faster than monolinguals.

Updating (N-back)

The model for the updating task is presented in Table 4. One bilingual child did not take part in this task and we needed to

exclude four multilingual children from the analysis in the 2-back task due to below chance performance. Outliers with log (RT) > 7.9 and log(RT) < 6.3 were removed (0.9% of data points, corresponds to approximately M ± 2 SD). The final data set consisted of 2585 observations from 125 participants. There was significant variance in intercepts across participants and RTs were significantly longer in the 2-back task compared to the 1-back. Overall, multilinguals performed significantly slower than monolinguals and a trend indicates that the increase in difficulty from 1- to 2-back is smaller for multilinguals than monolinguals.

Homogeneous groups

Background variables

Pairwise matching resulted in a subsample of n = 21 children per group, who did not differ in any background variables (cf. Table 2). The multilingual group included n = 3 trilinguals and n = 18 bilinguals. The reported frequency of language use at home was the following: 24% spoke *predominantly one or more languages other than German*, 33% spoke *mainly German* and 24% spoke *German and another language similarly often* (19% are missing data). Four bilinguals spoke Turkish, and the remaining spoke Albanian, Arabic, Bosnian, Chinese, English, Hungarian, Polish, Punjabi, Russian or Spanish. Trilinguals spoke English and Mandinka, Albanian and Polish, and Turkish and Zaza.

Table 4. Regression models for the EF tasks

		Heterogeneous sample			Homogeneous sample		
		b	SE	t	b	SE	t
Go/Nogo	(Intercept)	7.11	0.02	287.03	7.09	0.04	162.19
	Block 2	−0.08	0.02	−3.87	−0.09	0.03	−2.64
	Multilingualism	0.11	0.04	2.89	0.07	0.06	1.21
	Block 2 : Multilingualism	−0.02	0.03	−0.64	0.01	0.05	0.12
BST	(Intercept)	6.84	0.02	407.63	6.86	0.03	267.69
	Congruency	−0.06	0.01	−6.15	−0.09	0.02	−4.73
	Incongruency	0.08	0.01	5.42	0.08	0.03	3.11
	Mixed	0.05	0.02	3.00	0.04	0.02	1.57
	Multilingualism	−0.05	0.03	−1.88	−0.07	0.04	−2.03
	Congruency : Mixed	0.02	0.02	0.91	0.03	0.03	1.16
	Incongruency : Mixed	−0.06	0.02	−3.19	−0.03	0.03	−1.02
	Congruency : Multilingualism	−0.02	0.02	−1.08	−0.01	0.03	−0.26
	Incongruency : Multilingualism	0.03	0.02	1.22	0.00	0.04	0.09
	Mixed : Multilingualism	−0.02	0.02	−1.07	−0.02	0.03	−0.67
	Congruency : Mixed : Multilingualism	0.04	0.03	1.52	0.04	0.04	1.03
	Incongruency : Mixed : Multilingualism	0.00	0.03	0.01	0.00	0.04	0.05
N-back	(Intercept)	7.08	0.02	312.51	7.06	0.04	192.88
	2-back	0.10	0.02	4.56	0.03	0.04	0.73
	Multilingualism	0.08	0.03	2.26	0.03	0.05	0.62
	2-back : Multilingualism	−0.05	0.04	−1.33	0.06	0.05	1.13

Note. Fixed effects of the linear mixed models for the heterogeneous and homogeneous sample.

EF-tasks

Figure 2 displays the raw values for the response inhibition, interference inhibition and updating task of the homogeneous sample.

Response inhibition

For the Go/Nogo task, 805 observations from 42 participants were included in the model (see Table 4). There was significant variance in intercepts across participants and RTs decreased from the first to the second block. Multilingualism did not influence RTs.

Interference inhibition

The final model (see Table 4) for the BST comprises 3594 observations from 42 participants. There was significant variance in the intercepts across participants. Congruency influenced RTs significantly: In comparison to neutral trials, performance on congruent trials was faster and on incongruent trials slower. The influence of block type was marginally significant. Multilinguals responded overall significantly faster than monolinguals, but no interactions reached statistical significance.

Updating

The regression model for the N-back (see Table 4) consists of 899 observations of 42 participants. There was significant variance in the intercepts across participants, but no significant difference between tasks or groups.

Discussion

In this study, we first compared heterogeneous groups of mono- and multilingual children on their performance in three EF tasks, tapping into response inhibition, interference inhibition and updating. When group membership was confounded with differences in SES and German skills, multilinguals showed a disadvantage in the Go/Nogo and N-back task, but similar performance in the BST. After matching participants of both groups pairwise on possibly influential variables (i.e., age, gender, intelligence, SES and German lexicon size), multilinguals outperformed their monolingual peers in the BST; more specifically they answered overall faster. The differences in the Go/Nogo and the N-back task disappeared: multilinguals were no longer slower in response inhibition and in updating.

Conclusions from the heterogeneous groups

When we 'ignored' between-group differences in age, SES and lexicon size in our first analysis based on the heterogeneous groups, we found no bilingual advantage, but rather a disadvantage in updating and response inhibition. This was expected due to the known detrimental effects of these variables for the multilingual group. This concerns especially their lower SES, which negatively influences EF (Carlson et al. 2013; Morton & Harper, 2007). In general, the parents' SES, as measured by either financial or educational aspects, is related to the amount of
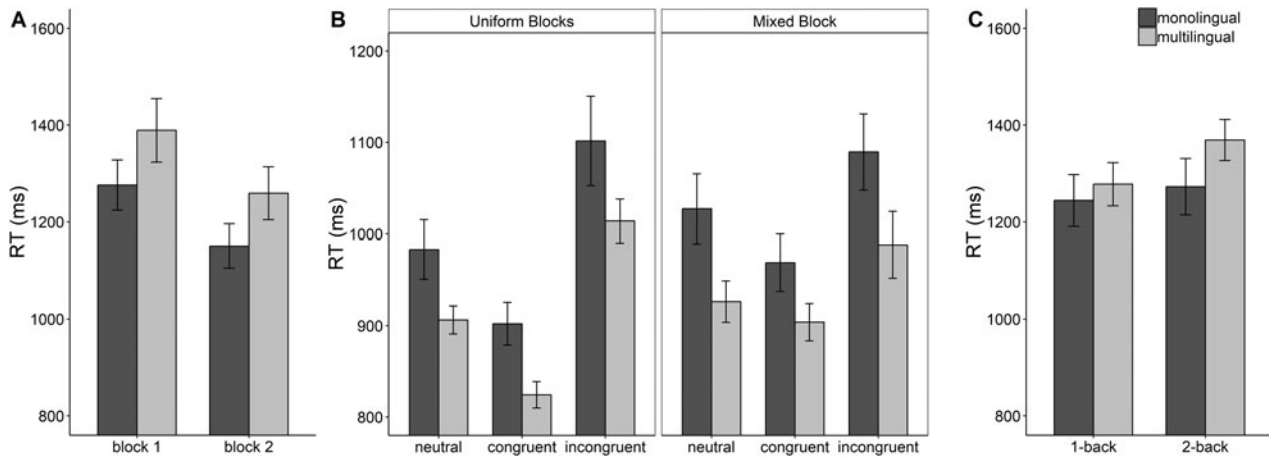
Fig. 2. Mean response latencies (with error bars indicating standard errors) of the A) go/nogo task, B) BST and C) N-back for the heterogeneous groups (mono-linguals - dark bars and multilinguals - light bars).

stimulating educational input children receive and is therefore a major factor for cognitive development including EF and, for example, lexicon (Calvo & Bialystok, 2014; Carlson et al. 2013).

Nevertheless, it is remarkable that multilinguals were apparently able to compensate for these detrimental effects and performed equally on the BST in the first comparison between heterogeneous groups. It seems that they were able to counter the negative influences of the background variables with better cognitive control. This pattern was also reported in Carlson and Meltzoff (2008) for Spanish–English preschool children who did not differ in their EF performance from monolingual peers despite disadvantages in SES, age and verbal ability. In our study, multilinguals have a more difficult starting position due to their parents' lower SES but also because of their poorer German skills, their later start of German acquisition and a disadvantage in the German education system (cf. Fereidooni, 2012). Therefore, one cannot expect a bilingual advantage in executive functions, because the combination of disadvantaging factors (linked to their multilingual migration background) likely overshadows a cognitive advantage. However, their comparable performance indicates that multilinguals can at least compensate for these disadvantages through superior EF.

Conclusions from pairwise matching

The matching procedures in the literature are in general applied on a group level (e.g., Duñabeitia et al., 2014; Kousaie et al., 2015). Our method of matching participants on an individual level, however, considers the distribution of multiple variables in each matched couple of mono- and multilinguals. A more rigorous approach like this was already demanded by other researchers, like Hilchey and Klein (2011): "the extent to which bilingualism is the complete, partial, or apparent cause of these data is an area that warrants further investigative work, and we urge future investigators of the BICA/ BEPA hypotheses to be assiduous in their efforts to match monolinguals and bilinguals on plausibly pertinent demographic factors" (p. 643). Our approach for matching participants aimed at controlling for as many important background variables as possible, as it seemed likely that it was a certain combination of background variables that made the difference. By reporting in detail the process of reducing our initial sample of 168 children to two homogeneous groups of 21 mono- and 21 multilinguals, we want to encourage

other researchers to reveal their methods of participant selection more clearly. In contrast to other matching strategies that focused, for example, only on reaching insignificance between group means, we carefully chose pairs of children who overlapped as well as possible in all our background variables. Thus, this strategy ensured that the language groups consisted of children who were similar on both the group and the individual level. This reduced the sample size tremendously, but "by reducing samples to ensure that groups are more comparable, (…) smaller studies can be more informative than larger studies, if they are better controlled." (Hope, 2015, p. 59). After closely following these recommendations, our results indicate indeed that this is an appropriate method to approach the bilingual advantage debate, since after minimizing the influence of all potentially influential background variables by pairwise matching, a multilingual advantage in the BST emerged and the differences in the other tasks were balanced out.

Matching on lexicon size had an additional effect: namely that we excluded children not only with particularly poor German skills but also with late AoA of German from the sample. Therefore, the remaining multilinguals supposedly underwent a more extensive training in their languages, which is important because the proficiency in the single languages (Luk & Bialystok, 2013), their usage (Luk & Bialystok, 2013; Yow & Li, 2015), and the AoA (Kapa & Colombo, 2013) influence EF. Their exhaustive bilingual experience may have resulted in superior performance in the BST. Crucially, future research should consider bilingual experience more thoroughly: in particular, because of the tremendously heterogeneous experiences in countries, speech communities, school settings, and so forth.

Superior interference inhibition

Since multilinguals outperformed their monolingual peers only in the BST but not in the Go/Nogo task, our results support the dissociation between interference and response inhibition (Luk, Anderson, Craik, Grady & Bialystok, 2010; Martin-Rhee & Bialystok, 2008). The crucial difference between these concepts is that the information in Go/Nogo tasks is not by itself 'conflicting' but requires only execution or suppression of a motor response. In contrast, in the BST, every stimulus contains potentially conflicting information on shape and color that needs to be processed to allow for correct responses. The increased attentional

demands due to the potentially conflicting information probably involves the neural networks also required in bilingual language processing: constant attention towards the relevant stimulus feature (or the target language) and inhibition of the irrelevant information (or the non-target language). Thus, speaking multiple languages affects only interference but not response inhibition; a finding consistent with previous literature (Esposito et al., 2013; Martin-Rhee & Bialystok, 2008).

Multilinguals showed globally reduced RT in the BST. This result indicates that not solely inhibition, but rather executive processing is enhanced in multilinguals. Hilchey et al. (2015) term this pattern BILINGUAL EXECUTIVE PROCESSING ADVANTAGE (BEPA) and Bialystok (2017) SUPERIOR EXECUTIVE ATTENTION. Both comprise not only advanced inhibition of irrelevant information but also advanced attention towards relevant information. This interpretation is in accordance with the theory on bilingual language processing: these two aspects of EF – inhibition and attention – are essential for bilinguals to comprehend and produce language (see introduction for details) and their constant usage trains the associated EF components. Consequently, the presence of potentially conflicting information in interference inhibition tasks seems to be the key feature that allows multilinguals to reveal their EF advantage.

If only inhibition was modulated by bilingualism, the advantage would manifest itself in a smaller interference effect and particularly lower RT in incongruent conditions (Hilchey et al. refer to this pattern as BICA). Esposito et al. (2013) found this pattern in the BST when they compared bilingual 4-year-olds with monolingual peers. Differences in age and therefore cognitive development between our sample and Esposito and colleagues' might cause the discrepancy in outcome patterns.

Updating

Multilinguals answered overall slower in the N-back task, but after matching participants no difference was found. This finding is in line with other studies showing no influence of bilingualism on working memory: for example, in the Listening Span and Pattern Recall Task (Namazi & Thordardottir, 2010), or in Backward Digit Span and Counting Recall (Engel de Abreu, 2011). However, Blom et al. (2014), who used a similar methodology, found opposite results: they compared visuospatial and verbal working memory between monolingual Dutch and Turkish–Dutch children, who had a migration background, lower SES and smaller Dutch vocabulary size (in that way, a sample similar to ours). Initially, both groups performed similarly, but, when SES and vocabulary were controlled, bilinguals outperformed monolinguals. These opposite results may stem from differences in samples or task impurity, which impedes precise measurement of EF due to increasing chances of type I and type II errors (Friedman, 2016). Task impurity results from variance in stimuli characteristics (word, digit, or picture), response modality (motor or verbal) or their dynamics (they assess either the storage capacity, as in span tasks, or the ability to manipulate working memory content, like the N-back task that requires concurrent updating of memory content and dropping of irrelevant information; Conway, Kane, Bunting, Hambrick, Wilhelm & Engle, 2005). These factors might conceal common features of N-back and other working memory tasks and cause the heterogeneity of research findings. Our results do not support a bilingual advantage in updating, but its existence cannot be excluded (Calvo, Ibáñez & García, 2016).

The role of high cognitive demands

Since training studies with EF tasks (Diamond, 2014) and other studies on the bilingual advantage (Costa et al., 2009; Martin-Rhee & Bialystok, 2008) usually found a bilingual advantage in the more demanding conditions, we expected this also in our study. However, neither the increased EF demands from 1- to 2-back nor from uniform to mixed block in the BST yielded a bilingual advantage. Without pairwise matching, no advantage could arise because of the strong influence of the other background variables, which had a negative influence on the multilinguals' performance. With pairwise matching, however, the more difficult experimental conditions did not yield additional challenges for the homogeneous groups anymore. We conclude that matching led to a sample with on average higher mental faculties.

Comparing response latencies on the uniform versus the mixed block in the BST, we observed that the additional difficulty of having to switch between different types of stimuli (neutral, congruent, incongruent) seems to have been compensated by overall longer RTs, particularly on the easier trial types (neutral, congruent), rather than on the more difficult, incongruent condition. This strategy has been observed as well in studies on language switching, in particular in trilingual contexts. When participants had to switch between three languages for naming digits, such mixed conditions were at the expense of the more dominant languages (Festman & Mosca, 2016).

Limitations

In this study, we assessed three different EF components with three tasks. Due to task impurity, however, further studies need to investigate which tasks measure exactly which EF components. Friedman (2016) advises researchers for example to include multiple tasks and latent variables to ensure precise measurement of EF. Equally important is controlling for influential factors in research on the bilingual advantage. As suggested by Hope (2015), our approach aimed at controlling for as many important background variables as possible. Future research should provide even more insights into their role on EF and their possibly differential impact on mono- versus multilinguals' processing (for one recent example of the differential role of EF on mono- vs. multilinguals' spelling, see Czapka et al., 2019). An alternative approach would be regression analyses based on even larger samples of mono- and multilinguals and to statistically control for the background variables in one model. The variable we could not balance between language groups was migration status. How it affects EF is still unknown, but further research investigating the interaction between EF and migration status is necessary to exclude uncontrolled influence.

Conclusion

The analysis of our heterogeneous sample of primary school children showed that multilinguals with disadvantages in SES and German skills underperformed in the updating and response inhibition tasks. Since their performance was equal in interference inhibition, we conclude that they compensated for the detrimental effects of the background variables by superior EF. Our comparison of the homogeneous groups supports this conclusion. Following this initial analysis, we meticulously matched participants on age, gender, intelligence, SES and German lexicon size to achieve groups that were comparable on a group and individual level. The performance in the updating and response inhibition

tasks was now equal between those matched groups, and in the interference inhibition task, multilinguals showed globally reduced RT indicating a bilingual executive processing advantage.

Author ORCIDs. 🄸🄳 Julia Festman, 0000-0003-0156-3713

References

Adesope OO, Lavin T, Thompson T and Ungerleider C (2010) A Systematic Review and Meta-Analysis of the Cognitive Correlates of Bilingualism. *Review of Educational Research* 80, 207–245. doi:10.3102/0034654310368803

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

Bak TH, Long MR, Vega-Mendoza M and Sorace A (2016) Novelty, Challenge, and Practice: The Impact of Intensive Language Learning on Attentional Functions. *PLoS ONE* 11(4), e0153485. https://doi.org/10.1371/journal.pone.0153485

Barac R and Bialystok E (2011) Cognitive development of bilingual children. *Language Teaching* 44, 36–54. doi:10.1017/S0261444810000339

Best JR and Miller PH (2010) A Developmental Perspective on Executive Function. *Child Development* 81, 1641–1660. doi:10.1111/j.1467-8624.2010.01499.x.A

Bialystok E (2015) Bilingualism and the Development of Executive Function: The Role of Attention. *Child Development Perspectives* 9, 117–121. doi:10.1111/cdep.12116

Bialystok E (2017) The bilingual adaptation: How minds accommodate experience. *Psychological bulletin* 143, 233–262. doi:10.1037/bul0000099

Bialystok E, Barac R, Blaye A and Poulin-Dubois D (2010) Word mapping and executive functioning in young monolingual and bilingual children. *Journal of Cognition and Development* 11, 485–508.

Bialystok E and Luk G (2012) Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition* 15, 397–401. doi:10.1017/S1366728909990423

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

Bialystok E and Viswanathan M (2009) Components of executive control with advantages for bilingual children in two cultures. *Cognition* 112(3), 494–500. https://doi.org/10.1016/j.cognition.2009.06.014

Bildungsberichterstattung A (2014) *Bildung in Deutschland 2014 - Ein indikatorengestützter Bericht mit einer Analyse. Bildungsbericht 2014.* Retrieved from www.bildungsbericht.de

Blom E, Küntay AC, Messer MH, Verhagen J and Leseman PPM (2014) The benefits of being bilingual: Working memory in bilingual Turkish – Dutch children. *Journal of Experimental Child Psychology* 128, 105–119. doi:10.1016/j.jecp.2014.06.007

Calvo A and Bialystok E (2014) Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition* 130, 278–288. doi:10.1016/j.cognition.2013.11.015

Calvo N, Ibáñez A and García AM (2016) The Impact of Bilingualism on Working Memory: A Null Effect on the Whole May Not Be So on the Parts. *Frontiers in Psychology* 7, 1–4. doi:10.3389/fpsyg.2016.00265

Carlson SM and Meltzoff AN (2008) Bilingual experience and executive functioning in young children. *Developmental Science* 11(2), 282–298. https://doi.org/10.1111/j.1467-7687.2008.00675.x

Carlson SM, Zelazo PD and Faja S (2013) Executive function. In PD Zelazo (Ed.), *The Oxford Handbook of Developmental Psychology, Vol. 1: Body and mind* (pp. 706–743). New York: Oxford University Press.

Conway A, Kane M, Bunting M, Hambrick Z, Wilhelm O and Engle R (2005) Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review* 12, 769–786. https://doi.org/10.3758/BF03196772

Costa A, Fuentes LJ and Vivas AB (2010) The impact of bilingualism on the executive control and orienting networks of attention. *Bilingualism: Language and Cognition* 13, 315–325. doi:10.1017/S1366728909990010

Costa A, Hernandez M, Costa-Faidella J and Sebastián-Gallés N (2009) On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition* 113, 135–149. doi:10.1016/j.cognition.2009.08.001

Czapka S, Klassert A and Festman J (2019) Executive functions and language: Their differential influence on mono- vs. multilingual spelling in primary school. *Frontiers in Psychology* 10, 97. doi:10.3389/fpsyg.2019.00097

de Bruin A, Treccani B and Della Sala S (2014) Cognitive advantage in bilingualism: an example of publication bias? *Psychological Science* 26, 99–107. https://doi.org/10.1177/0956797614557866

Diamond A (2013) Executive functions. *Annual Review of Psychology* 64, 135–68. doi:10.1146/annurev-psych-113011-143750

Diamond A (2014) Want to Optimize Executive Functions and Academic Outcomes?: Simple, Just Nourish the Human Spirit. In *Minnesota Symposia on Child Psychology* (Vol. 37, p. 205–232). NIH Public Access.

DIMDI. (n.d.) ICD-10-GM Version 2016. Retrieved September 20, 2016, from www.dimdi.de/static/de/klassi/icd-10-gm/kodesuche/onlinefassungen/htmlgm2016/block-f70-f79.htm

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

Engel de Abreu PMJ (2011) Working memory in multilingual children: Is there a bilingual effect? *Memory* 19(5), 529–537. doi:10.1080/09658211.2011.590504

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

Esposito AG, Baker-Ward L and Mueller ST (2013) Interference suppression vs. response inhibition: An explanation for the absence of a bilingual advantage in preschoolers' Stroop task performance. *Cognitive Development* 28, 354–363. doi:10.1016/j.cogdev.2013.09.002

Fereidooni K (2012) Kinder mit Migrationshintergrund im deutschen Schulwesen–Benachteiligung aus (Bildungs-) politischen Ursachen?. *Gesellschaft–Wirtschaft–Politik* (GWP) 61, 363–371.

Festman J and Mosca M (2016) Influence of preparation time on language control. A trilingual digit-naming study. In JW Schwieter (Ed.), *Cognitive Control and Consequences of Multilingualism* (pp. 145–171). John Benjamins Publishing Company. https://doi.org/10.1075/bpa.2.07fes

Friedman NP (2016) Research on individual differences in executive functions: Implications for the bilingual advantage hypothesis. *Linguistic Approaches to Bilingualism*. doi:10.1075/lab.15041.fri

Glück CW (2011) *WWT 6-10. Wortschatz- und Wortfindungstest für 6- bis 10-Jährige* (2nd ed.). München: Urban & Fischer.

Green DW (1998) Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1, 67–81. doi:10.1017/S1366728998000133

Hagmann-Von Arx P, Petermann F and Grob A (2013) Konvergente und diskriminante Validität der WISC-IV und der Intelligence and Development Scales (IDS) bei Kindern mit Migrationshintergrund. *Diagnostica* 59, 170–182.

Hilchey MD and Klein RM (2011) Are there bilingual advantages on non-linguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review* 18, 625–58. doi:10.3758/s13423-011-0116-7

Hilchey MD, Saint-Aubin J and Klein R (2015) Does bilingual exercise enhance cognitive fitness in traditional non-linguistic executive processing

tasks? In JW Schwieter (Ed.), *The Cambridge Handbook of Bilingual Processing* (pp. 286–613). Cambridge University Press.

Hope TMH (2015) The Bilingual Cognitive Advantage: No Smoke without Fire? *AIMS Neuroscience* **2**, 58–65. doi:10.3934/Neuroscience.2015.2.58

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

Kousaie S, Laliberté C, López Zunini R, Taler V, Zunini RL and Taler V (2015) A behavioral and electrophysiological investigation of the effect of bilingualism on lexical ambiguity resolution in young adults. *Frontiers in Human Neuroscience* **9**(December), 1–14. https://doi.org/10.3389/fnhum. 2015.00682

Kroll JF, Dussias PE, Bogulski CA and Kroff JRV (2012) Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. *Psychology of Learning and Motivation* **56**, 229–262. https://doi.org/10.1016/B978-0-12-394393-4.00007-8

Leerhoff H, Rehkämper K, Rockmann U, Brunner M, Gärtner H and Wendt W (2013) *Bildung in Berlin und Brandenburg 2013*. (Autorengruppe Regionale Bildungsberichterstattung Berlin-Brandenburg, Ed.).

Luk G, Anderson JAE, Craik FIM, Grady C and Bialystok E (2010) Distinct neural correlates for two types of inhibition in bilinguals: Response inhibition versus interference suppression. *Brain and Cognition* **74**, 347–357. https://doi.org/10.1016/j.bandc.2010.09.004

Luk G and Bialystok E (2013) Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology* **25**, 605–621. doi:10.1080/20445911.2013.795574

Luo L, Craik FIM, Moreno S and Bialystok E (2013) Bilingualism interacts with domain in a working memory task: Evidence from aging. *Psychology and Aging* **28**, 28–34. doi:10.1037/a0030875

Martin-Rhee MM and Bialystok E (2008) The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: Language and Cognition* **11**, 81–93. doi:10.1017/S1366728907003227

Miyake A and Friedman NP (2012) The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current Directions in Psychological Science* **21**, 8–14. doi:10.1177/0963721411429458

Morton JB and Harper SN (2007) What did Simon say? Revisiting the bilingual advantage. *Developmental Science* **10**, 719–726. doi:10.1111/j.1467-7687.2007.00623.x

Mueller ST and Esposito AG (2014) Computerized Testing Software for Assessing Interference Suppression in Children and Adults: The Bivalent Shape Task (BST). *Journal of Open Research Software* **2**, e3.

Namazi M and Thordardottir E (2010) A working memory, not bilingual advantage, in controlled attention. *International Journal of Bilingual Education and Bilingualism* **13**, 597–616. doi:10.1080/13670050.2010.488288

Niklas F, Schmiedeler S, Pröstler N and Schneider W (2011) Die Bedeutung des Migrationshintergrunds, des Kindergartenbesuchs sowie der Zusammensetzung der Kindergartengruppe für sprachliche Leistungen von Vorschulkindern. *Zeitschrift Für Pädagogische Psychologie* **25**, 115–130.

Paap KR and Greenberg ZI (2013) There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology* **66**, 232–258. doi:10.1016/j.cogpsych.2012.12.002

Paap KR, Johnson HA and Sawi O (2014) Are bilingual advantages dependent upon specific tasks or specific bilingual experiences? *Journal of Cognitive Psychology* **26**, 615–639. doi:10.1080/20445911.2014.944914

Paap KR, Johnson HA and 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. doi:10.1016/j.cortex.2015.04.014

Paap KR and Sawi O (2014) Bilingual advantages in executive functioning: Problems in convergent validity, discriminant validity, and the identification of the theoretical constructs. *Frontiers in Psychology* **5**, 1–15. https://doi.org/10.3389/fpsyg.2014.00962

Petermann F and Daseking M (2012) *Zürcher Lesetest - II (ZLT-II)*. Bern: Huber.

Poarch GJ and van Hell JG (2012) Executive functions and inhibitory control in multilingual children: Evidence from second-language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology* **113**, 535–551.

R Core Team. (2015) R: A Language and Environment for Statistical Computing. Vienna.

Scaltritti M, Peressotti F and Miozzo M (2017) Bilingual advantage and language switch: What's the linkage? *Bilingualism: Language and Cognition* **20** (1), 80–97. https://doi.org/10.1017/S1366728915000565

Snodgrass JG and Vanderwart M (1980) A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory* **6**, 174–215.

Stocco A and Prat CS (2014) Bilingualism trains specific brain circuits involved in flexible rule selection and application. *Brain and Language* **137**, 50–61. https://doi.org/10.1016/j.bandl.2014.07.005

UNESCO Institute for Statistics. (2012) *The International Standard Classification of Education 2011. Comparative Social Research*. Montreal: UNESCO Institute for Statistics. doi:10.1108/S0195-6310(2013)0000030017

Weiß RH and Osterland J (2013) *CFT 1-R Grundintelligenztest Skala 1 - Revision*. Göttingen: Hogrefe.

Yow WQ and Li X (2015) Balanced bilingualism and early age of second language acquisition as the underlying mechanisms of a bilingual executive control advantage: Why variations in bilingual experiences matter. *Frontiers in Psychology* **6**, 1–12. doi:10.3389/fpsyg.2015.00164