Multilingual Language Control and Executive Function: A Replication Study

Gregory J. Poarch *

Department of English Linguistics, University of Münster, Münster, Germany

Recent discussion has called into question whether navigating and controlling multiple languages in daily life influences the development of executive function. Given the dearth in replications of studies that have documented differences in executive function between multilingual and monolingual children, the present study replicates a study on executive function in children (Poarch and Van Hell, 2012a) with a child population from the same educational and socio-economic background. Two executive function tasks (Simon and Flanker) were administered to 163 children aged 5–13 years who were either monolingual second language (L2) learners of English or multilinguals [German-English bilinguals or German-Language X bilingual third language (L3) learners of English]. While the Simon task yielded no differences between groups, the Flanker task differed significantly across groups with multilinguals showing enhanced conflict resolution over L2 learners. While the children's performance on the two tasks yielded diverging results, the outcome is partially in line with the view that enhanced executive function in multilingual children arises from their permanent need to monitor, control, and shift between multiple languages. These findings are discussed against the backdrop of varying inhibitory processes invoked by the specific nature of the two tasks and of developmental trajectories of executive function.

Keywords: executive function, simon task, flanker task, second language learners, bilinguals, third language learners

INTRODUCTION

There is a growing body of research documenting that children who grow up with and regularly use multiple languages exhibit differential non-verbal executive function compared to children who only grow up and use one language. Such differences between multilingual and monolingual children are assumed to be linked to the lifelong multilingual experience of having to control and use multiple languages in daily life (for reviews, see Bialystok et al., 2012; Kroll and Bialystok, 2013; Baum and Titone, 2014; Valian, 2015; Bialystok, 2017; Poarch and Van Hell, 2017; Poarch, 2018). While there is ample experimental evidence in support of the notion that sustained and long-term multilingual experience positively affects executive function development in children (e.g., Carlson and Meltzoff, 2008; Poarch and Van Hell, 2012a; Poarch and Bialystok, 2015), there are now also studies that have yielded no executive function differences between multilingual and monolingual children (Duñabeitia et al., 2013; Antón et al., 2014; Gathercole et al., 2014).

Given these mixed findings, and in order to move forward in addressing the question of whether speaking more than one language on a regular basis indeed impacts the development of executive function, there is a need to replicate previous findings of executive function differences between groups in similar populations of children, and in doing so to possibly identify more specifically under which conditions multilingual children profit from their language control experience, and to assess whether the specific experimental measures used so far in the research field are unequivocally appropriate to adequately tap executive function processes. The present study attempts to address these issues by closely replicating a published study (Poarch and Van Hell, 2012a) with a very similar population from the same environment (extended to a larger age range) and using the same types of experimental measures. Such an approach is also warranted in light of the limited reproducibility of research in psychological science (Open Science Collaboration, 2015).

Executive Function and Multilingualism

Our cognitive system is geared toward making choices in daily life between alternative and competing responses (cf. Keye et al., 2009). The mechanism responsible for detecting situations in which such conflicting information is present, needs to be processed, and subsequently resolved is subsumed under the so-called executive function system. This system incorporates cognitive functions such as selective attention, updating information, shifting between sets of information, and monitoring for and resolving conflict (see, e.g., Botvinick et al., 2001; Engle, 2002; Miyake and Friedman, 2012; Diamond, 2013) and develops from early childhood until it reaches maturity during adolescence (Anderson, 2002). The theoretical basis of multilingualism affecting domain-general non-verbal cognitive processing is grounded in the finding that the processes subserving multilingual language control and non-verbal cognitive control show extensive overlap (Declerck et al., 2017; but see Calabria et al., 2015; Branzi et al., 2016, for evidence of less overlap) and that multilinguals need to cognitively control multiple competing languages and are exposed to nearly constant cross-language activation and interaction (e.g., during lexical processing; Thierry and Wu, 2007; Poarch and Van Hell, 2012b). Such control processes, which are also drawn on during bilingual language processing (e.g., Filippi et al., 2015) or when switching from one language to another (e.g., Anderson et al., 2018a), induce repetitive cognitive load that over time impacts the neural networks responsible for and subserving executive function (e.g., Calabria et al., 2018). These processes are also assumed to influence the development and efficacy of executive function (see Green and Abutalebi, 2013; for comprehensive reviews, see Bialystok, 2017; Antoniou, 2019).

There are numerous studies that have reported executive function differences between groups of multilingual and monolingual children matched on a variety of language and social background variables (e.g., Carlson and Meltzoff, 2008; Engel de Abreu et al., 2012; Morales et al., 2013; Blom et al., 2014, 2017; Ladas et al., 2015; Poarch and Bialystok,

2015; Crivello et al., 2016; De Cat et al., 2018; Thomas-Sunesson et al., 2018; for a review of research with children, see Poarch and Van Hell, 2017). The study most relevant to the present study, and the one described in detail at this point, is that by Poarch and Van Hell (2012a) who administered two executive function tasks (the Simon task and a variant of the Flanker task) to four groups of children (monolinguals, L2 learners, bilinguals, and trilinguals) aged 5–8. Bilinguals and trilinguals were defined as children who regularly used multiple languages (see Surrain and Luk, 2017, for how bilinguals are characterized in the literature). The study aimed to extend previous research that had compared only monolingual and bilingual children and to investigate executive function in children who were matched on proficiency in their first language (L1), on socio-economic status, while differing on language backgrounds and proficiency in their second language (L2). In the Simon task (Experiment 1), bilinguals and trilinguals showed significantly faster conflict resolution than monolinguals, and marginally so than L2 learners. Furthermore, bilinguals and trilinguals did not differ in their performance, and L2 learners and monolinguals did not differ either. The performance in the Flanker-type task yielded similar results, with bilinguals and trilinguals outperforming L2 learners in resolving conflict induced by the incongruent condition (see Description of tasks and measures below). Note that there was no monolingual participant group in Experiment 2. These findings were interpreted as indicating enhanced inhibitory control for bilinguals and trilinguals over L2 learners (and monolinguals in Experiment 1) stemming from the necessity for multilingual children to control their developing and interacting languages. Training language control processes regularly and repeatedly may boost the multilingual children's shifting of attention, task monitoring, and conflict resolution in these tasks. Alternatively, it may also modulate the impact of distracting information during task performance.

However, as indicated above, there are studies reporting no differences between multilingual and monolingual children in executive function task performance (Duñabeitia et al., 2013; Antón et al., 2014; Gathercole et al., 2014; Ross and Melinger, 2017). As such, the latter studies can be seen to challenge the assumption that multilingualism has an effect on the development of executive function and have fuelled the discussion on whether and how multilingual language experience can impact executive function (see, e.g., Poarch and Van Hell, 2017), and whether the executive function tasks used are ideally equipped to measure the efficacy of the executive function system (see Valian, 2015; Poarch and Van Hell, in press).

Description of Tasks and Measures

There are a number of experimental paradigms that tap non-verbal cognitive processes, two of which have been used ubiquitously in the field of research on multilingualism and executive function: the Eriksen Flanker task (1974) and the Simon task (Simon and Rudell, 1967). Both tasks are thought to induce cognitive conflict during task performance, requiring selective attention to identify conflict and subsequent cognitive resources for conflict resolution (see, e.g., Hommel, 2011;

Wöstmann et al., 2013), albeit in slightly different manners. While the Flanker task uses arrays of arrows that are either congruent or incongruent to measure resistance to the interference of flanking distractors (Friedman and Miyake, 2004), the Simon task uses colored squares to induce conflict by a spatial stimulus-response mismatch in incongruent trials compared to an absence of a mismatch in congruent trials.

Note that in Poarch and Van Hell (2012a) a modified and more elaborate version of the Flanker task was used, the Attentional Networks Task (ANT; Fan et al., 2002; Rueda et al., 2004). In essence, the ANT is a Flanker task (with the customary inhibitory control component that requires inhibiting distractors) with added executive function components, namely alerting and orienting. However, these additional components are disregarded in the present study in order to focus on the main question at hand as to whether multilinguals and monolingual differ in conflict monitoring and inhibitory control.

In both tasks, beyond inspecting overall reaction times in the congruent and incongruent conditions, a difference score as an index of inhibitory control is calculated (the congruent condition reaction time subtracted from the incongruent condition reaction time). The difference score magnitude indicates how strongly distracted individuals are in the incongruent condition compared to the congruent condition. A larger magnitude indexes poorer interference control (for a more detailed account of how performance in these tasks can be modeled, see Botvinick et al., 2001; Keye et al., 2009).

Finally, since the Simon task and the Flanker task are used to tap participants' conflict monitoring and inhibitory control, performance on both tasks can be expected to correlate positively. For conflict monitoring, overall processing speed across the tasks should correlate, for inhibitory control, performance on incongruent trials and the difference score should correlate across tasks (see, e.g., Keye et al., 2009; Wöstmann et al., 2013). In contrast, if there is no correlation across task performance, the tasks may not entirely tap the same executive function components (see Fan et al., 2003; Valian, 2015).

To the author's knowledge, there are only two studies with children that have correlated performance across the two tasks of interest (Ross and Melinger, 2017; Poarch and Van Hell, in press). Ross and Melinger (2017) found child bilinguals, bidialectals, and monolinguals to not differ on overall performance and the calculated difference scores in the Simon and Flanker tasks. Critically, the congruent and incongruent reaction times correlated significantly across tasks indicating convergent validity. The difference scores indexing inhibitory control were, however, not analyzed separately. In contrast, Poarch and Van Hell (in press) re-analyzed the data from their original study (Poarch and Van Hell, 2012a) and found that neither congruent and incongruent conditions nor the difference score correlated across tasks, which in turn calls the convergent validity across tasks into question (see also Paap and Greenberg, 2013). The inconsistent convergent validity for these executive function tasks (see also Keye et al., 2009) indicates that one (or both) of the tasks may not fully or only partially measure the efficacy of the conflict monitoring mechanism.

The Present Study

The main objective of the present study is to replicate previous work by Poarch and Van Hell (2012a), using the same task types and experimental set-up (Simon and Flanker), with a very similar population from the same environment (second language learners, bilinguals and bilingual third language learners), and in an extension, to also focus on children from a wider age range. Bilinguals and bilingual third-language learners are defined here as regular users of either two or three languages (Surrain and Luk, 2017). Both groups of children have been found to exhibit similar effects on executive function development compared to monolingual children, irrespective of the number of languages controlled on a daily basis (Poarch and Van Hell, 2012a; Poarch and Bialystok, 2015). Accordingly, similar executive function task performance by bilinguals and bilingual third language learners was expected. Hence, for the purpose of the present study, in the initial analyses the two groups were collapsed into a single group of multilinguals (subsequently, the two groups were also analyzed separately to confirm that their performance was indeed similar) and the following predictions were made:

1) If the cognitive effort of constantly controlling multiple languages has an effect on executive function development, then executive function task performance should differ between monolingual second language learners with little language control experience and multilinguals with more extensive language control experience.

2) The difference in performance between multilinguals and second language learners could be displayed by: (a) better overall performance by multilingual compared to second-language learners, which would amount to more efficient selective attention and task monitoring in executive function tasks in which participants are faced with congruent and incongruent stimuli. Such a performance difference has been interpreted as enhanced general-domain executive function in multilinguals (e.g., Martin-Rhee and Bialystok, 2008; Yang et al., 2011; Kapa and Colombo, 2013); and/or (b) a smaller difference between performance on congruent and incongruent stimuli, yielding a smaller difference score magnitude and better inhibitory control for multilinguals, which would amount to enhanced domain-specific executive function (e.g., Engel de Abreu et al., 2012; Poarch and Van Hell, 2012a; Poarch and Bialystok, 2015; Yang and Yang, 2016).

Furthermore, employing two tasks ubiquitously used in past research to tap cognitive processing in bilingual and monolingual children allowed for correlational analyses of the children's performance on the two tasks. Note that only very few studies so far have used these executive function tasks in children and, critically, have subsequently correlated task performance (Ross and Melinger, 2017; Poarch and Van Hell, in press).

MATERIALS AND METHODS

Participants

Participants were 163 children, 5- to 13-years old, who attended private primary and secondary German-English immersion

schools in Frankfurt, Germany. Four children were excluded due to incomplete data sets and/or background information. Thus, of the remaining 159 children, 77 children were German monolingual second language learners of English (henceforth L2 learners; 43 girls), 34 German-English bilinguals (12 girls), and 48 German-Language X third-language learners of English (henceforth L3 learners; 30 girls). The children's mean age was 9.7 years ($SD = 2.3$; range $= 5.2$–13.3 years).

Signed consent was provided by the children's parents[1], who also completed an earlier version of the Language and Social Background Questionnaire (LSBQ; Anderson et al., 2018b), in which the home language environment and proficiency in each language is assessed. The L2 learners were all native speakers of German and had been learning English for an average of 1.8 years ($SD = 1.5$). The bilingual children lived in homes in which German and English were the primary languages, with German being the main language outside the home, and German and English used at school. They had been learning English in educational contexts for an average of 3.0 years ($SD = 1.6$). The L3 learners spoke two languages at home (one of which being German), German and English at school, and had been learning English for an average of 1.8 years ($SD = 1.5$). The home languages spoken apart from German included Arabic (5), Croatian (2), Danish (2), Dutch (3), Eritrean (1), Greek (2), Hebrew (2), Hindi (1), Italian (4), Japanese (3), Lithuanian (1), Polish (3), Portuguese (2), Russian (4), Serbian (1), Spanish (5), Swedish (2), Turkish (3), Urdu (1), Vietnamese (1).

Parents were asked to rate their children's daily language usage on a set of 5-point scales that extended from "All German" (0) to "Only other language" (4). An average score of 2 indicates that home communication was divided equally between German and other languages. The mean score across these scales for L2 learners was 0.7 ($SD = 0.5$), for bilinguals it was 2.1 ($SD = 1.1$), and for L3 learners it was 1.9 ($SD = 0.9$), indicating that the monolinguals' homes functioned primarily in German, while those of the bilinguals and L3 learners showed a more balance use of German and English or German and another language, $F_{(2, 156)} = 55.75$, $p < 0.001$, with subsequent Tukey post-hoc analyses confirming the assumption that bilinguals and L3 learners did not differ significantly, $p = 0.48$, whereas both differed significantly from L2 learners, $ps < 0.001$. Parents' highest levels of education (on a 5-point scale: 1 = not completed high school to 5 = graduate or professional degree) were collapsed across both parents and used to index socio-economic status (SES). There were no differences between groups, $F < 1$, $p > 0.80$. Background measures are reported in **Table 1**.

As mentioned above, bilinguals and L3 learners were expected to perform similarly on the executive function tasks, and were thus subsumed under the label *multilinguals*. This resulted in subsequent comparisons of two instead of three groups.

---

[1]There is no ethics committee available for experimental studies conducted with human participants at the Faculty of Philology, University of Münster. The present study is in accordance with local legislation and the institutional requirements and follows the Code of Ethics "Rules of Good Scientific Practice" of the University of Münster (2002) and The European Code of Conduct for Research Integrity (European Federation of Academies of Sciences and Humanities, 2017).

TABLE 1 | Mean scores (and standard deviations) for background measures by language group.

	L2 learners (n = 77)	Bilinguals (n = 34)	L3 learners (n = 48)
Sex (female/male)	43/34	12/22	30/18
Age (years)	9.78 (2.25)	9.14 (2.29)	9.98 (2.28)
Length of Immersion (English)	1.83 (1.61)	3.04 (1.63)	1.79 (1.46)
Home language environment[a]	0.68 (0.51)	2.07 (1.08)	1.87 (0.89)
Parents' education[b]	3.46 (0.72)	3.52 (0.65)	3.53 (0.75)
PC usage (hours per day)	0.68 (0.54)	0.76 (0.54)	0.69 (0.50)
TROG-D standard score[c]	105.12 (12.31)	104.44 (17.09)	103.77 (16.78)
TROG-2 standard score	73.42 (14.69)	99.88 (11.71)	81.15 (16.20)
Raven's standard score	98.42 (11.42)	103.82 (12.50)	100.00 (10.21)

[a]Home language environment was quantified using a 5-point scale where 0, "All German"; 2, "Half German; half other language(s)"; 4, "Only other language."
[b]Education was quantified using a 5-point scale where 1, no high school diploma; 2, high school graduate; 3, some college or college diploma; 4, bachelor's degree; 5, graduate degree; score collapsed across parents.
[c]TROG-D German standard score calculated by transforming TROG-D T-score (range 20 to 80; M = 50, SD = 10) to TROG-2 English standard score (range of 55–145; M = 100, SD = 15). Formula for converting T-scores into standard scores: b = {[(a–a mean)/a SD] x b SD} + b mean, where a = T-score and b = standard score.

MATERIALS AND PROCEDURE

The background measures and experimental tasks were completed by the children in one session of approximately 45 min. First, one of the language proficiency tasks was administered, followed by one of the executive function tasks, the Raven's test, the other executive function task, and finally the other language proficiency task. The order of the language proficiency tasks and executive function tasks was counterbalanced. The children were informed before the experiment session began that they could choose to discontinue being tested at any time during the testing session. Each child was tested individually in a quiet room of their schools by a trained experimenter. Once the session was completed, the children received a small gift for the participation.

Background Measures
Test for Reception of Grammar
The Test for Reception of Grammar measures the receptive language proficiency of children. It was originally created by Bishop for English (TROG-2; Bishop, 2003), and is also available in revised and amended version for German (TROG-D; Fox, 2006). While the materials used in both test versions have some overlap, half the items are different. To counteract any spillover effects, the two tests were administered in a counterbalanced manner at the beginning and at the end of the test battery.

Raven's Colored Progressive Matrices
The Raven's CPM test (Raven et al., 1998) is a measure of non-verbal visuospatial reasoning. Participants are shown two arrays of colored pictures: one picture forms a pattern and a second one depicts potential components of the pattern. Participants must indicate the picture in the second array that best matches and

fits into the picture in the first array. Results are calculated as standard scores corrected for age.

Executive Functions Tasks

The executive function tasks were the Simon task (Simon and Rudell, 1967) and the Flanker task (Eriksen and Eriksen, 1974).

Simon Task

In the Simon task, the children see single colored squares on the computer screen and need to press a left or right button to indicate the color of the square. The position of each square on the screen renders a condition either congruent (e.g., a red-color square on the *left* calls for a left button press) or incongruent (e.g., a red-color square on the *right* calls for a left button press). Incongruent trials induce response conflict through a spatial stimulus-response mismatch, the resolution of which requires participants to draw on inhibitory processes for conflict resolution. In contrast, congruent trials with a spatial stimulus-response match induce no conflict. Each trial was initiated with a fixation cross at screen center 350 ms prior to stimulus onset, followed by a blank screen for 150 ms, after which the stimulus was displayed. Each stimulus remained on screen until a participant response or for a maximum of 3,000 ms. Before each next trial, an inter-trial interval of 850 ms ensued. All trials were counterbalanced with left/right responses. The experiment was presented in four blocks. First, there was a block of 12 practice trials to make participants familiar with the experiment. After this, there were three mixed blocks of 42 trials (14 central, congruent, and incongruent trials each), presented in a randomly generated order by the E-prime program.

Flanker Task

In the Flanker task, the children need to indicate the direction of a target arrow (pointing left or right) in the middle of an array of five arrows, using two buttons on a serial response box, Depending on which variant of the task is used, there are up to four types of trials. Baseline trials display a single arrow in the middle of the screen, while in neutral trials, two diamonds each flank the central arrow. These trial types were not used in the present study since they are sometimes reported in research but rarely analyzed (similarly to the central condition in the Simon task). Congruent trials show the flanking arrows pointing in the same direction as the target arrow, while incongruent trials have target and flanking arrows pointing in opposite directions. Each trial was initiated with a fixation cross at screen center 350 ms prior to stimulus onset, followed by a 150 ms blank, and then immediately by a stimulus. Each stimulus remained on screen until a participant response or for a maximum of 3,000 ms. The experiment was presented in five blocks. First, there was a block of 12 random congruent and incongruent practice trials to familiarize participants with the experiment. Then, there were four mixed blocks of 32 trials (16 congruent and 16 incongruent) presented in a randomly generated order by E-prime. Prior to each next trial, an inter-trial interval of 850 ms ensued. Only RTs of correct responses were included in the analysis.

By subtracting the performance in the congruent condition from that of the incongruent condition, a difference score indexing inhibitory control is calculated in both the Simon and the Flanker task. The magnitude of each difference score indicates the distraction by the induced conflict experienced by individuals. Larger difference scores indicate less efficient conflict resolution and interference control.

RESULTS

Results from the demographic background, German and English receptive grammar, and non-verbal intelligence measures are presented in **Table 1**.

T-tests comparing the two groups' scores for German and English receptive grammar and non-verbal intelligence showed no difference in either German receptive grammar, $p = 0.65$, or in non-verbal intelligence, $p = 0.11$, while the children did differ in English receptive grammar, $p < 0.001$. One-way ANOVAs comparing the three original groups confirmed these results: German receptive grammar, $p > 0.50$, non-verbal intelligence, $p > 0.10$, English receptive grammar, $p < 0.001$ (Tukey *post-hoc* comparisons, all $ps < 0.01$), with the bilinguals showing the highest scores, followed by the L3 learners, and the lowest scores by L2 learners.

Mean response times (RT) and mean accuracy rates were calculated for each condition of the two executive function tasks. Central trials in the Simon task were part of the experimental set-up; however, they are conventionally not compared in subsequent analyses and are thus not reported here.

Data Trimming Procedure

Incorrect responses (Simon: 3.9% for the congruent condition, 9.6% for the incongruent condition; Flanker: 1.8% for the congruent condition, 5.1% for the incongruent condition) were excluded from the RT analysis, as were outliers with RTs shorter than 200 ms (Simon: 0.6% for the congruent condition, 0.6% for the incongruent condition; Flanker: 0.7% for the congruent condition, 1.1% for the incongruent condition). Contrary to Poarch and Bialystok (2015), RTs above 2,000 ms were not considered outliers (see Zhou and Krott, 2016, for rationale; see also De Cat et al., 2018). RT and accuracy data for both tasks are presented in **Table 2**.

Simon Task Results

RTs and accuracies on the two critical trial types in the Simon task, the congruent and incongruent trials, were analyzed using repeated measures mixed ANOVAs with trial type (congruent and incongruent) as within-group variable and language group (L2 learners, multilinguals) as between-group variable, and given the substantial age range, age was entered as a covariate. The RT analysis yielded a significant main effect of trial type, $F_{(1, 156)} = 68.24$, $\eta^2 = 0.28$, $p < 0.001$, no significant effect of language group, $F_{(1, 156)} < 1.5$, $\eta^2 < 0.01$, $p = 0.23$, and a significant effect of age, $F_{(1, 156)} = 143.52$, $\eta^2 = 0.48$, $p < 0.001$. Furthermore, there was no significant interaction between trial type and language group, $F_{(1, 156)} < 1.1$, $\eta^2 < 0.01$, $p = 0.34$, and a significant interaction between trial type and age, $F_{(1, 156)} = 15.29$ $\eta^2 = 0.06$, $p < 0.001$, with the children in the middle age range showing less performance

TABLE 2 | Mean RT and accuracy scores (and standard deviations) in Simon and Flanker task by language group.

	Language Group			
	L2 learners	**Multilinguals**	**L2 learners**	**Multilinguals**
Condition/Task	**Simon**		**Flanker**	
RT				
(a) Congruent	592 (122)	577 (128)	713 (246)	670 (197)
(b) Incongruent	649 (129)	641 (142)	798 (269)	725 (198)
(c) Conflict (b–a)	57	64	85	55
ACCURACY				
(a) Congruent	0.965 (0.038)	0.959 (0.050)	0.981 (0.029)	0.980 (0.034)
(b) Incongruent	0.906 (0.075)	0.902 (0.071)	0.942 (0.076)	0.952 (0.055)
(c) Conflict (b–a)	0.059	0.057	0.039	0.028

overlap across groups than the younger and older children, who showed a large performance overlap in both conditions. The non-significant interaction between trial type and language group was confirmed by the similar conflict magnitudes (incongruent condition RTs—congruent condition RTs) for L2 learners (57 ms) and multilinguals (64 ms).

The accuracy analysis similarly yielded a significant main effect of trial type, $F_{(1, 156)} = 11.40$, $\eta^2 = 0.07$, $p < 0.001$, none of language group, $F_{(1, 156)} < 1$, $\eta^2 < 0.01$, $p > 0.50$, and a significant main effect of age, $F_{(1, 156)} = 10.37$, $\eta^2 = 0.06$, $p = 0.002$. Furthermore, there were no significant interactions, $Fs_{(1, 156)} < 1.3$, $\eta s^2 < 0.01$, $ps > 0.26$. The results indicate that the groups performed similarly overall (no domain-general executive function difference) and displayed similar effect magnitudes (no domain-specific inhibitory difference), and as such did not differ in resolving conflict in the Simon task.
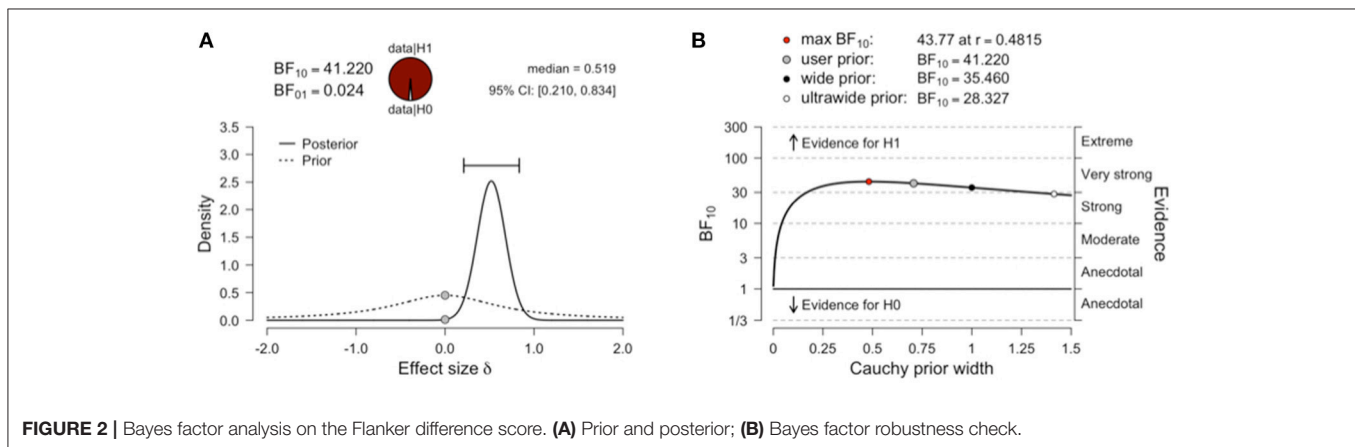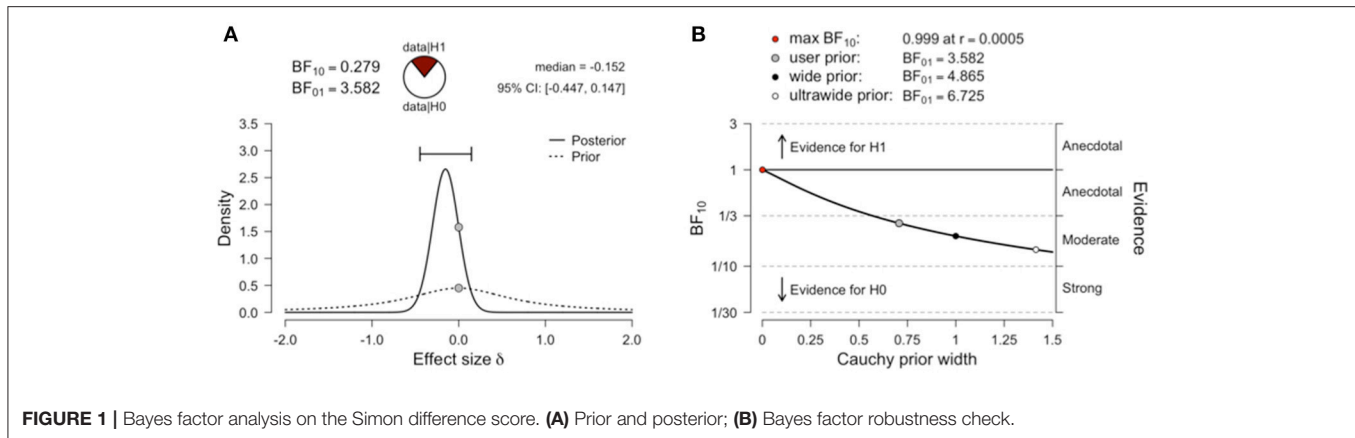
Flanker Task Results

Subsequently, performance on the congruent and incongruent trials of the Flanker task was analyzed in the same way as for the Simon task. The RT analysis yielded a significant main effect of trial type, $F_{(1, 156)} = 24.96$, $\eta^2 = 0.13$, $p < 0.001$, a main effect of language group, $F_{(1, 156)} = 7.77$, $\eta^2 = 0.02$, $p = 0.006$, and a main effect of age, $F_{(1, 156)} = 187.34$ $\eta^2 = 0.53$, $p < 0.001$. Furthermore, there was a significant interaction between trial type and language group, $F_{(1, 156)} = 12.59$, $\eta^2 = 0.06$, $p < 0.001$, but none between trial type and age, $F_{(1, 156)} < 1.7$, $\eta^2 < 0.01$, $p = 0.20$. The interaction between trial type and language group was further investigated through a separate one-way ANOVA on the conflict magnitudes, $F_{(1, 157)} = 12.19$, $\eta^2 = 0.07$, $p < 0.001$, showing a larger conflict for L2 learners (85 ms) than for multilinguals (55 ms). This result was further confirmed by comparisons of performance on congruent and incongruent conditions separately. While the groups did not differ significantly in the congruent condition, $F_{(1, 157)} = 1.44$, $\eta^2 < 0.01$, $p = 0.23$, they did so marginally in the incongruent condition, $F_{(1, 157)} = 3.76$, $\eta^2 = 0.02$, $p = 0.054$, which is assumed to have driven the significant main effect of language group.

The accuracies were at ceiling performance and the analysis yielded no main effect of trial type, $F_{(1, 156)} = 1.6$, $\eta^2 = 0.01$, $p = 0.21$, none of language group, $F_{(1, 156)} < 1$, $\eta^2 < 0.01$, $p = 0.56$, but a significant main effect of age, $F_{(1, 156)} = 11.82$, $\eta^2 = 0.07$, $p < 0.001$. Furthermore, there were no significant interactions, $F_{(1, 156)} < 1.9$, $\eta^2 < 0.02$, $p > 0.18$. As such, the results show no overall faster performance for multilinguals compared to L2 learners. However, they do indicate that multilinguals exhibit enhanced conflict resolution over L2 learners in the Flanker task and thus better domain-specific inhibitory control.

To tease apart whether the collapsed group of multilinguals also differed from the L2 learners when separated into the original two groups of bilinguals and L3 learners, a repeated measures mixed ANOVAs with trial type (congruent and incongruent) as within-group variable, language group (L2 learners, bilinguals, L3 learners) as between-group variable, and age as a covariate was conducted on the Flanker RTs only. The RT analysis yielded significant main effects of trial type, $F_{(1, 156)} = 21.93$, $\eta^2 = 0.12$, $p < 0.001$, of language group, $F_{(1, 156)} = 4.90$, $\eta^2 = 0.03$, $p = 0.009$, and of age, $F_{(1, 156)} = 190.32$, $\eta^2 = 0.54$, $p < 0.001$. Furthermore, there was a significant interaction between trial type and language group, $F_{(1, 156)} = 6.26$, $\eta^2 = 0.07$, $p = 0.002$, but none between trial type and age, $F_{(1, 156)} < 1.7$, $\eta^2 < 0.01$, $p = 0.21$. The interaction between trial type and language group was further investigated through a separate one-way ANOVA on the conflict magnitudes, yielding significant differences between groups, $F_{(1, 157)} = 6.09$, $\eta^2 = 0.07$, $p = 0.003$. Tukey *post-hoc* comparisons showed that bilinguals (57 ms) and L3 learners (54 ms) both resolved conflict significantly faster than L2 learners (85 ms), $p = 0.034$ and $p = 0.006$, respectively. Critically, bilinguals and L3 learners did not differ significantly, $p = 0.96$. The results confirm the two-group comparison above and indicate that bilinguals and L3 learners showed smaller effect magnitudes and were thus better at resolving conflict than L2 learners in the Flanker task.

Bayes Analyses

Finally, in an attempt to confirm the results obtained from the repeated measures ANOVA and to better adjudicate between the null hypothesis (H$_0$), which means that the groups of children did not differ significantly in their performance, and the alternative hypothesis (H$_1$), namely that the groups did indeed differ, Bayes factor analyses were performed (Wagenmakers et al., 2016) using JASP (JASP Team, 2018). Bayes factors indicate the weighted evidence either for or against specific effects of interest, which is displayed using BF$_{01}$ for evidence in favor of the null hypothesis (H$_0$) vs. BF$_{10}$ for evidence in favor of the alternative hypothesis (H$_1$) (for more detailed information on Bayesian inference, see Wagenmakers et al., 2018). For example, Bayes factors below 1 provide little evidence for the effects of interest, whereas Bayes factors above 30 provide very strong evidence for such effects (see **Figures 1**, **2**). For the difference score obtained in the Simon task, the Bayes factor with a BF$_{10}$ value of 0.28 indicated moderate evidence for the null hypothesis (see **Figure 1**), which means the difference scores across groups were similar. In contrast, for the Flanker task difference scores,

FIGURE 1 | Bayes factor analysis on the Simon difference score. **(A)** Prior and posterior; **(B)** Bayes factor robustness check.

FIGURE 2 | Bayes factor analysis on the Flanker difference score. **(A)** Prior and posterior; **(B)** Bayes factor robustness check.

there was strong to very strong evidence for the alternative hypothesis, indicating that the language groups differed, with a BF_{10} value of 41.22 (see **Figure 2**). The latter Bayes factor indicates that the data are 41.22 times more likely under H_1 than under H_0.

Correlational Analyses Simon Task and Flanker Task

The present study employed two executive function tasks that are customarily used to tap individuals' inhibitory control. Hence, one could hypothesize that performance on one task should correlate with that on the other (see Poarch and Van Hell, in press, for a more detailed rationale). To test this hypothesis, RT performance from both tasks on the congruent condition, the incongruent condition, and the resulting difference score (i.e., the conflict magnitudes also referred to as the Simon effect and the Flanker effect) were entered into a correlational analysis (see **Table 3**).

Within-Task Correlations

The Simon task congruent and incongruent conditions correlated significantly, $r = 0.95$ $p < 0.001$, as did the incongruent condition and the Simon effect, $r = 0.40$, $p < 0.001$, while the congruent condition and the Simon effect did not, $r = 0.10$,

$p = 0.20$. For the Flanker task, the congruent and incongruent conditions correlated significantly, $r = 0.97$ $p < 0.001$, as did the incongruent condition and the Flanker effect, $r = 0.37$, $p < 0.001$, and the congruent condition and the Flanker effect correlated marginally, $r = 0.14$, $p = 0.07$.

Cross-Task Correlations

The Simon and Flanker congruent conditions, $r = 0.71$, $p < 0.001$, and the incongruent conditions, $r = 0.70$, $p < 0.001$, correlated significantly. Critically, however, the Simon and Flanker effects did not correlate, $r = 0.07$, $p = 0.37$.

Finally, the Simon and Flanker effects were entered into a correlational analysis with the Home Language Environment score as an index for how multilingual the language environment of the children was outside of their educational context. While the Home Language Environment score and the Simon effect showed no significant correlation, $r = -0.06$, $p = 0.47$, the Home Language Environment score and the Flanker effect correlated significantly, $r = -0.16$, $p = 0.05$. Evidently, the more multilingual the children's environment was, the better they were at resolving conflict in the Flanker task, but not in the Simon task. Hence, the tasks may be tapping different components of executive function and inhibitory control (see Discussion for a more detailed interpretation). As Keye et al. (2009) have also

TABLE 3 | Pearson correlations between performance in task conditions and measures of inhibitory control.

Conditions and congruency effects	Simon			Flanker			
	1	2	3	4	5	6	7
1. Simon congruent	—	**0.954*****	0.102	**0.713*****	**0.688*****	0.077	−0.051
2. Simon incongruent		—	**0.396*****	**0.726*****	**0.703*****	0.092	−0.065
3. Simon effect			—	**0.225****	**0.228****	0.071	−0.061
4. Flanker congruent				—	**0.972*****	**0.142°**	−0.107
5. Flanker incongruent					—	**0.370****	**−0.137°**
6. Flanker effect						—	**−0.156***
7. Home lang. environment							—

Cross-correlations significant at $p < 0.05$*; at $p < 0.01$**; at $p < 0.001$***; and marginally significant at $p < 0.10$°; all in bold.

pointed out, the conflicts induced in both tasks are likely caused by more than one source of variance, which may make it is less likely to find a correlation of the conflicts across tasks.

DISCUSSION

The rationale for the present study was to explore whether the sustained cognitive control exerted on a daily basis by multilingual children in order to control their languages affects the development of their non-verbal executive function differently than that of monolingual children, and, in doing so, to replicate an earlier study by Poarch and Van Hell (2012a) with a very similar population living in the same language environment but extended to a wider age range. While the original study had focused on children aged 5–8, the present study tested 5- to 13-year old children. For this purpose, two executive function tasks were administered to the children to investigate whether their performance would differ across groups in their task monitoring (i.e., overall speed) and in their resolution of conflict (i.e., the difference score).

The Simon task data yielded no difference between groups, with multilinguals and monolinguals performing similarly both in overall speed and accuracy and in the obtained difference score. In contrast, the Flanker task showed that multilinguals and monolinguals differed significantly in their efficacy to resolve conflict, notably, and critically driven by differing performance in the incongruent condition, in that multilinguals displayed significantly smaller difference scores than monolinguals. While the Simon tasks results are not in line with those of the previous study, the Flanker results corroborate the earlier findings.

In light of these mixed findings, two issues will be highlighted and discussed in the following: (1) the nature of the population tested and matching of groups, and (2) the type of tasks used to tap executive function.

First, previous mixed findings have, amongst other explanations, been attributed to various factors inherent in

comparing groups experimentally, such as whether or not multilingual and monolingual children had been adequately matched on first language proficiency and socio-economic status (Paap et al., 2015). However, as Poarch and Van Hell (2017) have pointed out, the matching of children groups has not been overtly systematically different—in both research documenting differences between groups and that reporting null-results—to serve as a sufficient explanation for the mixed results (see also Baum and Titone, 2014; Bialystok, 2017). In the present study, the groups of children all attended private immersion schools, were meticulously matched on age, socio-economic status, fluid intelligence, PC usage, and L1 proficiency. The groups did differ, however, on the background variables L2 proficiency and home language environment, which are exactly those that could be assumed to differentiate multilinguals from monolinguals. Additional information on multilingual language usage patterns following the Adaptive Control Hypothesis by Green and Abutalebi (2013) may, in the future, offer a more fine-grained assessment of multilingual individuals and offer insight into within-group differences based on distinct contexts of multilingual interaction. According to Green and Abutalebi, single language, dual language, and dense code-switching contexts in a multilingual's life require differing degrees of cognitive control and thus also pose varying demands on the executive function system (see also Yang et al., 2016). However, for researchers to utilize such information, multilingual participants would need to be able to validly indicate which of these contexts pervade their lives. Moreover, a caveat to most research conducted in the field so far is that there are other lifestyle variables that have an effect on the development of executive function and may thus also influence performance on executive function tasks. Musical expertise (Peretz and Zatorre, 2005; Zuk et al., 2014; Schroeder et al., 2016) has been shown to be one of these variables, as has physical exercise (Best, 2010), dietary intake (Kim and Wang, 2017), circadian rhythm (Hahn et al., 2012), and sleep quality (Kuula et al., 2015). Future research could take all these additional variables into account and possibly an array of others (see Bak and Robertson, 2017), although the measurement of all of these may prove rather cumbersome in the scope of experimental research conducted in the field. What is striking, however, is that the effects on executive function of these diverse lifestyle variables seem to be less controversial than those of using multiple languages in daily life (cf. Bak, 2016).

Second, two prominent tasks in multilingualism research, the Simon and the Flanker task, have in the past been interchangeably and ubiquitously used to investigate executive function, and more specifically, conflict resolution, inhibitory control, and task monitoring. However, on closer inspection, the two tasks display differences in task demands that may inadvertently draw on both overlapping and non-overlapping subcomponents of executive function during task performance. According to Botvinick et al. (2001), both tasks can be described using the conflict monitoring and control theory, in which a conflict detector in the brain's ACC is triggered by a conflict signal. In the prefrontal cortex, control processes are then engaged to focus on relevant stimulus features in the task, which is stimulus location in the Simon task and feature dimension in

the Flanker task. Subsequently, stimulus-response compatibility is determined, upon which initiation of the correct response follows. While in both tasks, performance depends on whether the condition is compatible or incompatible, performance is modulated differently: in the Simon task through stimulus-response compatibility and bi-dimensional perceptual and motor conflicts, whereas in the Flanker task through stimulus–stimulus compatibility and uni-dimensional perceptual conflict (e.g., Keye et al., 2009; Ambrosi et al., 2016; see also Posner, 1980; Abrahamse and Van der Lubbe, 2008; Snyder et al., 2015). As such, these differences in how conflict is elicited may engage partially differing cognitive processes and induce varying cognitive loads during task performance, possibly also modulated depending on the age of the participant. Given the disparate developmental trajectories of the various executive function subcomponents in children (Anderson, 2002), one may adduce that children at varying ages may be differentially cognitively taxed during task performance of the Simon and the Flanker. The findings by Poarch and Van Hell (2012a), who found differences between groups in both the Simon and the Flanker for 5- to 8-year-old children, and the results of the present study with 5- to 13-year-old children, who differed only in the Flanker task, speaks to the effect of age on task performance and its development.

The correlational analyses conducted in the present study are thus informative as they indicate significant correlations across task conditions, corroborating the results reported by Ross and Melinger (2017), who found the performance of their groups of children to correlate across tasks (see also Poarch et al., 2018, for adults). However, in the present study, similarly to Poarch and Van Hell (in press), the difference score did not correlate significantly across tasks (see Kousaie and Phillips, 2012; Paap and Greenberg, 2013, for adults). The mixed findings from these correlational analyses are thus inconclusive as to whether these two measures of executive function tap conflict resolution, inhibitory control, and task monitoring similarly, which could be expected according to Miyake and Friedman (2012) if the same underlying cognitive processes were engaged during task performance. The present study's correlational results indicate the engagement of similar subcomponents of task monitoring across tasks (i.e., correlation of the two conditions) but separable subcomponents of inhibitory control (i.e., non-correlation of difference score). Furthermore, while the home language environment as an index of degree of multilingualism correlated significantly with the Flanker task difference score, this was not the case for that of the Simon task. The partially

diverging cognitive demands posed by the two tasks may thus be critical in whether or not differential performance emerges in multilinguals and monolinguals and whether their performance correlates (Macnamara and Conway, 2014; Ambrosi et al., 2016; Qu et al., 2016; for a more detailed discussion, see Poarch and Van Hell, 2017, and Poarch and Van Hell, in press). This may, all the more so, be the case for individuals such as children in whom executive function development is still ongoing (Anderson, 2002).

CONCLUSION

The present study aimed at replicating earlier research in a population from the same language environment using the same experimental design. The results offer partially corroborating evidence of systematic differences in executive function between multilingual and monolingual children aged 5–13. Given the debate on the findings in the executive function and multilingualism literature, culminating in titles such as "There is no coherent evidence for a bilingual advantage in executive processing" (Paap and Greenberg, 2013), the present findings partially replicate earlier findings and tentatively support the view that multilingualism indeed has an effect on executive function task performance, albeit depending on which tasks is used. The differing performance of the groups across tasks was hypothesized to be driven by factors such as differences in induced cognitive load and task complexity. Furthermore, differences in individuals' language backgrounds, language usage patterns, and other lifestyle variables may have a crucial impact on the course of executive function development in children (Baum and Titone, 2014; Van Hell and Poarch, 2014). Future research may want to draw on more sensitive measures of executive function and aim at testing children longitudinally to better trace the development of executive function over time.

AUTHOR CONTRIBUTIONS

GP conception, design, data collection, statistics, and writing.

ACKNOWLEDGMENTS

REFERENCES

Abrahamse, E. L., and Van der Lubbe, R. H. J. (2008). Endogenous orienting modulates the Simon effect: critical factors in experimental design. *Psychol. Res.* 72, 261–272. doi: 10.1007/s00426-007-0110-x

Ambrosi, S., Lemaire, P., and Blaye, A. (2016). Do young children modulate their cognitive control? sequential congruency effects across three conflict tasks in 5-to-6 year-olds. *Exp. Psychol.* 63, 117–126. doi: 10.1027/1618-3169/a000320

Anderson, J. A. E., Chung-Fat-Yim, A., Bellana, B., Luk, G., and Bialystok, E. (2018a). Language and cognitive control networks

in bilinguals and monolinguals. *Neuropsychologica* 117, 352–363. doi: 10.1016/j.neuropsychologia.2018.06.023

Anderson, J. A. E., Mak, L., Keyvani Chahi, A., and Bialystok, E. (2018b). The language and social background questionnaire: assessing degree of bilingualism in a diverse population. *Behav. Res. Methods* 50, 250–263. doi: 10.3758/s13428-017-0867-9

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychol.* 8, 71–82. doi: 10.1076/chin.8.2.71.8724

Antón, E., Duñabeitia, J. A., Estévez, A., Hernández, J. A., Castillo, A., Fuentes, L., and Carreiras, M. (2014). Is there a bilingual advantage in the ANT

task? Evidence from children. *Front. Psychol.* 5:398. doi: 10.3389/fpsyg.2014.00398

Antoniou, M. (2019). The advantages of bilingualism debate. *Ann. Rev. Linguist.* 5, 1–21. doi: 10.1146/annurev-linguistics-011718-011820

Bak, T. (2016). Cooking pasta in La Paz. *Linguist. Approach. Biling.* 6, 699–717. doi: 10.1075/lab.16002.bak

Bak, T. H., and Robertson, I. (2017). Biology enters the scene. a new perspective on bilingualism, cognition, and dementia. *Neurobiol. Aging* 50, 3–4. doi: 10.1016/j.neurobiolaging.2016.10.020

Baum, S., and Titone, D. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Appl. Psycholinguist.* 35, 857–894. doi: 10.1017/S0142716414000174

Best, J. R. (2010). Effects of physical activity on children's executive function: contributions of experimental research on aerobic exercise. *Dev. Rev.* 30, 331–551. doi: 10.1016/j.dr.2010.08.001

Bialystok, E. (2017). The bilingual adaptation: how minds accommodate experience. *Psychol. Bull.* 143, 233–262. doi: 10.1037/bul0000099

Bialystok, E., Craik, F. I. M., and Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends Cogn. Sci.* 16, 240–250. doi: 10.1016/j.tics.2012.03.001

Bishop, D. V. M. (2003). *The Test for Reception of Grammar, Version 2 (TROG-2).* London: Harcourt Assessment.

Blom, E., Boerma, T., Bosma, E., Cornips, L., and Everaert, E. (2017). Cognitive advantages of bilingual children in different sociolinguistic contexts. *Front. Psychol.* 8:552. doi: 10.3389/fpsyg.2017.00552

Blom, E., Küntay, A. C., Messer, M., Verhagen, J., and Leseman, P. (2014). The benefits of being bilingual: working memory in bilingual Turkish- Dutch children. *J. Exp. Child Psychol.* 128, 105–119. doi: 10.1016/j.jecp.2014.06.007

Botvinick, M. M., Braver, T. S., Barch, D., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 625–652. doi: 10.1037/0033-295X.108.3.624

Branzi, F. M., Calabria, M., Boscarino, M. L., and Costa, A. (2016). On the overlap between bilingual language control and domain-general executive control. *Acta Psychol.* 166, 21–30. doi: 10.1016/j.actpsy.2016.03.001

Calabria, M., Branzi, F. M., Marne, P., Hernandez, M., and Costa, A. (2015). Age-related effects over bilingual language control and executive control. *Biling. Lang. Cogn.* 18, 65–78. doi: 10.1017/S1366728913000138

Calabria, M., Costa, A., Green, D. W., and Abutalebi, J. (2018). Neural basis of bilingual language control. *Ann. N. Y. Acad. Sci.* 1426, 221–235. doi: 10.1111/nyas.13879

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

Crivello, C., Kuzyk, O., Rodrigues, M., Friend, M., Zesiger, P., and Poulin-Dubois, D. (2016). The effects of bilingual growth on toddlers' executive function. *J. Exp. Child Psychol.* 141, 121–132. doi: 10.1016/j.jecp.2015.08.004

De Cat, C., Gusnanto, A., and Serratrice, L. (2018). Identifying a threshold for the executive function advantage in bilingual children. *Stud. Second Lang. Acquis.* 40, 119–51. doi: 10.1017/S0272263116000486

Declerck, M., Grainger, J., Koch, I., and Philipp, A. M. (2017). Is language control just a form of executive control? Evidence for overlapping processes in language switching and task switching. *J. Mem. Lang.* 95, 138–145. doi: 10.1016/j.jml.2017.03.005

Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750

Duñabeitia, A., Hernández, J.A., Antón, E., Macizo, P., Estévez, A., Fuentes, L., et al. (2013). The inhibitory advantage in bilingual children revisited: myth or reality? *Exp. Psychol.* 61, 234–251. doi: 10.1027/1618-3169/a000243

Engel de Abreu, P. M., Cruz-Santos, A., Tourinho, C. J., Martin, R., and Bialystok, E., (2012). Bilingualism enriches the poor: enhanced cognitive control in low-income minority children. *Psychol. Sci.* 23, 1364–1371. doi: 10.1177/0956797612443836

Engle, R. W. (2002). Working memory capacity as executive attention. *Curr. Dir. Psychol. Sci.* 11, 19–23. doi: 10.1111/1467-8721.00160

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149. doi: 10.3758/BF03203267

European Federation of Academies of Sciences and Humanities (2017). *The European Code of Conduct for Research Integrity.* Berlin: ALLEA.

Fan, J., Flombaum, J. I., McCandliss, B. D., Thomas, K. M., and Posner, M. I. (2003). Cognitive and brain consequences of conflict. *Neuroimage* 18, 42–57. doi: 10.1006/nimg.2002.1319

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., and Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *J. Cogn. Neurosci.* 14, 340–347. doi: 10.1162/089892902317361886

Filippi, R., Morris, J., Richardson, F. M., Bright, P., Thomas, M. S. C., Karmiloff-Smith, A., et al. (2015). Bilingual children show an advantage in controlling verbal interference during spoken language comprehension. *Biling. Lang. Cogn.* 18, 490–501. doi: 10.1017/S1366728914000686

Fox, A. V. (2006). *TROG-D: Test zur Überprüfung des Grammatikverständnisses.* [Test for reception of grammar - German version]. Idstein: Schulz-Kirchner.

Friedman, N. P., and Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *J. Exp. Psychol. General* 133, 101–135. doi: 10.1037/0096-3445.133.1.101

Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Viñas Guasch, N., Jones, L. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschooler through older adults on card sorting, Simon, and metalinguistic tasks. *Front. Psychol.* 5:11. doi: 10.3389/fpsyg.2014.00011

Green, D. W., and Abutalebi, J. (2013). Language control in bilinguals: the adaptive control hypothesis. *J. Cogn. Psychol.* 25, 515–530. doi: 10.1080/20445911.2013.796377

Hahn, C., Cowell, J. M., Wiprzycka, U. J., Goldstein, D., Ralph, M., Hasher, L., et al. (2012). Circadian rhythms in executive function during the transition to adolescence: the effect of synchrony between chronotype and time of day. *Dev. Sci.* 15, 408–416. doi: 10.1111/j.1467-7687.2012.01137.x

Hommel, B. (2011). The Simon effect as tool and heuristic. *Acta Psychol.* 136, 189–202. doi: 10.1016/j.actpsy.2010.04.011

JASP Team (2018). *JASP (Version 0.9.0.1) [Computer Software].* Amsterdam.

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

Keye, D., Wilhelm, O., Oberauer, K., and van Ravenzwaaij, D. (2009). Individual differences in conflict-monitoring: testing means and covariance hypothesis about the Simon and the Eriksen Flanker task. *Psychol. Res.* 73, 762–776. doi: 10.1007/s00426-008-0188-9

Kim, J. Y., and Wang, S. W. (2017). Relationships between dietary intake and cognitive function in healthy Korean children and adolescents. *J. Lifestyle Med.* 7, 10–17. doi: 10.15280/jlm.2017.7.1.10

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

Kroll, J. F., and Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *J. Cogn. Psychol.* 25, 497–514. doi: 10.1080/20445911.2013.799170

Kuula, L., Pesonen, A. K., Martikainen, S., Kajantie, E., Lahti, J., Strandberg, T., and Räikkönen, K. (2015). Poor sleep and neurocognitive function in early adolescence. *Sleep Med.* 16, 1207–1212. doi: 10.1016/j.sleep.2015.06.017

Ladas, A. I., Carroll, D. J., and Vivas, A. B. (2015). Attentional processes in low-socioeconomic status bilingual children: are they modulated by the amount of bilingual experience? *Child Dev.* 86, 557–578. doi: 10.1111/cdev.12332

Macnamara, B. N., and Conway, A. R. A. (2014). Novel evidence in support of the bilingual advantage: influences of task demands and experience on cognitive control and working memory. *Psychon. Bull. Rev.* 21, 520–525. doi: 10.3758/s13423-013-0524-y

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

Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Curr. Dir. Psychol. Sci.* 21, 8–14. doi: 10.1177/0963721411429458

Morales, J., Calvo, A., and Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *J. Exp. Child Psychol.* 114, 187–202. doi: 10.1016/j.jecp.2012.09.002

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

Paap, K. R., and Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cogn. Psychol.* 66, 232–258. doi: 10.1016/j.cogpsych.2012.12.002

Paap, K. R., Johnson, H. A., and Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex* 73, 265–278. doi: 10.1016/j.cortex.2015.04.014

Peretz, I., and Zatorre, R. J. (2005). Brain organization for music processing. *Annu. Rev. Psychol.* 56, 89–114. doi: 10.1146/annurev.psych.56.091103.070225

Poarch, G. J. (2018). "How the use and control of multiple languages can affect attentional processes," in *The Interactive Mind: Language, Vision and Attention*, eds N. Mani, R. M. Mishra, and F. Huettig (Basingstoke: Macmillan), 53–60.

Poarch, G. J., and Bialystok, E. (2015). Bilingualism as a model for multitasking. *Dev. Rev.* 35, 113–124. doi: 10.1016/j.dr.2014.12.003

Poarch, G. J., and Van Hell, J. G. (2012a). Executive functions and inhibitory control in multilingual children: evidence from second language learners, bilinguals, and trilinguals. *J. Exp. Child Psychol.* 113, 535–551. doi: 10.1016/j.jecp.2012.06.013

Poarch, G. J., and Van Hell, J. G. (2012b). Cross-language activation in children's speech production: evidence from second language learners, bilinguals, and trilinguals. *J. Exp. Child Psychol.* 111, 419–438. doi: 10.1016/j.jecp.2011.09.008

Poarch, G. J., and Van Hell, J. G. (2017). "Bilingualism and the development of executive function in children: the interplay of languages and cognition," in *Executive Function: Development Across the Life Span (Frontiers of Developmental Science)*, eds S. Wiebe and J. Karbach (New York, NY: Taylor and Francis), 172–187.

Poarch, G. J., and Van Hell, J. G. (in press). "Does performance on executive function tasks correlate? Evidence fromchild trilinguals, bilinguals, and second language learners," in *Bilingualism, Executive Function, and Beyond: Questions and Insights. Studies in Bilingualism* (John Benjamins), 57.

Poarch, G. J., Vanhove, J., and Berthele, R. (2018). The effect of bidialectalism on executive function. *Int. J. Biling.* 1–17. doi: 10.1177/1367006918763132

Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 23, 3–25. doi: 10.1080/00335558008248231

Qu, L., Low, J. J. W., Zhang, T., Li, H., and Zelazo, P. D. (2016). Bilingual advantage in executive control when task demands are considered. *Biling. Lang. Cogn.* 19, 277–293. doi: 10.1017/S1366728914000376

Raven, J. C., Raven, J., and Court, J. H. (1998). *Raven's Colored Progressive Matrices.* San Antonio, TX: The Psychological Corporation.

Ross, J., and Melinger, A. (2017). Bilingual advantage, bidialectal advantage or neither? Comparing performance across three tests of executive function in middle childhood. *Dev. Sci.* 20:e12405. doi: 10.1111/desc.12405

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., et al. (2004). Development of attentional networks in children. *Neuropyschologica* 42, 1029–1040. doi: 10.1016/j.neuropsychologia.2003.12.012

Schroeder, S. R., Marian, V., Shook, A., and Bartolotti, J. (2016). Bilingualism and musicianship enhance cognitive control. *Neural Plast.* 4058620, 1–11. doi: 10.1155/2016/4058620

Simon, J. R., and Rudell, A. P. (1967). Auditory S–R compatibility: the effect of an irrelevant cue on information processing. *J. Appl. Psychol.* 51, 300–304. doi: 10.1037/h0020586

Snyder, H. R., Miyake, A., and Hankin, B. L. (2015). Advancing understanding of executive function impairments and psychopathology: bridging the gap between clinical and cognitive approaches. *Front. Psychol.* 6:328. doi: 10.3389/fpsyg.2015.00328

Surrain, S., and Luk, G. (2017). Describing bilinguals: a systematic review of labels and descriptions used in the literature between 2005–2015. *Biling. Lang. Cogn.* 1–15. doi: 10.1017/S1366728917000682

Thierry, G., and Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12530–12535. doi: 10.1073/pnas.0609927104

Thomas-Sunesson, D., Hakuta, K., and Bialystok, E. (2018). Degree of bilingualism modifies executive control in Hispanic children in the US. *Int. J. Biling. Educ. Biling.* 21. 197–206. doi: 10.1080/13670050.2016.1148114

University of Münster (2002). *Code of Ethics "Rules of Good Scientific Practise."* Available online at: https://www.uni-muenster.de/imperia/md/content/wwu/senat/pdf/kodex_en_22_11_2017.pdf

Valian, V. (2015). Bilingualism and cognition. *Biling. Lang. Cogn.* 18, 3–24. doi: 10.1017/S1366728914000522

Van Hell, J. G., and Poarch, G. J. (2014). How much bilingual experience is needed to affect executive control? *Appl. Psycholinguist.* 35, 925–928. doi: 10.1017/S0142716414000265

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., and Morey, R. D. (2018). Bayesian inference for psychology. Part I: theoretical advantages and practical ramifications. *Psychon. Bull. Rev.* 25, 35–57. doi: 10.3758/s13423-017-1343-3

Wagenmakers, E.-J., Morey, R. D., and Lee, M. D. (2016). Bayesian benefits for the pragmatic researcher. *Curr. Dir. Psychol. Sci.* 11, 19–23. doi: 10.1177/0963721416643289

Wöstmann, N. M., Aichert, D. S., Costa, A., Rubia, K., Moller, H. J., and Ettinger, U. (2013). Reliability and plasticity of response inhibition and interference control. *Brain Cogn.* 81, 82–94. doi: 10.1016/j.bandc.2012.09.010

Yang, H., Hartanto, A., and Yang, S. (2016). The complex nature of bilinguals' language usage modulates task-switching outcomes. *Front. Psychol.* 7:560. doi: 10.3389/fpsyg.2016.00560

Yang, S., and Yang, H. (2016). Bilingual effects on deployment of the attention system in linguistically and culturally homogeneous children and adults. *J. Exp. Child Psychol.* 146, 121–136. doi: 10.1016/j.jecp.2016.01.011

Yang, S., Yang, H., and Lust, B. (2011). Early childhood bilingualism leads to advances in executive attention: dissociating culture and language. *Biling. Lang. Cogn.* 14, 412–422. doi: 10.1017/S1366728910000611

Zhou, B., and Krott, A. (2016). Data trimming procedure can eliminate bilingual cognitive advantage. *Psychon. Bull. Rev.* 23, 1221–1230. doi: 10.3758/s13423-015-0981-6

Zuk, J., Benjamin, C., Kenyon, A., and Gaab, N. (2014). Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PLoS ONE* 9:e99868. doi: 10.1371/journal.pone.0099868