HHS Public Access
Author manuscript
J Exp Child Psychol. Author manuscript; available in PMC 2022 August 01.

Published in final edited form as:
J Exp Child Psychol. 2021 August ; 208: 105127. doi:10.1016/j.jecp.2021.105127.

Does Long-Term Dual-Language Immersion Affect Children's Executive Functioning?

Anne Neveu, **Kimberly Crespo**, **Susan Ellis Weismer**, **Margarita Kaushanskaya**
University of Wisconsin-Madison

Dual-Language Immersion (DLI) programs bring together native English speakers and English language learners and allow them to acquire language and literacy skills in two languages. Although DLI experience has been associated with enhanced academic and language outcomes in both the native (L1) and second language (L2) (Lindholm-Leary & Genesee, 2014; Marian, Shook, & Shroeder, 2013), less is known about whether DLI experience yields executive function (EF) advantages (Carlson & Meltzoff, 2008; Kalia, Daneri, & Wilbourn, 2019). In this study, we examined the effect of DLI experience on the development of executive function skills in majority-language children - native speakers of the dominant language in society, - to ensure group comparability with monolinguals, especially in terms of socio-economic status and native language skills. The bilingual children in our study were native-English-speaking children attending Spanish-English DLI programs in the US. We measured children's progress over the course of one year, to better capture the trajectory of EF development in intensive L2 immersion.

Dual-Language Development and Executive Function

Dual-Language Immersion (DLI) programs allow children to develop proficiency in a second language while maintaining and developing their skills in their native language and learning grade-appropriate content (Calderón & Minaya-Rowe, 2003; Cloud, Genesee, & Hamayan, 2000). Two-Way Immersion (TWI) is a type of DLI that brings together children who are native speakers of the dominant language in society, i.e. majority-language speakers, and learners of that language, i.e. minority-language speakers. In TWI, classroom makeup consists of a ratio of approximately 50:50 majority- and minority-language speakers, facilitating interactions with native-speakers of both languages. TWI leads to positive outcomes for both groups of children, in terms of native language development, second language acquisition and proficiency attainment, and academic achievement (Calderón & Minaya-Rowe, 2003; Esposito & Bauer, 2018; Genesee, 2004; Lindholm-Leary & Genesee, 2010; Lindholm-Leary & Genesee, 2014). However, intensive L2 exposure may also be associated with some linguistic challenges in the L1.

Correspondence concerning this article should be addressed to Anne Neveu, Language Acquisition & Bilingualism Lab, Waisman Center Rm 476, 1500 Highland Ave. Madison, WI 53705. aneveu@wisc.edu.

Declarations of interest: none.

Indeed, compared to monolinguals in non-immersion programs, majority-language students in 90:10 DLI programs (where children in the first year of the program receive 90% exposure in the minority language) develop their L1 reading, writing, speaking and listening skills at a slower rate (Genesee, 2004). A similar process has been observed in young adults, tested 3 months into foreign language immersion: while comprehension and production fluency increase in the L2, access to the L1 is reduced (Linck, Kroll, & Sunderman, 2009). The lag in L1 development in bilingual children might be explained by the fact that a bilingual's two languages compete for selection (Costa et al., 2003; Kroll, Bobb, & Hoshino, 2014; Marian & Spivey, 2003; Poarch & Van Hell, 2012), creating additional cognitive demands (e.g. Bialystok, 2009; Kalia et al., 2019; Poarch & Van Hell, 2012). A potential consequence of these increased demands in the linguistic domain that require children to practice inhibiting one language while using the other is a cognitive functioning advantage. In general, bilingual children frequently need to switch between their two languages and therefore need a control mechanism to activate the relevant language in context while 'tuning down' or inhibiting the other (e.g. Bialystok, 2007; Bialystok & Craik, 2010; Green, 1998). Inhibiting, shifting, switching, and monitoring are some of the processes that constitute executive functioning (EF). EFs can be defined as "general-purpose control mechanisms that modulate the operation of various cognitive subprocesses and […] regulate the dynamics of human cognition (Miyake et al., 2000, p. 50)." Simultaneous bilingualism - acquisition of two languages in parallel from birth or within a few years from birth (Kohnert, 2010) – may have an impact on children's executive functioning (e.g. Bialystok, 1999; Bialystok, 2010; Bialystok & Martin, 2004; Carlson & Meltzoff, 2008; Martin-Rhee & Bialystok, 2008), although the literature on bilingual EF advantages has become increasingly contentious (e.g. Morton & Harper, 2007; 2009; Namazi & Thordardottir, 2010; Paap & Greenberg, 2013).

Studies testing the effects of bilingualism on EFs often measure executive functioning using nonverbal tasks that require applying attentional control, to focus on relevant information to complete the task while inhibiting irrelevant information. Many tasks have been used to measure inhibitory control, which is both the ability to suppress a dominant response and the capacity to resist interference from distracting input (e.g., Tiego Testa et al., 2018). For example, simultaneous bilingual children aged 5 to 8 years old were shown to perform better than monolinguals on the nonverbal Simon task, which requires suppression of interfering information (e.g. children must press the left button when seeing the word 'left', even if the word is presented on the right side of the screen) (Martin-Rhee & Bialystok, 2008). In another nonverbal task, the flanker, children must focus on the direction of a target item at the center of a line of five items (e.g. fish or arrows), inhibiting the direction in which the neighboring items are pointing. These items may be pointing in the same direction (congruent trials) or in opposite directions (incongruent trials). Bilingual children aged 8 to 11 outperformed monolinguals on incongruent trials, suggesting better inhibitory control (Poarch & Bialystok, 2015). Simultaneous bilingual children have also been shown to outperform their monolingual peers on measures of task shifting. One such measure is the Dimensional Change Card Sort (DCCS) task (Zelazo, Frye, & Rapus, 1996) – a widely used omnibus executive control task that requires switching, shifting and monitoring rules to sort cards (Bialystok, 1999; Bialystok & Martin, 2004).

The DCCS involves three conditions: a pre-shift, in which participants know and apply one rule only; a post-shift, in which participants must apply the new rule; and a mixed condition, in which either of the two rules must be applied. From these condition, three cost measures can be derived: shifting costs, switching costs and mixing costs. Shifting costs are derived from the difference between the post-shift and the pre-shift condition and index the capacity to overcome rule perseveration (Frye, Zelazo, & Palfai, 1995). Switching costs are derived from the difference between switch and non-switched trials in the mixed condition and index the ability to switch back and forth between two rules (Prior & MacWhinney, 2010). Mixing costs are derived from the difference between non-switch trials in the mixed condition and pre-shift trials; they index the ability to monitor the application of a rule, while knowing more than one rule (Prior & MacWhinney, 2010). Evidence supporting a bilingual advantage on these cost measures has been found in several studies (e.g. Bialystok, 1999; Bialystok & Martin, 2004). For instance, in Okanda, Moriguchi, and Itakura (2010), bilingual children were more accurate than monolingual children on the shifting measure of the DCCS. Prior and MacWhinney (2010) showed that bilinguals exhibited lower switching costs, indicating that they were faster than monolinguals to respond to switch trials. Barac and Bialystok (2012) found that bilinguals exhibited lower mixing costs than monolinguals, suggesting a bilingual advantage on monitoring skills. However, bilingual advantages on the DCCS cost measures have not always been replicated (e.g. Morton & Harper, 2007; Paap & Greenberg, 2013). For example, Paap and Greenberg (2013) did not find any significant group differences between bilinguals and monolinguals on the same task used in Prior and MacWhinney (2010).

Similarly, mixed results have been found on the go/no-go task when comparing bilingual and monolingual children’s response inhibition abilities. The go/no-go paradigm requires participants to respond to stimuli with the exception of specific items, the “no-go” stimuli. Both the flanker and the go/no-go measure inhibition, albeit different aspects of it: the flanker measures the ability to resist interference, while the go/no-go measures the ability to inhibit prepotent responses (Bunge et al., 2002; Kaushanskaya et al., 2017). In a study using both behavioral and electrophysiological measures to assess potential differences between bilingual and monolingual five-year-old children on simple (gift delay) and complex response inhibition (go/no-go) tasks, bilinguals outperformed monolinguals on the go/no-go task, suggesting higher executive control capacity for the bilinguals (Barac, Moreno & Bialystok, 2016). However, this bilingual advantage was not found when comparing monolingual and bilingual groups of children who were on average 9 years old (Bonifacci et al., 2011).

Large-scale studies have found mixed evidence of the putative bilingual advantage on executive functioning. For example, in a study involving 252 monolinguals and 252 bilinguals matched on various knowledge and cognitive measures, children completed a verbal and a nonverbal Stroop task (Duñabeitia et al., 2014). Both tasks involved inhibiting irrelevant information when responding to stimuli. Findings showed that both groups performed similarly on each task, suggesting no differences in inhibitory control capacity across monolinguals and bilinguals. It has been suggested that inconsistencies in the degree to which bilingual and monolingual groups are matched on several knowledge and cognitive measures can explain some of the discrepant results in the bilingual EF literature.

Challenges in Comparing Simultaneous Bilinguals and Monolinguals

In a review of studies on bilingualism and non-linguistic executive function published between 2010 and 2014, Hilchey, Saint-Aubin and Klein (2015) conclude that the variability in sociolinguistic factors across monolinguals and bilinguals across the lifespan limits replication of studies and confidence in associating bilingualism with specific cognitive advantages. Similarly, Paap, Johnson and Sawi (2015) discuss factors that interact with measures of EFs in bilinguals, such as socio-economic status (SES) and cultural differences. For example, Engel de Abreu et al. (2012) found that bilinguals outperformed monolinguals on monitoring and inhibiting measures. However, the bilinguals lived in Luxembourg while the monolinguals lived in Portugal, which introduces confounds in terms of SES and cultural differences that might account for some of the bilingual advantage observed. Indeed, in two large-scale studies by Antón et al. (2014) and Duñabeitia et al. (2014), involving Basque-Spanish bilinguals of the same age and older than in Engel de Abreu et al. (2012), matched on SES and all native residents of Spain, no advantages were found for the bilinguals on executive control measures. Regarding cultural differences, a study by Carlson and Choi (2009) showed that when comparing Korean-English bilinguals living in the United States and American English monolinguals on six different measures of executive processing, bilinguals exhibited cognitive advantages. However, when these bilinguals were compared with Korean monolinguals, advantages disappeared, suggesting that cultural differences over and above mono- or bilingualism might be contributing the executive processing differences observed.

One way to limit these confounds is to compare children living in a similar environment, but who differ in their language experiences. Such an environment is created by DLI classrooms, which provide majority-language children with extensive L2 experience. In the present study, English-speaking children attending Spanish-English DLI classrooms and English-speaking children attending English-only classrooms all spoke English as their first language, came from the same town, attended the same schools, and shared a cultural background. Their families were highly comparable in their SES, and the children were highly similar in their English language skills. A comparison of these two groups on EF skills therefore affords a rare opportunity to test the effect of bilingual language experience on EFs outside of the socio-cultural and linguistic correlates of bilingualism inherent in testing simultaneous bilingual children (vis à vis monolingual children).

DLI and Executive Functions

The Adaptive Control Hypothesis (ACH) is useful in formulating specific predictions regarding how the language environment of majority-language speakers attending DLI might shape specific EFs (Green & Abutalebi, 2013). The ACH rests on the assumption that language comprehension and production require control processes. It follows that in the context of bilingual interactions, how the two languages are used poses variable demands on cognitive control, compelling language control processes to adapt to these demands. Green and Abutalebi identify eight types of control processes, which are: goal maintenance, conflict monitoring, interference suppression, salient cue detection, selective response inhibition, task disengagement, task engagement, and opportunistic planning. They delineate three contexts in which these control processes will adapt differently: the single-, dual- and code-switched language contexts. In a “dual-language context”, a bilingual’s two languages are used with different interlocutors, in the same environment (e.g., classroom or family). As such, language switching is more likely to take place between conversations rather than within an utterance. In the “single-language context”, a bilingual uses one language at home and one language outside the home, and therefore switching between languages will not be frequent. In the “dense code-switching context”, bilingual speakers frequently mix both languages within an utterance. The authors emphasize that goal maintenance, conflict monitoring, and interference suppression are all necessary for a bilingual to choose what language to speak. As a result, both the single- and dual-language contexts make demands on these processes over and above those imposed by the dense code-switching context. However, the dual-language context makes additional demands in that the bilingual speaker must detect salient cues to control which language to use with a given interlocutor. If the inappropriate language in context is selected due to a robust prepotent response, the speaker will need to apply selective response inhibition, to first trigger “task disengagement” (stop using the language) and subsequently generate “task engagement” (start using the other language). In the code-switching context, as dual-language use is expected at the utterance level, no such additional demands are made on these control processes. However, the speaker must apply “opportunistic planning” to adapt words from one language to fit in the frame of another language, a control process that is not readily needed in either of the two other contexts.

When considered from the perspective of the ACH, a DLI classroom for majority-language speakers in part resembles a dual-language context, because the majority-language child has the opportunity to speak the majority language with native speakers of the majority language, and to speak the minority language with native speakers of the minority language. It is unclear however whether children actually maintain such a separation of their languages outside the classroom. Furthermore, for the majority-language children, the exposure to the minority language happens only at school, with most children experiencing little input in the minority language in their homes. Thus, the DLI classroom also has characteristics of a single-language context for majority-language speakers. Given this conceptualization of DLI for majority-language speakers, who are likely to engage in minimal code-switching, it can be hypothesized that the EFs most likely to be impacted by their language experience would be monitoring, resisting interference and inhibiting prepotent responses.

However, similar to studies with simultaneous bilinguals, studies on DLI bilinguals, including majority- and minority language children, present mixed evidence regarding cognitive advantages for bilingual children as a result of immersion (Carlson & Meltzoff, 2008; Kalia et al., 2019). In a study comparing monolingual native English-speakers with both majority- and minority language speakers who had been in English-Spanish DLI for at least 9 months, children age 5 to 9 were administered two measures of executive function: the DCCS and a Lexical Stroop Sort task (Kalia et al., 2019). The DLI bilinguals outperformed the monolinguals on both EF tasks. In addition, there were no differences in performance on the EF tasks between minority- and majority language children in DLI (Kalia et al., 2019). In Poarch and Van Hell (2012), four groups of children (5–8 years old) were tested. Monolinguals were compared with simultaneous bilinguals, majority-language speaker DLI bilinguals (immersed on average for 1.3 years), and trilinguals on the Simon task and the Attentional Networks Task (ANT), measuring conflict detection and resolution ability. The simultaneous bilinguals and trilinguals, but not majority language DLI bilinguals, performed better on both EF tasks than the monolinguals, suggesting that a threshold of extended practice in inhibitory control must be passed to begin observing enhanced executive functioning in bi- or trilingual children.

Similarly, Carlson and Meltzoff (2008) compared monolinguals, simultaneous bilinguals and majority-language speaker DLI bilingual children (ages 4–6) who had been immersed for 6 months on a battery of EF tasks. After statistically controlling for verbal ability, SES and age, simultaneous bilinguals demonstrated significantly better performance on the EF tasks over the monolinguals and DLI bilinguals, who did not statistically differ from each other. In Kaushanskaya, Gross, and Buac (2014), monolingual children and majority-language speaker DLI bilinguals (immersed for approximately two years on average), ages 5–7, were compared on measures of word-learning, verbal short-term and working memory and a measure of nonverbal task-shifting. No differences were found between groups on the task-shifting and verbal short-term memory measures. However, DLI bilinguals outperformed the monolinguals on the verbal working memory and word-learning measures. These findings suggest that although exposure to the L2 in DLI may not translate to executive control advantages in children, it does support the development of verbal memory and word-learning ability.

Overall, these studies involved children who had been in immersion for six months up to three years, and less is known about how children perform in later years of the DLI. It is possible that the absence of EF advantages in DLI bilinguals who are majority-language speakers observed in prior studies is related to a relatively short period of time they have spent in the immersion classrooms. Therefore, in the present study, we examined how older children with more DLI experience (four years) might perform on executive function tasks. We focused on the majority-language group to avoid the issues previously discussed involved in matching bilingual and monolingual groups. In our sample, native speakers of the dominant language in DLI, the majority-speakers, were similar to their monolingual peers in all aspects but classroom environment.

Current Study

We examined the effect of intensive second language (L2) exposure on the development of inhibiting, shifting, switching and monitoring functions in children who had been in dual-language immersion for an average of four years at the beginning of the study. We focused on majority-language speakers in DLI in order to avoid the difficulty of matching children on socioeconomic status or language proficiency. We tested children twice, over the period of one year, to examine the possibility that children's maturation may contribute to differential progress in executive functioning depending on group – monolingual or bilingual.

To assess the potential impact of dual-language immersive experience on executive functioning, we tested children in the 8 to 10 years-old age range. Within this age range, we expected children's language functioning to largely stabilize, and children to gain a significant degree of Spanish experience and skill. At the same time, their executive functions would still be maturing (Anderson, Jacobs, & Anderson, 2010; Stuss, 1992), allowing for the possibility of movement in the EF skills as the result of language experience. The children in the two groups (DLI and mainstream classrooms) were native speakers of English, characterized by highly similar SES, ethnicity, language, and cultural profiles. At the time of the study, the bilingual children had been immersed in the DLI program for approximately 4 years. We predicted that if intensive L2 exposure in DLI affects performance on executive function measures, then DLI bilinguals would perform better than monolinguals on these measures. Moreover, we examined the developmental trajectory of EFs in this age range and predicted that both groups would perform better on executive function skills over time.

The benefit of the longitudinal design is its potential to capture group differences that may be moderated by children's maturation levels and cognitive skills. EF tasks are notoriously sensitive to participants' level of cognitive functioning, and the same task may yield optimal levels of performance in one age group, and lead to floor effects in another age group. We considered the possibility that one potential reason for murky findings within the bilingual EF literature might be the different degree of sensitivity of the same EF task at different ages and levels of cognitive maturity. We therefore tested children twice over a 1-year period, expecting that should our particular EF tasks be more or less sensitive to bilingual experience at different ages, we would be better able to capture this effect if we tested children longitudinally. The theoretical ramification of the longitudinal design is that it enables testing of the persistence of the bilingual effect on EFs. If DLI experience affords a lasting effect on EFs, we should observe bilingual advantages on our EF tasks at both testing time points. We thus conducted a one-year longitudinal study to capture the potential long-term effects of DLI immersion on children's EF skills. The children in the DLI programs became bilingual as a result of four years in DLI, and experienced five years in DLI by the end of the study.

We focused on the EF measures that have played a central role in the previous bilingual EF literature – inhibitory control (as indexed by flanker and go/no-go tasks) (Gunnerud et al., 2020), and task-shifting (as captured by the different cost indexes of the DCCS task). These particular EFs are expected to be salient in a DLI context when considered within the theoretical framework of the ACH (Green & Abutalebi, 2013). We did not make distinct predictions for each of the EF measures. Rather, we were interested in whether the effects of DLI experience would be specific to any one of these EFs, or whether they would generalize across tasks. In general, we hypothesized that children in the DLI must consistently inhibit one of their languages (since the two languages are present sequentially in the classroom and since code-switching is generally absent in DLI classrooms, at least in teachers' communication, e.g. Spooner & Arias Olsen, 2017; Wei & Martin, 2009). We therefore hypothesized that should DLI experience have an effect on EFs, this effect would be the strongest for the inhibitory control measures and for the monitoring aspect of EF, and less so for the shifting and switching measures.

Methods

Participants

Seventy monolingual children and fifty bilingual children were recruited from schools in Madison, WI for the initial testing session (Year 1). They were first pairwise matched on age (p = .68) and nonverbal IQ (p = .85). After these steps, 1 bilingual and 21 monolinguals could not be pairwise matched, leaving a sample of 49 monolinguals pairwise matched to 49 bilinguals. In this remaining sample, 7 pairs could not be included because one of the two matched participants did not come back in Year 2. In the remaining 42 pairs, a further 9 pairs were removed due to missing data from participants in one or both of the groups. Our analyses focused on the resulting groups, which included 33 monolingual children (17 females) with an average age of 9.17 years (SD = 1.03) and 33 English-Spanish bilingual children (15 females) with an average age of 9.27 years (SD = 0.94). We ran independent samples t-tests on demographic measures (age, mother's years of education and nonverbal IQ, and on English Core, Receptive and Expressive language), separately for monolinguals and bilinguals (to ensure that the larger sample of seventy monolinguals and fifty bilinguals did not differ from the smaller sample used in this study. For the bilingual sample, we additionally compared Spanish Core, Receptive and Expressive language and English and Spanish exposure. There were no significant differences across the original larger samples and the smaller matched samples on any of these measures, showing that the children who were selected for this study were largely representative of the overall sample.

Bilingual children were native speakers of English from English-speaking families and had been exposed to Spanish at age 4 or 5 through DLI, for an average of 4.14 years (SD = 1.09) at the first testing session. As a group, they did not experience Spanish exposure at home. The DLI program was structured following the 90:10 Spanish/English exposure at the time of testing, where 90% of classroom instruction was in Spanish, and 10% was in English during the first year (grade 1). This ratio evolved by 10% increment every year until 4^{th} grade, and by 5^{th} grade instruction was evenly split between English and Spanish. After four years in DLI, the average score on the *Test de Vocabulario en Imágenes Peabody* (*TVIP*, Dunn, Padilla, Lugo, & Dunn, 1986) was 80.79, indicating that the bilingual children were proficient users of Spanish. DLI bilingual children with more than 5% exposure to a third language during a typical week were excluded. Monolingual children were native speakers of English, with less than 5% exposure to another language during the week. These children did not receive any education in another language, and all of their very minimal exposure to another language took place in the community. For all children, the exclusionary criteria included a diagnosis of language impairment, learning disability, psychological or behavioral disorder, neurological impairment or other developmental disabilities. Both groups of children passed a hearing screening at 20dB, at 1000, 2000 and 4000 Hz.

We case-matched the monolingual and bilingual groups on age and IQ, indexed by the Wechsler Intelligence Scale for Children, 4th Edition (WISC-IV; Wechsler, 2003). At the first testing session, primary caregivers of all children filled out a background questionnaire about the child's family, medical and educational histories, including maternal years of education. We used maternal years of education as a proxy for socio-economic status (SES), as it is a proxy used in many studies (e.g. Barac, Moreno, & Bialystok, 2016; Ensminger & Fothergill, 2003; Miech, Essex, & Goldsmith, 2001), although we acknowledge that this metric does not fully capture the complex construct of SES. Independent sample t-tests showed that the groups did not significantly differ on SES and English Core, Receptive and Expressive language, indexed by the Clinical Evaluation of Language Fundamentals, 4th Edition (CELF-4; Semel, Wiig, & Secord, 2003.

All parents filled out the Language Experience and Proficiency Questionnaire (LEAP-Q), which probes for cultural identification and ethnicity information. With respect to cultural identification, 81.82% of the participants in the monolingual group and 90.91% of the participants in the bilingual group identified as US American. With respect to ethnicity, 84.85% of the monolinguals and 84.85% of the bilinguals identified as white. The breakdown by ethnicity across groups is shown in Appendix, Table A.5.

At the first visit, we additionally conducted a detailed interview of the bilingual children's dual-language acquisition, language development and language exposure. On average and during a typical week, bilingual children were exposed to English 75.34% of the time and to Spanish 24.66% of the time. Background information on all participants is presented in Table 1 and bilingual participants' characteristics are presented in Table 2.

Procedure

In both Year 1 and Year 2, testing sessions lasted between two and three hours and trained bilingual English-Spanish research assistants administered the Spanish standardized assessments to the bilingual children. Nonverbal executive function measures were administered both Year 1 and Year 2.

Inhibition.—We used two measures of inhibition, a child-appropriate flanker task and a go/no-go task. The flanker task has been widely used in previous research on inhibition skills in children (e.g., Barac, Bialystok, Castro, & Sanchez, 2014; Weintraub et al., 2013). Likewise, the go/no-go paradigm has been widely used to measure response inhibition (see Cragg & Nation, 2008 for a review).

In the flanker task, children must resolve a conflict by focusing on the direction of the middle stimulus while ignoring the surrounding stimuli. The child-friendly version of the task uses fish for the target stimuli, and seaweed for the stimuli surrounding the middle fish in the neutral condition. In the congruent condition, the middle fish is surrounded by four fish swimming in the same direction, two on each side of the target fish, whereas in the incongruent condition, the surrounding fish are swimming in the opposite direction as the target fish. The child is asked to hit the right or left button on a serial response box to indicate in which direction the middle fish is swimming. Stimuli are presented for a maximum duration of 1700ms, followed by a 1000ms inter-stimulus interval (ISI). Six untimed practice trials are provided with a fixation cross ("+" symbol) under the middle fish to help the child learn to orient to the middle target fish, followed by six timed practice trials without an orienting symbol under the middle fish. Nonverbal feedback is provided during the practice trials. Forty-eight timed test trials follow the practice trials. To develop a response habit, making inhibition more difficult, the majority of the trials (50%) are congruent. Incongruent and neutral trials each make up 25% of the remaining trials. An accuracy score is computed based on the number of wrong and omitted button presses. Reaction time (RT) is measured on correct responses.

In the go/no-go task, children have to respond to a target stimulus only, and refrain from responding to any non-target stimuli, using a serial response box. The child is instructed to respond as quickly and as accurately as possible. A fixation slide to orient the child's visual attention to the middle of the screen is presented for 550ms. Stimuli are then presented for up to 1300ms with a 1000ms ISI following a child's response. Children receive 8 practice trials (6 go trials and 2 no-go trials) with feedback, followed by 80 test trials. Go stimuli are presented for 75% of the trials. As for the flanker task, an accuracy score is computed based on the number of wrong and omitted button presses. Reaction time (RT) is measured on correct responses.

Task-Shifting.—To measure children's task-shifting ability, we used an adaptation of the widely used Dimensional Change Card Sort (DCCS) task (e.g. Bialystok & Martin, 2004; Frye, Zelazo & Palfai, 1995, Zelazo, Frye, & Rapus, 1996; Zelazo et al., 2003) and Bialystok and Martin's "color-shape game" (2004). To keep the task nonverbal, we omitted spoken cues and used patches of color or grey shapes at the top of the screen to cue sorting rules. The cues remained on the screen throughout the trial to reduce working memory demands. The stimuli consisted of circles and squares, either red or blue. We started the task by training children on the color dimension. If the child made more than one error during the four practice trials, practice was repeated. Children were then presented with 5 test trials on the dimension on which they had just been trained (this was the pre-switch condition). Children were then taught the new sorting rule (by shape instead of color) without practice and completed 5 test trials for shape (post-switch condition). In the mixed condition, which alternated between color and shape sorting rules, there were 30 pseudorandomized trials: 23 trials in the most recently used dimension and 7 interspersed trials that required children to shift to the other dimension. This setup resulted in 13 trials in which the child was asked to switch sorting rules and 17 trials in which the child was asked to use the same sorting rule as in the previous trial.

Cues in all three conditions (pre-switch, post-switch, and mixed) appeared at the top of the screen for 500ms, followed by the target stimulus (a red circle or a blue square) that appeared below the cue. Stimulus and cue remained on the screen for a maximum of 10s, during which time the child responded. At the bottom of the screen, the child saw a box on the left-hand side labeled with a red square and a box on the right-hand side labeled with a blue circle. The left and right buttons on the serial response box were correspondingly labeled with a red square and a blue circle. For the color sorting rule, the child was instructed to put all the red ones in the box with the red square by pressing the button on the left and to put all the blue ones in the box with the blue circle by pressing the button on the right. For the shape sorting rule, the child was instructed to put all the squares in the box with the red square by pressing the button on the left and to put all the circles in the box with the blue circle by pressing the button on the right. Children were instructed to respond as quickly and as accurately as possible. Each trial was followed by an 800ms ISI. We collected reaction times and accuracy measures and derived three cost variables indexing shifting performance: shifting costs (pre-shift minus post-shift trials); switching costs (non-switch minus switch trials in the mixed condition); and mixing costs (pre-shift minus non-switch trials in the mixed condition).

Analyses

Reaction time data were only analyzed for correct responses. RTs below 150 ms and RTs that were more than 2.5 SDs above or below the individual participant's mean were excluded. Trimmed RTs were log-transformed. Following these data-trimming procedures, an average of 4.5% (min: 4% – max: 5%) of trials were removed for the RT analyses across tasks.

Item-level data were used in all analyses. Linear mixed effects models were constructed to analyze RT data in R, version 3.2.2 (R Core Team, 2015) using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015). Logistic mixed effect models were constructed to analyze dichotomous accuracy data (0, 1) to examine the extent to which predictors increased or decreased the likelihood (log-odds) of making an accurate response. For models examining task shifting skills and flanker performance, fixed effects included Group (Monolinguals vs. Bilinguals), Time (Year 1, Year 2), and Condition as well as all lower-order two-way interactions, and a three-way interaction between Group × Time × Condition. The effect of Condition compared performance between neutral and incongruent trials for the flanker task, and pre-switch, post-switch, and mixing trials for the DCCS. By-subject random intercepts and by-subject random slopes were included in all models for the effects of Time and Condition, and the interaction between Time and Condition (Barr, Levy, Scheepers, and Tily, 2013). Some models failed to converge when fit with the maximal random effects structure; model convergence was achieved by reducing the random effect structure (Brauer & Curtin, 2018).

No-go accuracy was selected as the primary outcome variable for the go/no-go task. A recent principal component analysis (PCA) showed that no-go accuracy loaded highly on the inhibition construct of executive function (Kaushanskaya, Park, Gangopadhyay, Davidson, & Ellis Weismer, 2017). The sample in that study was very similar to the monolingual sample in the present study. The mixed models examining no-go performance included fixed effects of Group (Monolinguals vs. Bilinguals), Time (Year 1, Year 2), and their interaction (Group × Time). By-subject random intercepts and by-subject random slopes were included for the effects of Time. All dichotomous predictor variables (Group, Time, Condition) were contrast-coded (−0.5, 0.5) in all models.

Results

Inhibition

The linear mixed effect model constructed to examine conflict resolution via flanker RTlog data included 1469 observations. A significant main effect of Time was observed ($B = -0.03$, $SE = 0.01$, $t = -3.73$, $\beta = -0.15$), such that children were overall faster in Year 2 than Year 1. A significant effect of Condition was also observed, ($B = 0.03$, $SE = 0.005$, $t = 7.08$, $\beta = 0.15$), such that children were overall faster in the Neutral trials compared to the Incongruent trials. All other main effects and interactions were not significant (see Appendix Table A.6). The logistic mixed effects model examining the effects of group, time and condition on flanker accuracy revealed a significant main effect of condition ($B = -0.58$, $SE = 0.29$, $t = -2.01$, $\beta = -0.29$) such that children overall were more accurate in the Neutral trials compared to the Incongruent trials (Table A.1).

No-go accuracy was examined in a logistic mixed effect model that included 2480 observations (Table A.2). A significant interaction between Group and Time ($B = -0.78$, $SE = 0.32$, $z = -2.42$, $p = .02$, $\beta = -0.19$) was observed, such that in Year 1, bilingual children ($M = 0.87$, $SD = 0.09$) were significantly more likely to accurately inhibit a response on no-go trials than monolingual children ($M = 0.79$, $SD = 0.17$), (*Odds Ratio (OR)* = 0.46, *95% CI* = 0.25 – 0.86). However, in Year 2, monolingual children ($M = 0.86$, $SD = 0.13$) performed similarly to bilingual children ($M = 0.85$, $SD = 0.12$) (Figure 1).

Task Shifting

Accuracy.—The logistic mixed effect model constructed to examine *shifting skills* included 647 observations (Table A.3). A significant main effect of Condition was observed ($B = -1.53$, $SE = 0.45$, $z = -3.39$, $p < .001$), such that children were significantly more likely to be accurate in the pre-switch phase than the post-switch phase (*OR* = 0.22, *95% CI* = 0.09 − 0.52). All other main effects and interactions were not significant (see Appendix Table A.7).

A total of 3762 observations were included in the logistic mixed effect model constructed to examine *switching skills*. A significant main effect of Condition was observed ($B = -0.49$, $SE = 0.13$, $z = -3.73$, $p < .001$), such that children were significantly more likely to be accurate in the stay trials than the switch trials in the mixing phase (*OR* = 0.61, *95% CI* = 0.47 − 0.79). All other main effects and interactions were not significant (see Appendix Table A.7).

A total of 3521 observations were included in the logistic mixed effect model constructed to examine *mixing/monitoring skills*. A significant main effect of Condition was observed ($B = -1.79$, $SE = 0.24$, $z = -7.33$, $p < .001$), such that children were significantly more likely to be accurate during the pre-switch phase than in stay trials in the mixing phase (*OR* = 0.17, *95% CI* = 0.10 − 0.27). All other main effects and interactions were not significant (see Appendix Table A.7).

RT_{log}.—The linear mixed effect model constructed to examine shifting skills included 611 observations (Table A.4). A significant main effect of Time was observed ($B = -0.06$, $SE = 0.15$, $t = -3.83$, $\beta = -0.16$), such that overall, children were faster in Year 2 than Year 1. The results also yielded a significant main effect of Condition ($B = 0.14$, $SE = 0.01$, $t = 10.10$, $\beta = 0.40$), such that children were significantly faster in pre-switch trials compared to post-switch trials. All other main effects and interactions were not significant (see Appendix Table A.7).

Switching skills were also examined via a linear mixed effect model that included 3176 observations. The results yielded a similar pattern of results observed for shifting skills. A significant main effect of Time was observed ($B = -0.08$, $SE = 0.02$, $t = -5.47$, $\beta = -0.19$), such that overall, children were faster in Year 2 than Year 1. A significant main effect of Condition was also observed ($B = 0.03$, $SE = 0.01$, $t = 3.44$, $\beta = 0.05$), such that children significantly slowed down for switch trials in the mixing phase. A significant interaction between Time and Condition was also observed ($B = -0.03$, $SE = 0.01$, $t = -2.37$, $\beta = 0.04$), such that the difference in reaction time between neutral and incongruent trials decreased in Year 2 from Year 1. All other main effects and interactions were not significant (see Appendix Table A.7).

The linear mixed effect model constructed to analyze mixing/monitoring skills included 3100 observations. A significant main effect of Time was observed ($B = -0.07$, $SE = 0.01$, $t = -5.23$, $\beta = -0.15$), such that overall, children were slower in Year 1 than Year 2. A significant main effect of Condition was also observed ($B = 0.27$, $SE = 0.01$, $t = 25.06$, $\beta = 0.50$), such that children significantly slowed down in the mixing phase compared to performance in the pre-switch phase. All other main effects and interactions were not significant (see Appendix Table A.7).

Discussion

In the present study, we evaluated whether intensive exposure to two languages within the context of DLI might affect the developmental trajectory of executive functions in majority-language bilingual children aged 8–10 years, compared with monolingual children. The two groups of children in the present study did not significantly differ on age, nonverbal IQ and SES, and came from highly similar cultural and educational backgrounds. Therefore, this study represents a rather pure test of the effects that bilingual experience per se (rather than social and cultural variables that go hand-in-hand with bilingualism) has on nonverbal executive functions.

We tested both groups on two measures of inhibition, the flanker and go/no-go tasks, and on measures of shifting, switching and monitoring with the Dimensional Change Card Sort task. We found a significant interaction between group and year only on the go/no-go accuracy results, such that in Year 1, bilingual children were significantly more likely to accurately inhibit a response on no-go trials than monolingual children. However, in Year 2, monolingual children performed similarly to bilingual children. We interpret this finding to indicate that whatever effects of bilingualism acquired through DLI has on executive functioning, they are limited in scope and short-lived.

The tendency of DLI to affect response inhibition skills specifically could be explained by the fact that in the classroom, language use is regimented. By their fourth year in DLI, bilingual children are using both languages an equal amount of time in the classroom, as per the 90:10 model (90% Spanish, 10% English in kindergarten, with 10% respective decrement/increment each year, to reach a 50:50 ratio by grade four). It is possible that this even split between languages requires bilingual children to inhibit the language not in use with great enough frequency to engender a generalized effect on response inhibition skills, specifically. This interpretation would align with the Adaptive Control Hypothesis (Green & Abutalebi, 2013), whereby the language function adapts to the context of interaction. The findings also align with Poarch and Van Hell (2012) who found an inhibition advantage in bilingual and trilingual children aged 5 to 8 years old, but not in majority-language speaker DLI bilinguals. Poarch and Van Hell (2012) suggest that a specific threshold in language use and exposure might need to be attained before any advantages in inhibition are observed.

The fact that the response inhibition advantage was found in Year 1 but not in Year 2 might suggest that bilingual children's response inhibition skills reached peak levels earlier than for monolinguals. However, over the next year, bilingual children's response inhibition skills plateaued, allowing monolingual children's response inhibition skills to "catch up." This interpretation would be consistent with the patterns of results observed in Barac et al. (2016), where 5-year-old bilinguals outperformed monolinguals on go/no-go accuracy, and in Bonifacci et al. (2011), where this effect was not observed in 9-year-old children. Interestingly, response inhibition skills did not reach ceiling by Year 2 in either group of children. This indicates a need to follow the children over a longer period of time in future studies, as the current study cannot shed light on the subsequent trajectory of response inhibition development. It is possible that with stable bilingual exposure (the 50:50 ratio of English to Spanish remained constant over the year that this study took place), came stability in response inhibition skills. That is, the two groups may continue to follow the same developmental trajectory of response inhibition skills. It is also possible that bilingual exposure may fluctuate in its effects on the developing response inhibition system (for a yet unknown reason), following a non-linear pattern, such that with additional maturation, the groups may diverge once again in their response inhibition skills.

Whatever the explanation for the pattern of results we have observed for response inhibition, we did not find a similar pattern of results on the other measure of inhibition, the flanker task, which indexes resistance to interference. This finding of null results for the flanker task is in line with Duñabeitia et al. (2014) and Anton et al. (2014) who did not find significant group differences on inhibition skills in children and young teenagers. Although different tasks were used in these studies - Duñabeitia et al. (2014) used a verbal and a numerical Stroop task, and Anton et al. (2014) an Attentional Network Test – these tasks require resisting the interference of a prepotent response, similarly to the flanker task.

The lack of group differences in shifting, switching and monitoring skills is also in line with previous studies (Gathercole et al., 2014; Morton & Harper, 2007), although children in Morton and Harper (2007) were slightly younger (6–7 years). Our results are however at odds with findings from Bialystok (1999) and Bialystok and Martin (2004) who found a bilingual advantage on the DCCS. Their sample of children was younger, ranging from approximately 3 to 6 years old, and did not have English as a first language. It is therefore possible that in our sample, executive function skills assessed by the DCCS may have already plateaued in the age range tested and in the specific context of DLI. Based on the Adaptive Control Hypothesis, we can hypothesize that the less strict language division for native English DLI bilinguals, where English is spoken at home and both English and Spanish are spoken at school (combination of the single- and dual-language contexts) requires less shifting, switching and monitoring than in English language learners who speak their L1 only at home and their L2 only at school (Green & Abutalebi, 2013).

In interpreting our largely null findings, at least two cautionary notes must be added. First, while we matched our two groups on maternal years of education, using it as a well-known proxy for SES, we acknowledge that a more complex measure of SES that takes into account income levels and occupations may be better at capturing variability among children attending versus not attending DLI. Future studies would benefit from incorporating such a measure into their procedure. Second, the null results may reflect the relatively small sample sizes in our two groups of participants, and the possibility of being underpowered. We ran power simulations to isolate the effect of group in our models and although it is somewhat difficult to interpret post-hoc power analyses of item-level data, our simulations suggest that we may indeed be underpowered to detect an effect of bilingualism. However, studies that included much larger samples of bilingual and monolingual children than ours also failed to observe a significant effect of bilingualism on EFs (Anton et al., 2014; Duñabeitia et al., 2014). At the same time, studies that included similar or smaller sample sizes than ours have reported significant effects of bilingualism on EFs in childhood (e.g. Bialystok, 1999; Bialystok & Martin, 2004; Kalia et al., 2019). Therefore, while we fully acknowledge the importance of robust sample sizes, large sample sizes are not a guarantee of a significant group effect. In the future, we would aim to replicate our pairwise matching approach and analytical strategy that takes into account item effects, but in a much larger sample.

To conclude, we suggest that testing native English speakers in DLI and comparing them with monolingual English speakers is a more rigorous test of bilingual effects on EFs as the background of English language learners can vary on several dimensions, such as cultural background and SES. Indeed, in previous studies of the influence of bilingualism on executive functioning, differences in socio-economic status as well as in other variables, such as ethnicity, often emerge across groups of monolinguals and bilinguals, introducing variability. Therefore, although evidence of a relationship between bilingualism and executive functioning performance has been found (e.g. Bialystok, 2007; Bialystok & Craik, 2010; Green, 1998; Carlson & Meltzoff, 2008; Kalia et al., 2019), results from these studies are mixed (e.g. Bialystok & Majumder, 1998; Carlson & Meltzoff, 2008; Paap et al., 2015, 2018). Moreover, few involve children in DLI (Bialystok & Barac, 2012; Carlson & Meltzoff, 2008; Kalia et al., 2019), and of these the focus is on performance in the first years in immersion, leaving open the question of how children perform on EF tasks in later years of DLI. Our results indicate that by the fourth year of DLI, bilingual children show a modest advantage in response inhibition, but this advantage disappears by the fifth year. The general absence of DLI effects on EFs is consistent with other studies of DLI (Kaushanskaya, Gross, & Buac, 2014; Poarch & Van Hell, 2012) as well as with the broad literature (e.g. Bialystok & Martin, 2004; Duñabeitia et al., 2014; Hilchey, Saint-Aubin, & Klein, 2015; Paap et al., 2015; Valian, 2015), suggesting that overall, classroom dual-language immersion has a minimal impact on EFs.

The finding that DLI does not seem to affect EFs concerns a specific and indirect aspect of the DLI experience and must not overshadow the overall benefits of DLI. DLI leads to a host of positive outcomes for children of both majority- and minority-languages, including attaining high levels of L1 and L2 proficiency (Calderón & Minaya-Rowe, 2003; Genesee, 2004; Lindholm-Leary & Genesee, 2010; Lindholm-Leary & Genesee, 2014; Neveu, Gangopadhyay, Ellis Weismer, & Kaushanskaya, under review) and developing cultural awareness and competence, overall supporting academic achievement (Calderón & Minaya-Rowe, 2003). This should reassure parents, educators and policymakers on the demonstrated benefits of DLI. Moreover, DLI might positively affect the EFs of children who start at lower levels of L1 proficiency (e.g., children from lower SES backgrounds), suggesting that more research is needed to understand the specific contexts in which bilingualism might affect EFs (Engel de Abreu et al., 2012; Filippi & Bright, 2020).

Acknowledgements

This research was supported by NIH grants R01 DC011750, R01 DC016015 and U54 HD090256. The authors would like to thank the families who participated in this study. We also thank the Language Acquisition and Bilingualism Lab and the Language Processes Lab members for their assistance with participant recruitment, data collection, data coding, and valuable comments on this paper.

Appendix

Table A.1.

Flanker Accuracy and RTLog Models

	Accuracy		**RTlog**	
	***B* (*SE*)**	***z***	***B* (*SE*)**	***t***
Intercept	3.61 (0.24)	14.97***	2.78 (0.008)	357.00***
Group	0.43 (0.36)	1.19	0.02 (0.02)	0.97
Time	−0.43 (0.47)	−0.91	−0.03 (0.009)	−3.73***
Condition	−0.58 (0.29)	−2.01*	0.03 (0.005)	7.08***
Group X Time	0.54 (0.67)	−0.80	−0.03 (0.02)	−1.43
Group X Condition	0.05 (0.57)	0.08	−0.007 (0.01)	−0.73
Time X Condition	−0.30 (0.57)	−0.53	0.009 (0.01)	0.90
Group X Year X Condition	−1.61 (1.14)	−1.41	0.006 (0.02)	0.30
Observations	1,533		1,469	
N	66		66	
Akaike Inf. Crit.	532.19		−2,556.47	
Bayesian Inf. Crit.	590.87		−2,492.96	

Table A.2.

No-Go Accuracy Model

	Accuracy	
	***B* (*SE*)**	***z***
Intercept	1.93 (0.11)	17.68***
Group	0.21 (0.21)	0.99
Time	0.18 (0.17)	1.04
Group X Time	−0.78 (0.32)	−2.42*
Observations	2,480	
N	62	
Akaike Inf. Crit.	2,043.18	

	Accuracy	
	***B* (*SE*)**	***z***
Bayesian Inf. Crit.	2,083.89	

Table A.3.

DCCS Accuracy Models

	Shifting		Switching		Mixing	
	***B* (*SE*)**	***z***	***B* (*SE*)**	***z***	***B* (*SE*)**	***z***
Intercept	3.34 (0.34)	9.91***	1.75 (0.09)	19.58***	2.92 (0.15)	19.68***
Group	0.02 (0.49)	0.05	0.01 (0.17)	0.05	0.11 (0.29)	0.39
Time	0.28 (0.77)	0.37	0.14 (0.14)	1.00	0.04 (0.28)	0.14
Condition	−1.53 (0.45)	−3.39***	−0.49 (0.13)	−3.73***	−1.79 (0.24)	−7.33***
Group X Time	−0.53 (0.91)	−0.58	−0.40 (0.28)	−1.45	−0.62 (0.54)	−1.14
Group X Condition	−0.59 (0.90)	−0.65	−0.10 (0.25)	−0.41	−0.13 (0.49)	−0.26
Time X Condition	0.73 (0.91)	0.80	0.25 (0.22)	1.14	−0.04 (0.49)	−0.07
Group X Year X Condition	0.36 (1.80)	0.20	0.79 (0.41)	1.90	−0.30 (0.98)	−0.30
Observations	647		3,762		3,521	
N	66		66		66	
Akaike Inf. Crit.	281.66		3,139.41		2,402.12	
Bayesian Inf. Crit.	330.86		3,226.67		2,469.95	

Table A.4.

DCCS RTlog Models

	Shifting		Switching		Mixing	
	***B* (*SE*)**	***t***	***B* (*SE*)**	***t***	***B* (*SE*)**	***t***
Intercept	2.79 (0.01)	266.66***	3.01 (0.01)	226.19***	2.86 (0.01)	276.71***
Group	0.02 (0.02)	0.97	−0.002 (0.03)	−0.09	0.006 (0.02)	0.31
Time	−0.06 (0.01)	−3.83**	−0.08 (0.02)	−5.47***	−0.07 (0.01)	−5.23***
Condition	0.14 (0.01)	10.10***	0.03 (0.01)	3.44**	0.27 (0.01)	25.06***
Group X Time	−0.04 (0.03)	−1.38	−0.02 (0.03)	−0.75	−0.02 (0.01)	−0.62
Group X Condition	0.05 (0.03)	1.90	−0.01 (0.01)	−0.80	−0.007 (0.02)	−0.32
Time X Condition	0.004 (0.02)	0.17	−0.03 (0.01)	−2.37*	−0.002 (0.02)	−0.09
Group X Year X Condition	−0.04 (0.04)	−0.95	0.006 (0.03)	0.20	−0.02 (0.04)	−0.41
Observations	611		3,176		3,100	
N	66		66		66	
Akaike Inf. Crit.	−499.82		−2,336.91		−2,368.27	
Bayesian Inf. Crit.	−433.60		−2,221.70		−2,253.53	

Table A.5.

Participants' Cultural Identification and Ethnicity Information, as Reported in the LEAP-Q

	Count	Ethnicity	Percentage
Monolinguals	28	White/Caucasian	84.85
	2	Black/African American	6.06
	1	White/Caucasian & Black/African American	3.03
	1	White/Caucasian & Black/African American & Asian	3.03
	1	White/Caucasian & Asian	3.03
	33		
Bilinguals	28	White/Caucasian	84.85
	2	White/Caucasian & Asian	6.06
	1	Black/African American	3.03
	1	Asian	3.03
	1	White/Caucasian & American Indian/Native Alaskan	3.03
	33		

Table A.6.

Raw Means and Standard Deviations for Flanker Results by Condition

	Monolingual Mean (SD)	DLI Bilingual Mean (SD)
Raw RT Incongruent	646.41 (203.89)	676.73 (200.56)
Raw RT Neutral	591.06 (161.05)	626.33 (179.84)
Accuracy Incongruent	0.94 (0.23)	0.96 (0.21)
Accuracy Neutral	0.96 (0.19)	0.97 (0.18)

Table A.7.

Raw Means and Standard Deviations for DCCS Results by Condition

	Monolingual Mean (SD)	DLI Bilingual Mean (SD)
Raw RT Pre-Switch	545.90 (204.29)	631.15 (658.47)
Raw RT Post-Switch	760.61 (518.56)	812.38 (413.25)
Raw RT Mix	1168.23 (716.29)	1179.91 (747.63)
Accuracy Pre-Switch	0.97 (0.18)	0.98 (0.16)
Accuracy Post-Switch	0.90 (0.30)	0.90 (0.29)
Accuracy Mix	0.85 (0.36)	0.85 (0.35)

References

Anderson V, Jacobs R, & Anderson PJ (2010). Executive functions and the frontal lobes: A lifespan perspective. Taylor & Francis.

Antón E, Duñabeitia JA, Estévez A, Hernández JA, Castillo A, Fuentes LJ, Davidson DJ, & Carreiras M (2014). Is there a bilingual advantage in the ANT task? Evidence from children. Frontiers in Psychology, 5, 1–12. 10.3389/fpsyg.2014.00398 [PubMed: 24474945]

Barac R, Bialystok E, Castro DC, & Sanchez M (2014). The cognitive development of young dual language learners: A critical review. Early Childhood Research Quarterly, 29(4), 699–714. [PubMed: 25284958]

Barac R, Moreno S, & Bialystok E (2016). Behavioral and electrophysiological differences in executive control between monolingual and bilingual children. Child Development, 87(4), 1277–1290. [PubMed: 27133956]

Barac R, & Bialystok E (2012). Bilingual effects on cognitive and linguistic development: Role of language, cultural background, and education. Child Development, 83(2), 413–422. [PubMed: 22313034]

Barr DJ, Levy R, Scheepers C, & Tily HJ (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language, 68(3), 255–278.

Bates D, Maechler M, Bolker B, & Walker S (2015). lme4: Linear mixed-effects models using Eigen and S. (R package version 1.1–9.).

Bialystok E (1999). Cognitive complexity and attentional control in the bilingual mind. Child development, 70(3), 636–644.

Bialystok E (2007). Cognitive effects of bilingualism: How linguistic experience leads to cognitive change. International Journal of Bilingual Education and Bilingualism, 10(3), 210–223.

Bialystok E (2010). Global–local and trail-making tasks by monolingual and bilingual children: Beyond inhibition. Developmental Psychology, 46(1), 93. [PubMed: 20053009]

Bialystok E (2009). Bilingualism: The good, the bad, and the indifferent. Bilingualism, 12(1), 3–11. 10.1017/S1366728908003477

Bialystok E, & Craik FIM (2010). Cognitive and linguistic processing in the bilingual mind. Current Directions in Psychological Science, 19(1), 19–23. 10.1177/0963721409358571

Bialystok E, & Majumder S (1998). The relationship between bilingualism and the development of cognitive processes in problem solving. Applied Psycholinguistics, 19(1), 69–85. 10.1017/s0142716400010584

Bialystok E, & Martin MM (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. Developmental Science, 7(3), 325–339. 10.1111/j.1467-7687.2004.00351.x [PubMed: 15595373]

Bonifacci P, Giombini L, Bellocchi S, & Contento S (2011). Speed of processing, anticipation, inhibition and working memory in bilinguals. Developmental Science, 14(2), 256–269. [PubMed: 22213899]

Brauer M, & Curtin JJ (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. Psychological Methods, 23(3), 389–411. 10.1037/met0000159 [PubMed: 29172609]

Bunge SA, Dudukovic NM, Thomason ME, Vaidya CJ, Gabrieli JD (2002). Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. Neuron 33(2), 301–311. [PubMed: 11804576]

Calderon ME, & Minaya-Rowe L (2003). Designing and implementing two-way bilingual programs: A step-by-step guide for administrators, teachers, and parents. Corwin Press.

Carlson SM, & Choi HP (2009). Bilingual and bicultural: Executive function in Korean and American children. Talk presented at the 2009 biennial meeting of the Society for Research in Child Development, Denver, CO, April 2–4, SRCD, 2009.

Carlson SM, & Meltzoff AN (2008). Bilingual experience and executive functioning in young children. Developmental Science, 11(2), 282–298. 10.1111/j.1467-7687.2008.00675.x [PubMed: 18333982]

Cloud N, Genesee F, & Hamayan E (2000). Dual language instruction: A handbook for enriched education. Heinle & Heinle.

Costa A, Colomé À, Gómez O, & Sebastián-Gallés N (2003). Another look at cross-language competition in bilingual speech production: Lexical and phonological factors. Bilingualism: Language and Cognition, 6(3), 167–179. 10.1017/s1366728903001111

Cragg L, & Nation K (2008). Go or no go? Developmental improvements in the efficiency of response inhibition in mid-childhood. Developmental Science, 11(6), 819–827. [PubMed: 19046150]

Duñabeitia JA, Hernández JA, Antón E, Macizo P, Estévez A, Fuentes LJ, & Carreiras M (2014). The inhibitory advantage in bilingual children revisited. Experimental Psychology, 61(3), 234–251. [PubMed: 24217139]

Dunn LM.; Padilla ER.; Lugo DE.; Dunn LM. (1986). Test de Vocabulario en Imágenes Peabody. Circle Pines, MN: American Guidance Service.

Engel de Abreu PM, Cruz-Santos A, Tourinho CJ, Martin R, & Bialystok E (2012). Bilingualism enriches the poor: Enhanced cognitive control in low-income minority children. Psychological Science, 23(11), 1364–1371. [PubMed: 23044796]

Ensminger ME, & Fothergill K (2003). A decade of measuring SES: What it tells us and where to go from here? In Bornstein H & Bradley RH (Eds.), Socioeconomic Status, Parenting, and Child Development (pp. 13–27). Mahwah: Lawrence Erlbaum.

Esposito AG, & Bauer PJ (2018). From bench to classroom: Collaborating within a dual-language education model. Journal of Cognition and Development, 20(2), 165–181. 10.1080/15248372.2018.1483374

Filippi R & Bright P (2020). A developmental approach to bilingual research: The effects of multi-language experience from infancy to old age. Talk presented at the 2020 Conference on Multilingualism, Reading, UK, June 23–25, COM2020.

Frye D, Zelazo PD, & Palfai T (1995). Theory of mind and rule-based reasoning. Cognitive Development, 10(4), 483–527.

Gathercole VCM, Thomas EM, Kennedy I, Prys C, Young N, Guasch NV, et al. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. Frontiers in Psychology, 5(11).

Genesee FH (2004). What do we know about bilingual education for majority language students? In Bhatia TK & Ritchie W (Eds.), Handbook of Bilingualism and Multiculturalism (pp 547–576). Malden, MA: Blackwell.

Green D (1998). Mental control of the bilingual lexico-semantic system. Bilingualism: Language and Cognition, 1, 67–81.

Green DW, & Abutalebi J (2013). Language control in bilinguals: The adaptive control hypothesis. Journal of Cognitive Psychology, 25(5), 515–530. 10.1080/20445911.2013.796377 [PubMed: 25077013]

Gunnerud HL, Ten Braak D, Reikerås EKL, Donolato E, & Melby-Lervåg M (2020). Is bilingualism related to a cognitive advantage in children? A systematic review and meta-analysis. Psychological Bulletin, 146(12), 1059. [PubMed: 32914991]

Hilchey MD, Saint-Aubin J, & Klein RM (2015). Does bilingual exercise enhance cognitive fitness in traditional non-linguistic executive processing tasks. In Schwieter JW (Ed.), The Cambridge Handbook of Bilingual Processing (pp. 586–613). Cambridge University Press. 10.1017/CBO9781107415324.004

Kalia V, Daneri MP, & Wilbourn MP (2019). Relations between vocabulary and executive functions in Spanish-English dual language learners. Bilingualism, 22(1), 1–14. 10.1017/S1366728917000463

Kaushanskaya M, Gross M, & Buac M (2014). Effects of classroom bilingualism on task-shifting, verbal memory, and word learning in children. Developmental Science, 17(4), 564–583. 10.1038/jid.2014.371 [PubMed: 24576079]

Kaushanskaya M, Park JS, Gangopadhyay I, Davidson MM, & Weismer SE (2017). The relationship between executive functions and language abilities in children: A latent variables approach. *Journal of Speech, Language, and* Hearing Research, 60(4), 912–923.

Kohnert K (2010). Bilingual children with primary language impairment: Issues, evidence and implications for clinical actions. Journal of Communication Disorders, 43(6), 456–473. [PubMed: 20371080]

Kroll JF, Bobb SC, & Hoshino N (2014). Two languages in mind: Bilingualism as a tool to investigate language, cognition, and the brain. Current Directions in Psychological Science, 23(3), 159–163. 10.1177/0963721414528511 [PubMed: 25309055]

Linck JA, Kroll JF, & Sunderman G (2009). Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second-language learning. Psychological Science, 20(12), 1507–1515. 10.1111/j.1467-9280.2009.02480.x [PubMed: 19906121]

Lindholm-Leary K; Genesee F (2014). Student outcomes in one-way, two-way, and indigenous language immersion education. Journal of Immersion and Content-Based Language Education, 2(2), 165–180.

Lindholm-Leary K, & Genesse F (2010). Alternative education programs for English learners. Improving Education for English Learners, 323–382. 10.1108/S0732-1317(2011)0000020013

Marian V, & Spivey M (2003). Competing activation in bilingual language processing: Within-and between-language competition. Bilingualism, 6(2), 97.

Marian V, Shook A, & Schroeder SR (2013). Bilingual two-way immersion programs benefit academic achievement. Bilingual Research Journal, 36(2), 167–186.

Martin-Rhee MM, & Bialystok E (2008). The development of two types of inhibitory control in monolingual and bilingual children. Bilingualism, 11(1), 81–93. 10.1017/S1366728907003227

Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, & Wager TD (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. Cognitive Psychology, 41(1), 49–100. [PubMed: 10945922]

Morton JB, & Harper SN (2007). What did Simon say? Revisiting the bilingual advantage. Developmental Science, 10(6), 719–726. [PubMed: 17973787]

Namazi M, & Thordardottir E (2010). A working memory, not bilingual advantage, in controlled attention. International Journal of Bilingual Education and Bilingualism, 13(5), 597–616.

Neveu A, Gangopadhyay I, Ellis Weismer S, & Kaushanskaya M (under review). Immersion in dual-language programs does not impede children’s native language acquisition. Journal of Child Language.

Okanda M, Moriguchi Y, & Itakura S (2010). Language and cognitive shifting: Evidence from young monolingual and bilingual children. Psychological Reports, 107(1), 68–78E. [PubMed: 20923050]

Paap KR, Anders-Jefferson R, Mason L, Alvarado K, & Zimiga B (2018). Bilingual advantages in inhibition or selective attention: More challenges. Frontiers in Psychology, 9, 1409. [PubMed: 30158887]

Paap KR, & Greenberg ZI (2013). There is no coherent evidence for a bilingual advantage in executive processing. Cognitive Psychology, 66(2), 232–258. 10.1016/j.cogpsych.2012.12.002 [PubMed: 23370226]

Paap KR, Johnson HA, & 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. 10.1016/j.cortex.2015.04.014 [PubMed: 26048659]

Poarch GJ, & Bialystok E (2015). Bilingualism as a model for multitasking. Developmental Review, 35, 113–124. [PubMed: 25821336]

Poarch GJ, & 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(4), 535–551. 10.1016/j.jecp.2012.06.013 [PubMed: 22892367]

Prior A, & MacWhinney B (2010). A bilingual advantage in task switching. Bilingualism, 13(2), 253–262. 10.1017/S1366728909990526

Semel E, Wiig EH, & Secord WA (2003). Clinical evaluation of language fundamentals (4th ed.). Psychological Corporation.

Spooner M, Arias Olsen I (2017). Code-switching in the DLI classroom: Considering alternatives to target-language-only instruction. In Dejonge-Kannan A, Spicer-Escalante K, Abell ML, Salgado E (Eds.), Perspectives on Effective Teaching in DLI and Foreign Language Classrooms (pp. 16–29). Utah State University.

Stuss DT (1992). Biological and psychological development of executive functions. Brain and Cognition, 20(1), 8–23. [PubMed: 1389124]

Team RC (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Tiego J, Testa R, Bellgrove MA, Pantelis C, & Whittle S (2018). A hierarchical model of inhibitory control. Frontiers in Psychology, 9, 1339. [PubMed: 30123154]

Valian V (2015). Bilingualism and cognition: A focus on mechanisms. Bilingualism, 18(1), 47.

Wechsler D (2003). Wechsler Intelligence Scale for Children (4th ed.). Psychological Corporation.

Wei L, & Martin P (2009). Conflicts and tensions in classroom codeswitching: An introduction. International Journal of Bilingual Education and Bilingualism, 12(2), 117–122.

Weintraub S, Dikmen Sureyya S., Heaton Robert K., Tulsky David S., Zelazo Philip D., Bauer Patricia J., Carlozzi Noelle E., Slotkin Jerry, Blitz David, Wallner-Allen Kathleen, Fox Nathan A., Beaumont Jennifer L., Mungas Dan, Nowinski Cindy J., Jennif RCG (2013). Cognition assessment using the NIH Toolbox. Neurology, 80(11 Supplement 3).

Zelazo PD, Müller U, Frye D, Marcovitch S, Argitis G, Boseovski J, … & Carlson SM (2003). The development of executive function in early childhood. Monographs of the Society for Research in Child Development, i–151. [PubMed: 12875200]

Zelazo PD, Frye D, & Rapus T (1996). An age-related dissociation between knowing rules and using them. Cognitive Development, 11(1), 37–63. 10.1016/S0885-2014(96)90027-1

Highlights

- Dual-Language Immersion majority-language children were compared to monolinguals.
- Bilinguals initially better inhibited a response than monolingual children.
- After a year, monolingual children performed similarly to bilingual children.
- Effects of DLI on executive functioning are limited in scope and transitory.

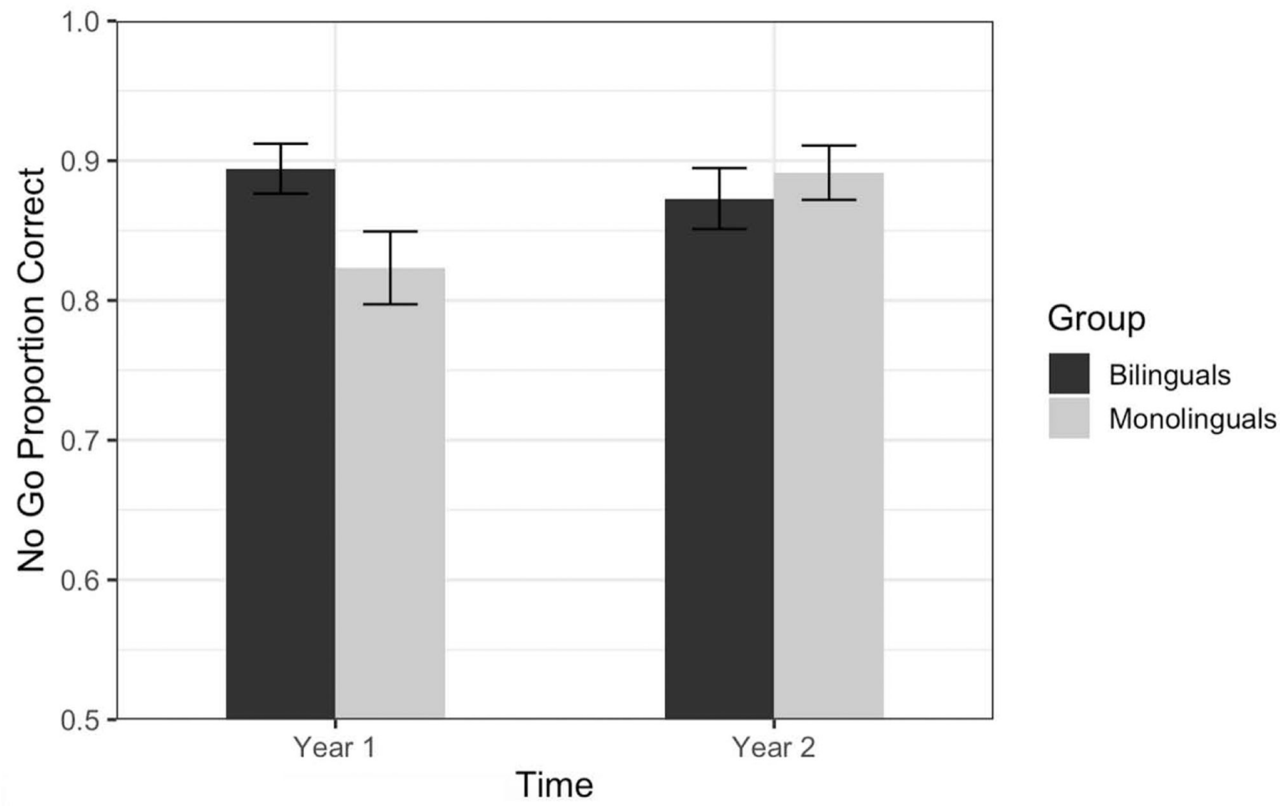

Figure 1.
Proportion of correct answers on the go/no-go task completed by bilingual and monolingual children, tested twice, one year apart. Error bars show standard errors.

Table 1.

Participant Characteristics in Year 1

	Monolingual Mean (SD)	DLI Bilingual Mean (SD)
Age (years)	9.17 (1.03)	9.27 (0.94)
N	33 (16M)	33 (18M)
Socioeconomic status[a]	17.06 (3.04)	16.98 (3.01)
Nonverbal IQ (WISC-IV)[b]	113.67 (13.33)	114.27 (11.79)
English Core Language (CELF-4)[c]	111.34 (10.90)	110.30 (11.41)
English Receptive Language (CELF-4)[c]	112.73 (12.54)	110.67 (13.01)
English Expressive Language (CELF-4)[c]	111.50 (11.34)	110.45 (12.77)
Spanish Core Language (CELF-4)[d]		86.48 (10.08)
Spanish Receptive Language (CELF-4)[d]		97.67 (12.47)
Spanish Expressive Language (CELF-4)[d]		81.00 (8.98)

[a]Indexed by total years of maternal education.

[b]Standard Score of Perceptual Reasoning Index from Wechsler Intelligence Scale for Children, 4th Edition.

[c]Standard Score from Clinical Evaluation of Language Fundamentals, 4th Edition.

[d]Standard Score from Clinical Evaluation of Language Fundamentals, Spanish 4th Edition.

Table 2.

Bilingual Characteristics in Year 1

	***n* = 33** Mean (SD)
Age of first Spanish Exposure (months)[a]	61.55 (8.67)
Amount of time in DLI (years)[b]	4.14 (1.09)
Dominant Language[c]	English: 97%, Both: 3%
English Exposure (%)[d]	75.34 (10.97)
Spanish Exposure (%)[e]	24.66 (10.97)
TVIP Score (Raw)	80.79 (10.78)

[a]Parent report: age when child started hearing Spanish.

[b]Parent report: age when child started dual-language immersion program.

[c]Parent report: child's dominant language.

[d][(Hours of English heard on a weekday × 5 days per week) + (Hours of English heard on Sat. & Sun.)] / (Hours child is awake per week)

[e]100 – percent of English exposure