






ARTICLE



<https://doi.org/10.1057/s41599-025-04360-7>

OPEN

# Language experience influences performance on the NIH Toolbox Cognition Battery: A cluster analysis

Ashley Chung-Fat-Yim <sup>1</sup>✉, Sayuri Hayakawa <sup>2</sup> & Viorica Marian <sup>1</sup>

Studies investigating the effects of bilingualism on cognitive function have often yielded conflicting results, which may stem in part from the use of arbitrary criteria to categorize participants into groups based on language experience. The present study addresses this limitation by using a machine learning algorithm, known as cluster analysis, to identify naturally occurring subgroups of participants with similar language profiles. In a sample of 169 participants with varying degrees of first- and second-language proficiencies and ages of acquisition, the cluster analysis yielded four bilingual subgroups: late-unbalanced, early-unbalanced, late-balanced, and early-balanced. All participants completed the NIH Toolbox Cognition Battery. Results revealed that early-balanced and early-unbalanced bilinguals scored higher than late-unbalanced bilinguals on the cognitive flexibility and inhibitory control subtests of the NIH Toolbox Cognition Battery, whereas late-unbalanced bilinguals scored higher than early-balanced bilinguals on the verbal working memory subtest of the NIH Toolbox Cognition Battery. Bilingual language experience did not impact performance on measures of processing speed, episodic memory, and English vocabulary. These findings demonstrate the utility of data-driven approaches to capture the variability in language experience that exists in the real world. We conclude that different bilingual experiences can shape a wide range of cognitive abilities, from working memory to inhibitory control.

<sup>1</sup>Northwestern University, Evanston, IL, USA. <sup>2</sup>Oklahoma State University, Stillwater, OK, USA. ✉email: [ashley.chungfatyim@northwestern.edu](mailto:ashley.chungfatyim@northwestern.edu)

## Introduction

Bilingualism is on the rise with more than half of the world's population using and speaking two or more languages on a regular basis (Marian, 2023). For example, in the United States, the number of dual-language programs has increased from 300 programs in 2001 to over 3600 programs in 2021 (Roberts, 2021). The language-learning app Duolingo reported over 500 million total users worldwide (Blanco, 2020). The impact of bilingualism on brain and cognitive development has been explored extensively (for reviews see Bialystok and Craik, 2022 and Marian et al. 2009). From better selective attention during infancy (Comishen et al. 2019; Kovács and Mehler, 2009) to delayed Alzheimer's symptoms in older age (e.g., Alladi et al. 2013; Bialystok et al. 2007), the cognitive consequences associated with bilingualism can be profound. However, several studies have challenged these claims by showing equivalent performance between monolinguals and bilinguals on some cognitive tasks (e.g., Paap and Greenberg, 2013; von Bastian et al. 2016). Considering that many linguistic variables contribute to defining and measuring bilingualism, the current study used a data-driven approach to categorize participants into groups based on age of acquisition and proficiency in the first and second languages. These categorizations were used to explore how variations in bilingual experiences influence performance on a variety of cognitive tasks, including executive function (inhibition and cognitive flexibility), language (vocabulary and reading), memory (working memory and episodic memory), and processing speed.

New approaches to studying the relationship between bilingualism and cognition have emerged alongside an evolving understanding of bilingualism. To compare performance between monolinguals and bilinguals on cognitive tasks, mathematical approaches like ex-Gaussian distribution analysis (i.e., examining response time distributions, Calabria et al. 2011; Tse and Altarriba, 2014; Zhou and Krott, 2018), drift-diffusion modeling (i.e., decomposing response times into parameters of decision and non-decision times, Ong et al. 2017; Soares et al. 2019), and confirmatory factor analysis (Anjomshoae et al. 2024) have been used. To quantify the variability of language experience in bilingual populations, researchers have implemented statistical techniques on various language measures, such as computing a factor score (DeLuca et al. 2019), examining bilingualism as a continuous factor (Dash et al. 2019; Dash et al. 2022), or applying data reduction approaches (e.g., aggregation, principal components analyses, or language entropy; Gullifer and Titone, 2020; Gullifer et al. 2021; Kałamała et al. 2022).

The present study aimed to implement a statistical technique known as cluster analysis (MacQueen, 1967) to identify different types of linguistic profiles. Through an unsupervised learning algorithm, cluster analysis classifies individuals with shared characteristics or minimal variability in profiles and assigns them into groups based on these criteria. For example, a subgroup of participants, such as those who had been exposed to two languages from an early age, may perform better on the cognitive flexibility measure by virtue of greater experience managing multiple languages than a subgroup of participants who acquired a second language more recently. Likewise, speakers with greater asymmetry in proficiency between their first and second languages may perform better on the inhibitory control measure because they often need to inhibit the stronger and more dominant language. Theoretically, the findings from the cluster analysis may help establish a link between the underlying cognitive process and language control. Thus, cluster analysis provides a more nuanced analytical approach for categorizing participants than relying on arbitrary thresholds in proficiency level, age of acquisition, or usage.

## Consequences of Bilingualism for Language and Cognition

Previous research has shown that bilingualism has consequences for language processing (Beatty-Martínez and Dussias, 2017; Sulpizio et al. 2020). Some studies have reported that bilinguals tend to know fewer words in each language (e.g., Bialystok et al. 2022), are generally slower to name objects on picture naming tasks (e.g., Gollan et al. 2005; Ivanova and Costa, 2008; Poarch and van Hell, 2012), and generate fewer words on verbal fluency tasks (e.g., Bialystok et al. 2008; Friesen et al. 2015; Sandoval et al. 2010) than their monolingual counterparts. Because bilinguals are dividing their time between multiple languages, the bilingual's weaker performance on verbal tasks may be attributed to their less frequent use of and exposure to each language compared to speakers of a single language. Additionally, bilinguals co-activate the lexicons from both of their languages even when only one language is needed (Marian and Spivey, 2003; Shook and Marian, 2019; Van Heuven et al. 1998; Wu and Thierry, 2010). Hence, the bilingual language system must manage competition between languages by selecting words in the intended language while suppressing or inhibiting those in the irrelevant language (Green, 1998). The amount of inhibition required to suppress the irrelevant language depends on the level of proficiency in each language. For bilinguals with a strong first language and weaker second language, the amount of inhibition needed to suppress the first language when operating in the second language will be larger than the amount of inhibition needed to suppress the second language when operating in the first language (Meuter and Allport, 1999).

Although the co-activation of competing languages may contribute to greater difficulty in some contexts (e.g., for word retrieval), the same mechanism has been proposed to underlie several reported benefits for cognitive function. Because bilinguals have two languages at their disposal for every communicative interaction, greater demands are placed on the fronto-parietal network to manage attention across two simultaneously active language systems. The predominant view is that language control in bilinguals is carried out by the executive control network (e.g., Anderson et al. 2018; Chen et al. 2021; De Baene et al. 2015). The accrued practice using executive control processes to ignore the irrelevant language while selecting the relevant language may enhance the domain-general executive control network for purposes beyond language.

Miyake and colleagues (2000) proposed a model of executive function consisting of three components: inhibiting prepotent responses, shifting between tasks, and updating working memory representations. On nonverbal executive control tasks of inhibition and shifting, bilinguals are generally faster and more accurate than monolinguals (see Bialystok, 2017 for a review). A wide range of linguistic factors, including the age of second language acquisition (Kapa and Colombo, 2013; Luk et al. 2011), second language proficiency (Rosselli et al. 2016; Tse and Altarriba, 2014; Videsott et al. 2012), and language control abilities (i.e., switchers versus non-switchers; Festman et al. 2010) have been shown to modulate executive control. In adults, language group effects are less pronounced on executive control tasks because accuracy tends to be at ceiling (Bialystok, 2017). Due to the lack of variability in performance and because both groups are operating at peak efficiency, bilingualism may not be associated with better performance on nonverbal measures of executive control in young adults.

Several meta-analyses have revealed that bilingualism was associated with better working memory abilities with a moderate effect size (Adesope et al. 2010; Grundy and Timmer, 2017; Monnier et al. 2022). Specifically, bilinguals generally achieve higher levels of performance than monolinguals when the

working memory task is *nonverbal* (e.g., Cockroft et al. 2019; Hansen et al. 2016; Luo et al. 2013), effortful (Antón et al. 2019; Jiao et al. 2019; Morales et al. 2013), and performed in the bilingual's native language (Grundy and Timmer, 2017). However, not all studies find that bilingualism leads to cognitive gains in working memory. Equivalent performance between language groups has been observed in children (McVeigh et al. 2019) and adults (D'Souza et al. 2018; Ratiu and Azuma, 2015) on working memory tasks.

Juggling multiple languages on a regular basis may enhance domain-general executive control, cascading into general improvements in processing speed, which has been described as the ability to identify, process, and formulate a response to information in a set amount of time (Kail and Salthouse, 1994). Donnelly and colleagues (2019) conducted a meta-analysis of 80 studies and found a small significant effect in favor of bilinguals for global response times (RTs) on nonverbal executive control tasks (e.g., Simon, Flanker, and Stroop tasks), with faster RTs for bilinguals than monolinguals on *both* congruent and incongruent trials. Furthermore, Luque and Morgan-Short (2021) found an association between degree of second-language proficiency in adult learners of Spanish and processing speed on the AX-CPT task, with greater proficiency in a second language associated with faster responses. Although some studies have shown that fluency in multiple languages influences how fast information is processed on executive control tasks, only a small number of studies have either included a pure measure of processing speed in the testing battery (Bonifacci et al. 2011; Ebert, 2021), accounted for processing speed in the analyses (Blumenfeld et al. 2016), or matched monolingual and bilingual participants on processing speed (Marton, 2016). Considering that the effect of bilingualism on pure measures of processing speed remains largely unknown, the current study examined how language experience affects processing speed.

Preliminary evidence suggests that bilingualism can additionally impact episodic memory. Episodic memory refers to the recollection of personal experiences, including details about the event's time and location. For bilinguals, episodic memory may include information about the linguistic context (Fernandez-Duque et al. in press; Schroeder and Marian, 2014). Bilinguals are more accurate at retrieving memories when the language at encoding is the same as the language at retrieval (Marian and Kaushanskaya, 2007). To date, only a few studies have directly compared bilinguals to monolinguals on episodic memory tasks (but see Ljungberg et al. 2013, 2020; Schroeder and Marian, 2012). For example, Schroeder and Marian (2012) found that bilinguals outperformed monolinguals on an episodic memory task, with earlier ages of second language acquisition and greater exposure to the second language leading to higher recall accuracy. These findings suggest that bilingualism can shape episodic memory.

### NIH Toolbox Cognition Battery

The NIH Toolbox Cognition Battery (Zelazo et al. 2013) assesses a wide range of cognitive abilities, including language, executive function, memory, and processing speed, in individuals from ages 3 to 85. The effects of bilingualism on performance in the NIH Toolbox Cognition Battery have been explored in various age populations, including children (Haft et al. 2019; Kuzyk et al. 2020; Nayak and Tarullo, 2020), adolescents (Dick et al. 2019; Kwon et al. 2021; Vaughn et al. 2021), and young adults (Marian et al. 2018a). In the young population, the differences between language groups on the individual subtests of the NIH Toolbox Cognition Battery appear to be small. Marian and colleagues (2018a) observed that bilinguals had higher scores on the Flanker

inhibitory control test, but lower scores on the List Sorting working memory test than monolinguals. Although unexpected given previous work showing a positive impact of bilingualism on working memory (e.g., Grundy and Timmer, 2017; Monnier et al. 2022), the lower scores on the List Sorting working memory test for bilinguals may have been due to their lower proficiency in English. Considering that the list sorting test involves recalling items in English, the verbal working memory task may be tapping into English comprehension or vocabulary knowledge rather than working memory.

### The Present Study

The present study uses cluster analysis to identify individuals with similar linguistic profiles. Clustering is an unsupervised machine learning algorithm for identifying subgroups of observations within a dataset (MacQueen, 1967). After mapping each participant within an  $n$ -dimensional coordinate space (e.g., by proficiency on the  $y$ -axis and age of acquisition on the  $x$ -axis), the algorithm begins by randomly selecting  $k$  centroid values from the dataset (corresponding to the specified number of clusters) and calculating the Euclidean distance between each remaining point and the closest centroid. The centroid is then reassigned as the mean of the cluster and the distance between each point and centroid is recalculated. Through an iterative process, the clusters and centroids are adjusted to minimize the distance between points within a given cluster and maximize the distance between points in different clusters. Ultimately, the goal is to identify subgroups of participants with similar language profiles to each other. This method has been implemented in almost every field from vision (e.g., Dhanachandra et al. 2015) to business (e.g., Anitha and Patil, 2022; Wang et al. 2005) to medicine (e.g., Juang and Wu, 2010; Khanmohammadi et al. 2017; Wismüller et al. 2002). In bilingualism research, cluster analysis has been used to identify language impairments in bilingual children (Hamann and Ibrahim, 2017) and to determine whether cognitive profiles can classify individuals as monolinguals and bilinguals (Jones et al. 2021). The results from the cluster analysis will uncover the types of cognitive resources recruited by bilinguals with similar linguistic profiles.

After identifying subgroups of participants, we will compare their performance on each subtest of the NIH Toolbox Cognition Battery that measure executive function (inhibitory control and cognitive flexibility), language (English reading and vocabulary), memory (episodic memory and working memory), and processing speed. For the cluster analysis, it is hypothesized that four clusters will emerge: early-unbalanced, early-balanced, late-balanced, and late-unbalanced bilinguals. Previous research has shown that bilinguals tend to perform worse than monolinguals on verbal measures (Bialystok et al. 2022; Gollan et al. 2005; Ivanova and Costa, 2008), but better on nonverbal executive control (e.g., inhibitory control and cognitive flexibility; Bialystok, 2017; Marian et al. 2009) and memory tasks (e.g., working memory and episodic memory; Grundy and Timmer, 2017; Monnier et al. 2022; Schroeder and Marian, 2012). Based on these findings, it is hypothesized that early-balanced bilinguals will perform worse than late-unbalanced-bilinguals and monolinguals on verbal measures, but perform better than both groups on nonverbal executive control and memory tasks. As proficiency increases in both languages (i.e., balanced bilingualism), the bilingual individual may need to recruit executive control to prevent intrusions from the nontarget language. Similarly, managing multiple languages from an early age (i.e., early bilingualism) allows for extensive practice in attending to and selecting the target language. We predicted that early bilingualism and more balanced proficiency would result in better

performance on nonverbal tasks of executive control, but worse performance on verbal tasks of executive control.

An open question, however, is the degree of interactivity between proficiency and age of acquisition. Although both variables were predicted to impact verbal and nonverbal cognitive function, few studies have examined the independent and interacting effects of earlier second language acquisition and more balanced language proficiency on cognitive performance within the same study and in young adult populations (Yow and Li, 2015). Thus, the examination of early-unbalanced and late-balanced bilinguals in addition to early-balanced and late-unbalanced bilinguals will make it possible to investigate whether proficiency and age of acquisition have additive or compensatory effects on cognitive function. For instance, if the *strength* of language co-activation and competition in daily life (which should increase with higher bilingual proficiency) and the *duration* of bilingual experience (i.e., age of bilingual acquisition) have independent positive effects on nonverbal executive control, we may expect an additive effect whereby early-balanced bilinguals outperform both early-unbalanced and late-balanced bilinguals. On the other hand, if the strength and duration of experience managing two languages each serve to strengthen nonverbal executive control via the same mechanism, we may expect a compensatory effect whereby either form of experience is sufficient to elicit improved performance (i.e., equivalent performance among early-balanced, early-unbalanced, and late-balanced bilinguals).

## Method

**Participants.** Data from 169 participants were compiled across four existing datasets from our lab (Chabal et al. 2022; Chen et al. 2017; Hayakawa et al. 2020; Marian et al. 2018b). These datasets were selected because they included data from the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al. 2007) and individual test scores from the NIH Toolbox Cognition Battery (Zelazo et al. 2013). The LEAP-Q was used to obtain information regarding each participant's language usage patterns as well as their level of proficiency, age of acquisition, and degree of exposure in each language. Participants rated their level of proficiency in each language on a scale from 0 (None) to 10 (Perfect) in speaking, understanding, and reading. They rated the extent to which several factors contributed towards learning each language (i.e., manner of acquisition) on a scale from 0 (Not a Contributor) to 10 (Most Important Contributor). The factors included interacting with friends, interacting with family, watching TV, reading, language tapes/self-instruction, and listening to the radio. Additionally, participants rated the extent to which they are exposed to each language (i.e., current exposure) on a scale from 0 (Never) to 10 (Always) in different contexts, including friends, family, TV, radio/music, reading, and language-lab/self-instruction. For each language, participants reported the number of years and months spent in a country, family, and school and/or work environment where the language is spoken (i.e., immersion).

Monolinguals were English native speakers with minimal proficiency in a second language (mean ratings of 0 or 1 out of 10 in second language proficiency averaged across speaking, understanding, and reading;  $N = 29$ ). The remaining 140 participants with self-reported second language proficiency of 2 or higher were proficient in English and had experience with one of the following languages: Spanish ( $N = 63$ ), Korean ( $N = 51$ ), French ( $N = 10$ ), Chinese ( $N = 4$ ), American Sign Language, German, Hebrew, Italian ( $N = 2$  each), Ikwerre, Polish, Russian, and Tagalog ( $N = 1$  each). Out of 140 bilinguals, English was the L1 (first language acquired) for 83 bilinguals and English was the L2

(second language acquired) for 57 bilinguals. On average, bilinguals rated their proficiency in their L1 as 9.35 ( $SD = 0.97$ ) and L2 as 6.28 ( $SD = 2.90$ ) out of 10. Monolinguals significantly differed from bilinguals on all L2 measures other than the duration immersion in an L2-speaking family. Monolinguals and bilinguals significantly differed on all L1 measures of proficiency, age of acquisition, current exposure, and immersion ( $ps < 0.05$ ). Sixty-one participants reported knowing three or more languages. All participants were adults between the ages of 18 and 36. Monolinguals and bilinguals did not differ in gender ( $M_s = 81.5\%$  and  $72.1\%$  female, respectively;  $p = 0.32$ ), age in years ( $M_s = 22.64$  and  $21.56$ , respectively;  $p = 0.13$ ), or years of education ( $M_s = 15.79$  and  $15.07$ , respectively;  $p = 0.10$ ). Language background information are reported for each language group in Table 1.

Participants provided consent and were debriefed about the objective of the study at the time of testing. The individual studies included in these analyses received ethics approval from Northwestern University's Institutional Review Board. Across studies, participants were recruited through university-wide mailing lists, flyers around campus, or the undergraduate research participant pool. Participants either received course credit or monetary compensation for their time.

**NIH Toolbox Cognition Battery (Zelazo et al. 2013).** The NIH Toolbox Cognition Battery is a standardized assessment tool used to measure a variety of different cognitive processes, including language, executive function, memory, and processing speed. The individual tests were administered and scored according to the guidelines in the testing manual. Raw scores were converted to normalized scaled scores by age (Casaletto et al. 2015). The battery took approximately 30 to 35 min to complete.

Reliability and validity were previously assessed by Weintraub et al. (2013). The authors ran interclass correlations between each test and well-established measures of the same construct (e.g., Picture Vocabulary Test from the NIH Toolbox Cognition Battery and the Peabody Picture Vocabulary Test – Fourth Edition by Dunn and Dunn, 2007). In adults, the interclass correlations were high, ranging from 0.72 for processing speed to 0.94 for inhibitory control. Furthermore, correlations for convergent validity were high, suggesting that each measure taps into the desired construct, while correlations for discriminant validity were low, indicating a weak relationship with measures tapping into different constructs (Weintraub et al. 2013).

**Executive Function.** Inhibitory control was measured using the Flanker task (Eriksen and Eriksen, 1974). In the adult version of the Flanker task, participants were shown a row of five arrows. The middle arrow pointed either in the same direction (←←←←←) or opposite direction (←←→←←) as the surrounding arrows. Participants selected the button that matched the direction the middle arrow was pointing to and completed four practice trials with feedback and 20 experimental trials in approximately 4 min.

Cognitive flexibility was measured using the Dimensional Change Card Sort test (DCCS; Zelazo, 2006). In the adult version of the DCCS test, participants sorted bivalent stimuli (e.g., blue ball or yellow truck) by either its shape or colour depending on the cue provided. Four practice trials and 30 experimental trials with color and shape cues intermixed were administered to participants. On average, the test took approximately 4 min to complete.

**Language.** Receptive vocabulary in English was measured using the Picture Vocabulary test and decoding skills in reading was measured using the Oral Reading Recognition test. In the Picture

**Table 1 Summary of L1 and L2 measures for monolinguals and bilinguals.**

	Measure	Group	L1	L2	L1 vs. L2
Proficiency (0-10)	Speaking	Monolingual	9.79 (0.49)	0.45 (1.5)	***
		Bilingual	9.36 (0.9)	6.02 (3.07)	***
	Reading	Monolingual	9.75 (0.65)	0.55 (1.88)	***
		Bilingual	9.17 (1.51)	6.33 (2.96)	***
	Understanding	Monolingual	9.79 (0.5)	0.55 (1.86)	***
		Bilingual	9.51 (0.78)	6.39 (3.03)	***
Age of Acquisition (in years)	Overall	Monolingual	0.36 (0.67)	19.76 (8.49)	***
		Bilingual	0.36 (0.83)	8.46 (4.82)	***
	Reading	Monolingual	3.86 (1.63)	18.96 (7.11)	***
		Bilingual	4.1 (1.89)	10.02 (4.24)	***
Manner of Acquisition (0-10)	Family	Monolingual	8.88 (1.65)	0.55 (2.13)	***
		Bilingual	9.27 (1.4)	2.75 (3.45)	***
	Friends	Monolingual	6.71 (2.6)	0.55 (2.08)	***
		Bilingual	6.9 (3.09)	6.77 (3.61)	
	Self-Instruction	Monolingual	1.48 (3.44)	0.03 (0.19)	
		Bilingual	1.79 (2.77)	3.4 (3.2)	***
	TV	Monolingual	5.46 (3.15)	0.28 (1.49)	***
		Bilingual	5.94 (2.89)	5.22 (3.12)	
	Radio/Music	Monolingual	3.67 (3.05)	0.17 (0.93)	***
		Bilingual	2.82 (2.74)	2.82 (3)	
	Reading	Monolingual	7.96 (1.97)	0.69 (2.58)	***
		Bilingual	7.3 (2.42)	7.21 (2.98)	
Current Exposure (0-10)	Family	Monolingual	9.62 (0.97)	0.59 (2.23)	***
		Bilingual	8.27 (2.54)	1.79 (2.36)	***
	Friends	Monolingual	9.75 (0.53)	0.21 (1.11)	***
		Bilingual	6.97 (3.26)	5.64 (3.87)	
	Self-Instruction	Monolingual	5.62 (4.73)	0 (0)	***
		Bilingual	2.05 (3.56)	1.88 (2.68)	
	TV	Monolingual	9.12 (2.17)	0.34 (1.86)	***
		Bilingual	6 (3.32)	4.42 (3.41)	**
	Radio/Music	Monolingual	9.04 (1.49)	0.24 (1.12)	***
		Bilingual	5.85 (3.18)	4.7 (3.36)	
	Reading	Monolingual	9.79 (0.51)	0.34 (1.86)	***
		Bilingual	5.5 (3.57)	6.25 (3.7)	
Immersion (in years)	Family	Monolingual	23.48 (4.31)	2.31 (6.93)	***
		Bilingual	20.49 (4.19)	3.76 (7.16)	***
	Country	Monolingual	23.72 (4.05)	0.83 (4.46)	***
		Bilingual	17.58 (7.13)	5.45 (7.22)	***
	School/Work	Monolingual	21.76 (5.13)	0.62 (3.34)	***
		Bilingual	16.08 (7.94)	5.65 (6.39)	***

Note. Monolinguals' L2 ages of acquisition were set to their current ages and other experience measures were set to 0 when no second language was known. Values in parentheses represent standard deviations. Asterisks represent significant differences between L1 and L2. \*\*\**p* < 0.001, \*\**p* < 0.01.

Vocabulary test, participants selected an image out of four options that best represented the word they heard. Two practice trials and approximately 25 items were administered to participants. On average, the test took 4 min to complete.

For the Oral Reading Recognition test, participants had 120 seconds to name out loud a series of letters or words that were presented on the screen one at a time. The number of items shown was based on the age of the participant, but on average, around 30 to 40 items were presented to participants.

*Memory.* Working memory was assessed using the List Sorting Working Memory test and episodic memory was assessed using the Picture Sequence Memory test. The List Sorting Working Memory test consisted of two conditions: 1-List condition and 2-List condition. In the 1-List condition, participants were shown pictures one at a time that belonged to the same category (e.g., fruits) followed by a blank screen. Once the blank screen appeared, participants named the pictures from smallest to largest. The 2-List condition followed a similar procedure as the 1-List condition, except that the pictures were from two different categories (e.g., animals and fruits). When the blank screen

appeared, participants named the pictures from smallest to largest from the first category (e.g., animals) followed by the second category (e.g., fruits). Participants were administered two practice trials with two pictures in each series. For the experimental trials, the number of pictures within each series increased by one on trials following a correct response (from two pictures to a maximum of seven pictures in a single sequence). The test ended when 2 trials of the same length failed or the maximum sequence of seven items was presented. On average, the test took approximately 7 min to complete.

In the Picture Sequence Memory test (episodic memory), participants were shown a sequence of pictures containing specific events or activities, such as going camping, having a birthday party, playing in the park, and so on. Once the sequence ended, the pictures were scrambled, and participants were instructed to move the pictures to the location they saw them move to on the screen. Participants were provided with one practice sequence, containing four pictures. If the participant succeeded in four trials, two experimental sequences were presented: one with 15 pictures and the other with 18 pictures. They completed the test in approximately 10 min.

**Processing Speed.** The Pattern Comparison Processing Speed test was used to measure processing speed. Two pictures appeared on the screen side-by-side. If the two pictures were identical, participants touched the “Yes” button. If the two pictures were different, participants touched the “No” button. Participants were administered six practice trials and named as many items as possible (of a possible 130) within 85 seconds.

### Statistical Analyses

**Cluster Analysis.** A cluster analysis was conducted to determine the optimal grouping among the 140 bilingual participants who reported proficiency greater than 1 (out of 10) in a second language. As information regarding manner of acquisition, exposure, and/or immersion were missing from 54 of the 140 bilingual participants, clusters were determined based on mean-centered measures of relative age of acquisition ( $L2 - L1$ ) and relative proficiency ( $L1 - L2$ ). Analyses were carried out using the Hartigan-Wong K-means clustering algorithm (Hartigan and Wong, 1979), executed with the *kmeans* R function with 20 random starting assignments and 4 *k* clusters (based on an examination of within-cluster sum of squares to determine the optimal number of groups).

To obtain more information about each group’s language usage patterns, we also examined whether the groups differed in their manner of language acquisition and exposure. Following an examination of correlations among LEAP-Q measures, four composite measures were created capturing 1) family acquisition and exposure (aggregated across manner of acquisition and current exposure ratings for family), 2) media and community language acquisition (aggregated across manner of acquisition ratings for friends, reading, TV, and music), 3) media and community language exposure (aggregated across current exposure ratings for friends, reading, TV, and music), and 4) language immersion (aggregated across the number of years immersed in a country, school, and work environment in which a given language was spoken) for each language.

**NIH Toolbox Analyses.** Separate linear models examined the effect of cluster group (monolinguals, late-unbalanced bilinguals, early-unbalanced bilinguals, late-balanced bilinguals, early-balanced bilinguals; see Results for additional detail) on each of the NIH Toolbox Cognition Battery measures. For each model, cluster group was sum coded with four contrasts (monolingual: 1, 0, 0, 0; late-unbalanced bilinguals: 0, 1, 0, 0; early-unbalanced bilinguals: 0, 0, 1, 0; late-balanced bilinguals: 0, 0, 0, 1; early-balanced bilinguals: -1, -1, -1, -1) to compare each group to the grand mean. Each group was compared to the grand mean because there was no group that could theoretically serve as the reference level. Following initial comparisons to the grand mean, we performed Tukey-adjusted planned pairwise comparisons to compare each group to each other using the *emmeans* R function (Lenth et al. 2018).

The bilingual clusters did not differ in gender ( $ps > 0.12$ ) or in age ( $ps > 0.090$ ). However, because monolinguals ( $M = 22.64$ ,  $SD = 8.46$ ) were older than late-unbalanced bilinguals ( $M = 21.00$ ,  $SD = 2.45$ ,  $p = 0.036$ ) and marginally older than early-balanced bilinguals ( $M = 21.22$ ,  $SD = 3.15$ ,  $p = 0.069$ ), all models included age as a covariate. Furthermore, as late-unbalanced bilinguals ( $M = 14.59$ ,  $SD = 2.65$ ) had fewer years of education relative to the grand mean ( $M = 15.19$ ,  $SD = 2.11$ ;  $Estimate = -0.65$ ,  $SE = 0.32$ ,  $t(163) = -2.02$ ,  $p = 0.045$ ), all models included mean-centered years of education as a covariate. Years of education did not significantly differ among monolinguals ( $M = 15.79$ ,  $SD = 2.08$ ), early-unbalanced bilinguals ( $M = 15.22$ ,  $SD = 2.00$ ), late-balanced bilinguals ( $M = 15.60$ ,

$SD = 1.83$ ), or early-balanced bilinguals ( $M = 15.02$ ,  $SD = 1.85$ ),  $ps > 0.12$ . All unbalanced bilinguals had English as their L1, whereas the balanced bilinguals had a mixture of English as their L1 or the other language as their L1. To account for the potential differences between balanced and unbalanced bilinguals in English proficiency, we included average English proficiency (aggregated across English speaking, understanding, and reading proficiency) as a covariate in all models. Moreover, as 61 of the participants reported knowing three or more languages, we also included number of known languages as a covariate in all models.

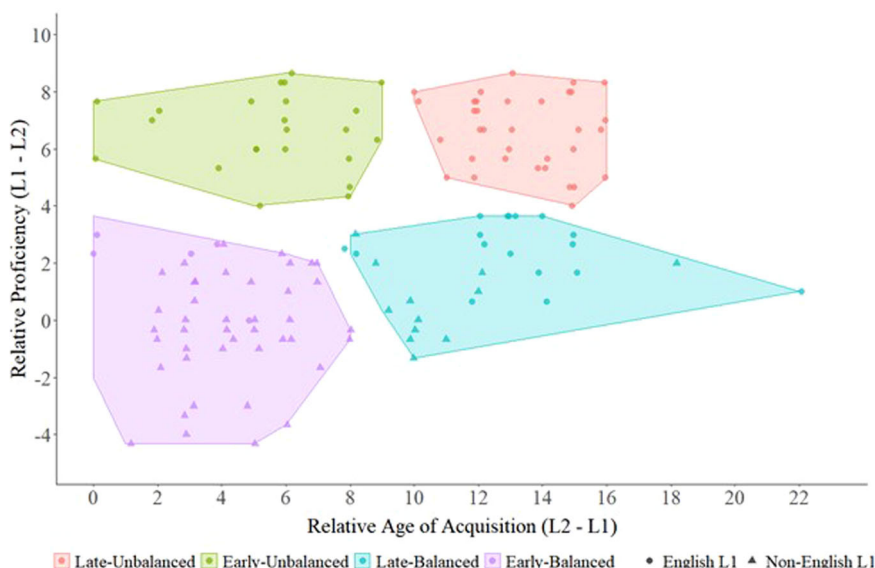
### Results

The cluster analysis based on relative language proficiency ( $L1 - L2$ ) and relative age of acquisition ( $L2 - L1$ ) yielded four bilingual groups, which broadly corresponded to late-unbalanced bilinguals ( $N = 34$ ), early-unbalanced bilinguals ( $N = 23$ ), late-balanced bilinguals ( $N = 29$ ), and early-balanced bilinguals ( $N = 54$ ; see Fig. 1 as well as Tables 2 and 3 for overviews of each group’s linguistic profile). The total number of participants contributing to each model varied depending on the availability of relevant NIH Toolbox Cognition Battery measures.

The difference between L1 and L2 age of acquisition (AoA) was significantly smaller for early (aggregated across balanced and unbalanced bilinguals;  $M = 4.21$ ,  $SD = 2.53$ ) compared to late bilinguals ( $M = 12.84$ ,  $SD = 2.5$ ;  $Estimate = -8.63$ ,  $SE = 0.7$ ,  $t(166) = -12.29$ ,  $p < 0.001$ ). The difference between L1 and L2 proficiency was significantly smaller for balanced bilinguals (aggregated across early and late bilinguals;  $M = 0.63$ ,  $SD = 2.02$ ) compared to unbalanced bilinguals ( $M = 6.63$ ,  $SD = 1.28$ ;  $Estimate = 6.00$ ,  $SE = 0.28$ ,  $t(166) = 21.43$ ,  $p < 0.001$ ). Although L1 and L2 AoA were closer for early than late bilinguals, L1 AoA was significantly earlier than L2 AoA for all four groups ( $ps < 0.001$ ). L1 proficiency was significantly higher than L2 proficiency for early-unbalanced, late-unbalanced, and late-balanced bilinguals ( $ps < 0.001$ ), whereas L1 and L2 proficiency did not differ for early-balanced bilinguals ( $p = 0.89$ ). English was the L1 for all late-unbalanced and early-unbalanced bilinguals. English was the L1 for 58.6% of late-balanced bilinguals and 16.7% of early-balanced bilinguals. English AoA was significantly earlier than non-English AoA for early- and late-unbalanced bilinguals ( $ps < 0.001$ ), did not differ for late-balanced bilinguals ( $p = 0.16$ ), and was significantly later than non-English AoA for early-balanced bilinguals ( $p < 0.001$ ). English proficiency was significantly greater than non-English proficiency for all four groups ( $ps < 0.007$ ).

### Executive Function

**Flanker Task: Inhibitory Control.** A total of 156 participants contributed to the model for flanker performance, including 22 monolinguals, 34 late-unbalanced bilinguals, 23 early-unbalanced bilinguals, 25 late-balanced bilinguals, and 52 early-balanced bilinguals. An additional 13 participants were not included in the analysis due to missing scores ( $n = 6$ ) or due to scores more than 2 standard deviations below the grand mean ( $n = 7$ ). Participants had a mean Flanker score of 115.37 ( $SD = 5.31$ ) and late-unbalanced bilinguals had significantly lower scores compared to the grand mean ( $Estimate = 2.50$ ,  $SE = 0.81$ ,  $t(147) = 3.11$ ,  $p = 0.0023$ ). Tukey-adjusted planned pairwise comparisons revealed that late-unbalanced bilinguals had significantly lower scores than early-unbalanced bilinguals ( $Estimate = 2.75$ ,  $SE = 1.37$ ,  $t(147) = 2.02$ ,  $p = 0.046$ , *Cohen’s d* = 0.55, 95% CI [-1.09, -0.01]), late-balanced ( $Estimate = 4.20$ ,  $SE = 1.35$ ,  $t(147) = 3.11$ ,  $p = 0.002$ , *Cohen’s d* = 0.84, 95% CI [-1.38, -0.30]), and early-balanced bilinguals ( $Estimate = 3.45$ ,  $SE = 1.17$ ,  $t(147) = 2.95$ ,  $p = 0.004$ , *Cohen’s d* = 0.69, 95% CI



**Fig. 1** Cluster analyses based on relative age of acquisition (L2 – L1) and proficiency (L1 – L2) yielded four bilingual groups comprised of late-unbalanced bilinguals (in red; N = 34), early-unbalanced bilinguals (in green; N = 23), late-balanced bilinguals (in teal; N = 29), and early-balanced bilinguals (in purple; N = 54). Circles represent participants whose L1 was English and triangles represent participants whose L1 was a non-English language.

**Table 2 Composite L1 and L2 language experience measures by cluster group.**

Measure	Late-Unbalanced	Early-Unbalanced	Late-Balanced	Early-Balanced
L1 AoA	0.06 (0.24)	0.39 (1.12)	0.14 (0.44)	0.67 (0.99)
L2 AoA	13.44 (1.76)	5.96 (3.04)	12.34 (3.17)	4.31 (2.33)
L1 Proficiency	9.73 (0.51)	9.90 (0.23)	9.47 (0.65)	8.81 (1.24)
L2 Proficiency	3.11 (1.20)	3.26 (1.33)	7.74 (1.31)	8.78 (1.27)
L1 Family Acquisition/Exposure	9.68 (0.93)	8.81 (1.44)	8.98 (1.35)	8.44 (1.89)
L2 Family Acquisition/Exposure	0.75 (1.27)	3.56 (3.53)	1.21 (1.95)	2.84 (2.85)
L1 Media/Community Acquisition	6.73 (2.05)	5.41 (2.14)	5.81 (1.76)	5.54 (1.78)
L2 Media/Community Acquisition	2.29 (1.43)	1.81 (1.00)	6.14 (1.52)	6.67 (1.65)
L1 Media/Community Exposure	8.98 (1.08)	8.78 (1.27)	5.47 (2.78)	5.10 (2.73)
L2 Media/Community Exposure	1.27 (1.00)	1.00 (0.89)	5.29 (2.61)	6.90 (2.27)
L1 Immersion	20.83 (2.45)	21.55 (3.73)	18.57 (7.13)	10.85 (7.56)
L2 Immersion	1.05 (2.44)	1.40 (2.68)	3.55 (3.98)	9.88 (7.06)

Note. Values in parentheses represent standard deviations.

[–1.15, –0.22]; see Fig. 2a). Late-balanced, early-balanced, and early-unbalanced bilinguals were not significantly different from each other,  $ps > 0.33$ . Furthermore, monolinguals did not differ from any of the bilingual groups, all  $ps > 0.15$ . Flanker scores declined with age ( $Estimate = 0.38$ ,  $SE = 0.17$ ,  $t(147) = 2.21$ ,  $p = 0.028$ ), but not with years of education, number of known languages, or average English proficiency, all  $ps > 0.25$ .

**Dimensional Change Card Sort (DCCS) Test: Cognitive Flexibility.** A total of 151 participants contributed to the model for DCCS performance, including 20 monolinguals, 33 late-unbalanced bilinguals, 23 early-unbalanced bilinguals, 24 late-balanced bilinguals, and 51 early-balanced bilinguals. An additional 18 participants were not included in the analysis due to missing scores ( $n = 9$ ) or due to scores greater than 2 standard deviations below the grand mean ( $n = 9$ ). Participants had a mean DCCS score of 112.09 ( $SD = 6.92$ ). None of the groups’ scores were significantly different from the grand mean,  $ps > 0.085$ . Tukey-adjusted planned pairwise comparisons revealed early-unbalanced bilinguals ( $Estimate = 3.75$ ,  $SE = 1.86$ ,  $t(142) = 2.01$ ,  $p = 0.046$ ,  $Cohen’s d = 0.55$ , 95% CI [–1.10, –0.01]) and early-balanced bilinguals ( $Estimate = 3.76$ ,  $SE = 1.60$ ,  $t(142) = 2.34$ ,

$p = 0.020$ ,  $Cohen’s d = 0.57$ , 95% CI [–1.04, –0.10]) scored significantly higher than late-unbalanced bilinguals (Fig. 2b). The other bilingual groups did not differ significantly from each other,  $ps > 0.27$ . Additionally, monolinguals scored marginally lower than early-balanced bilinguals ( $Estimate = 3.88$ ,  $SE = 2.08$ ,  $t(142) = 1.86$ ,  $p = 0.064$ ,  $Cohen’s d = 0.57$ , 95% CI [–1.18, 0.038]), but did not significantly differ from any of the other bilingual groups,  $ps > 0.091$ . DCCS scores did not vary by age, years of education, or number of known languages ( $ps > 0.29$ ), but did vary marginally by average English proficiency ( $Estimate = 1.40$ ,  $SE = 0.75$ ,  $t(142) = 1.85$ ,  $p = 0.066$ ).

**Language**

**Picture Vocabulary.** A total of 162 participants contributed to the model for English picture vocabulary performance, including 24 monolinguals, 34 late-unbalanced bilinguals, 22 early-unbalanced bilinguals, 29 late-balanced bilinguals, and 53 early-balanced bilinguals. An additional 7 participants were not included in the analysis due to missing scores ( $n = 4$ ) or due to scores greater than 2 standard deviations below the grand mean ( $n = 3$ ). Participants had a mean Picture Vocabulary score of 116.04

**Table 3 Composite L1 and L2 language experience measures by early versus late bilinguals and balanced versus unbalanced bilinguals.**

Measure	Early Bilinguals	Late Bilinguals	Early vs. Late	Balanced Bilinguals	Unbalanced Bilinguals	Balanced vs. Unbalanced
L1 Age of Acquisition	0.58 (1.03)	0.10 (0.35)	***	0.48 (0.88)	0.19 (0.74)	*
L2 Age of Acquisition	4.80 (2.65)	12.94 (2.55)	***	7.11 (4.67)	10.42 (4.38)	***
L1 Proficiency	9.14 (1.16)	9.61 (0.59)	**	9.04 (1.11)	9.80 (0.43)	***
L2 Proficiency	7.13 (2.85)	5.24 (2.64)	**	8.42 (1.37)	3.17 (1.24)	***
L1 Family Acquisition/Exposure	8.49 (1.83)	9.26 (1.23)	*	8.60 (1.76)	9.36 (1.19)	
L2 Family Acquisition/Exposure	2.95 (2.93)	1.03 (1.71)	**	2.38 (2.72)	1.88 (2.74)	
L1 Media/Community Acquisition	5.51 (1.81)	6.19 (1.91)		5.62 (1.77)	6.25 (2.14)	
L2 Media/Community Acquisition	5.97 (2.32)	4.65 (2.40)	*	6.52 (1.62)	2.10 (1.27)	***
L1 Media/Community Exposure	5.63 (2.88)	6.90 (2.83)	*	5.21 (2.73)	8.90 (1.13)	***
L2 Media/Community Exposure	6.06 (2.97)	3.74 (2.90)	**	6.45 (2.46)	1.16 (0.94)	***
L1 Immersion	13.92 (8.25)	19.80 (5.23)	***	13.55 (8.25)	21.11 (3.00)	***
L2 Immersion	7.43 (7.22)	2.20 (3.44)	***	7.67 (6.84)	1.19 (2.51)	***

Note: Asterisks in the Early vs. Late column indicate significant differences between early bilinguals (early-unbalanced + early-balanced) and late bilinguals (late-unbalanced + late-balanced), while asterisks in the Balanced vs. Unbalanced column indicate significant differences between balanced bilinguals (early-balanced) and unbalanced bilinguals (early-unbalanced + late-unbalanced). \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

( $SD = 12.80$ ). None of the groups' scores were significantly different from the grand mean,  $ps > 0.17$ . Tukey-adjusted planned pairwise comparisons revealed that scores did not significantly differ among monolinguals ( $M = 116.69$ ,  $SD = 14.05$ ), late-unbalanced bilinguals ( $M = 118.96$ ,  $SD = 10.74$ ), early-unbalanced bilinguals ( $M = 116.99$ ,  $SD = 10.50$ ), late-balanced bilinguals ( $M = 116.78$ ,  $SD = 11.57$ ), and early-balanced bilinguals ( $M = 113.07$ ,  $SD = 14.65$ ),  $ps > 0.087$ . Picture Vocabulary scores did not vary with age, years of education, number of known languages, or average English proficiency,  $ps > 0.14$ .

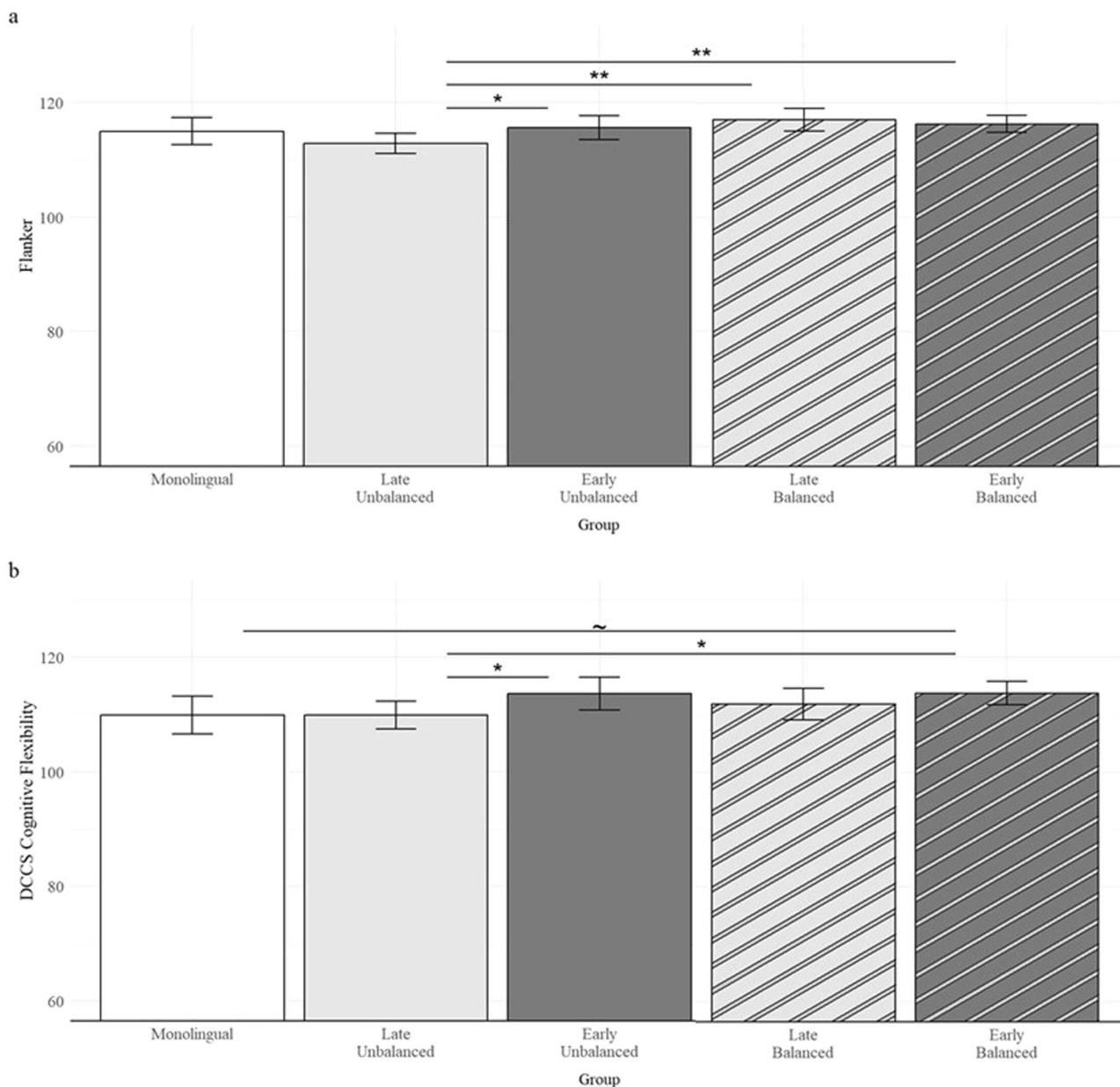
**Oral Reading Recognition.** A total of 95 participants contributed to the model for English oral reading performance, including 11 monolinguals, 28 late-unbalanced bilinguals, 16 early-unbalanced bilinguals, 13 late-balanced bilinguals, and 27 early-balanced bilinguals. An additional 74 participants were not included in the analysis due to missing scores ( $n = 68$ ) or due to scores greater than 2 standard deviations below the grand mean ( $n = 6$ ). Participants had a mean Oral Reading score of 120.82 ( $SD = 8.15$ ) and late-unbalanced bilinguals ( $M = 123.54$ ,  $SD = 6.97$ ) had marginally higher scores than the grand mean ( $Estimate = -2.97$ ,  $SE = 1.5$ ,  $t(86) = 1.98$ ,  $p = 0.051$ ). Tukey-adjusted planned pairwise comparisons revealed that oral reading performance was marginally higher for late-unbalanced bilinguals relative to late-balanced bilinguals ( $M = 118.22$ ,  $SD = 10.60$ ;  $Estimate = -4.90$ ,  $SE = 2.76$ ,  $t(86) = -1.78$ ,  $p = 0.079$ ,  $Cohen's d = 0.60$ , 95% CI  $[-0.076, 1.28]$ ). Early-unbalanced ( $M = 120.41$ ,  $SD = 7.80$ ) and early-balanced ( $M = 120.07$ ,  $SD = 8.31$ ) bilinguals did not differ significantly from any of the bilingual groups,  $ps > 0.22$ . The monolinguals ( $M = 119.39$ ,  $SD = 7.34$ ) did not differ from any of the bilingual groups,  $ps > 0.14$ . Oral Reading Recognition scores did not vary with age, years of education, number of known languages, or average English proficiency,  $ps > 0.25$ .

**Memory**

**List Sorting Working Memory.** A total of 159 participants contributed to the model for list sorting working memory performance, including 23 monolinguals, 34 late-unbalanced bilinguals, 23 early-unbalanced bilinguals, 28 late-balanced bilinguals, and 51 early-balanced bilinguals. An additional 10 participants were not included in the analysis due to missing scores ( $n = 8$ ) or due to scores greater than 2 standard deviations below the grand mean ( $n = 2$ ). Participants had a mean List Sorting Working Memory score of 107.55 ( $SD = 11.45$ ) and early-balanced bilinguals had significantly lower scores than the grand mean ( $Estimate = -4.11$ ,  $SE = 1.65$ ,  $t(150) = -2.50$ ,  $p = 0.014$ ). Tukey-adjusted planned pairwise comparisons revealed that early-balanced bilinguals scored significantly lower than late-unbalanced bilinguals ( $Estimate = -6.26$ ,  $SE = 2.56$ ,  $t(150) = -2.44$ ,  $p = 0.016$ ,  $Cohen's d = 0.56$ , 95% CI  $[0.10, 1.03]$ ) and marginally lower than monolinguals ( $Estimate = -6.02$ ,  $SE = 3.22$ ,  $t(150) = -1.87$ ,  $p = 0.064$ ,  $Cohen's d = 0.54$ , 95% CI  $[-0.034, 1.12]$ , see Fig. 3). All other bilingual groups did not differ from each other,  $ps > 0.13$ . Additionally, monolinguals did not differ from late-balanced, late-unbalanced, and early-unbalanced bilinguals,  $ps > 0.57$ . List Sorting Working Memory scores did not decline with age, years of education, number of known languages, or average English proficiency,  $ps > 0.14$ .

**Picture Sequence Episodic Memory.** A total of 100 participants contributed to the categorical model for picture sequence episodic memory performance, including 10 monolinguals, 28 late-unbalanced bilinguals, 16 early-unbalanced bilinguals, 15 late-balanced bilinguals, and 31 early-balanced bilinguals. An additional 69 participants were not included in the analysis due to



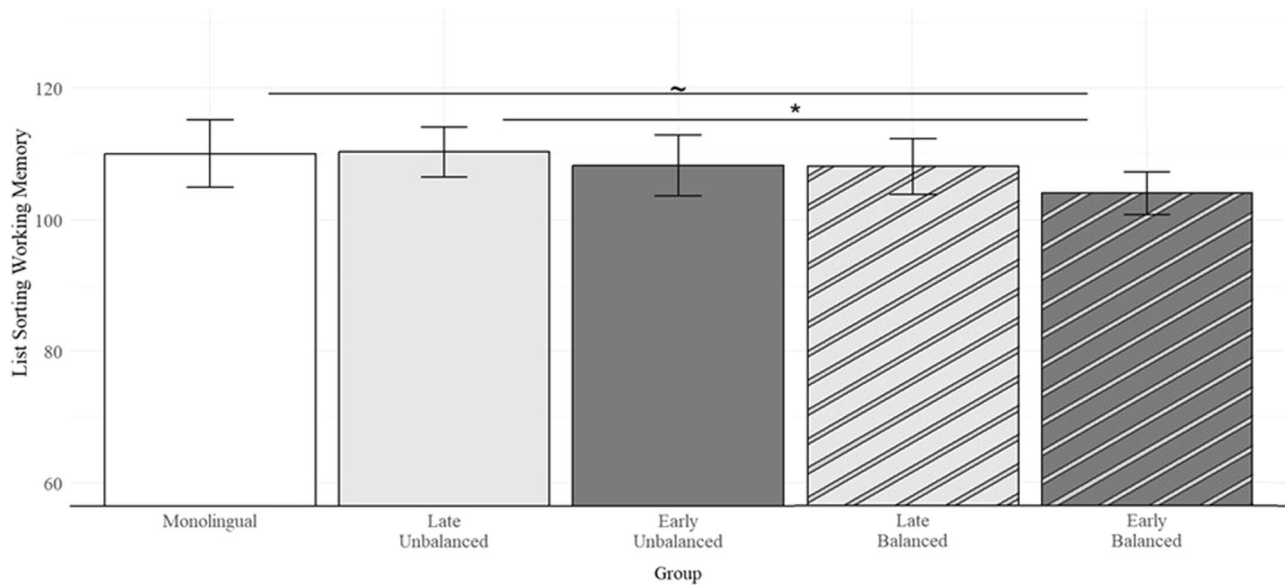


**Fig. 2 Mean scores on the executive function subtests of the NIH Toolbox Cognition Battery for each language group.** Flanker scores (a) were higher among early-balanced, late-balanced, and early-unbalanced bilinguals relative to late-unbalanced bilinguals. DCCS scores (b) were higher among early-unbalanced and early-balanced bilinguals relative to late-unbalanced bilinguals, and marginally higher among early-balanced bilinguals relative to monolinguals. Error bars represent 95% confidence intervals. \*\* $p < 0.01$ , \* $p < 0.05$ , ~ $p < 0.07$ .

missing scores ( $n = 65$ ) or due to scores greater than 2 standard deviations below the grand mean ( $n = 4$ ). Participants had a mean picture sequence episodic memory score of 112.82 ( $SD = 13.68$ ). Tukey-adjusted planned pairwise comparisons revealed that scores did not significantly differ among monolinguals ( $M = 113.50$ ,  $SD = 14.46$ ), late-unbalanced bilinguals ( $M = 111.69$ ,  $SD = 15.06$ ), early-unbalanced bilinguals ( $M = 114.39$ ,  $SD = 13.57$ ), late-balanced bilinguals ( $M = 115.48$ ,  $SD = 12.66$ ), and early-balanced bilinguals ( $M = 111.52$ ,  $SD = 13.26$ ),  $ps > 0.32$ . Picture Sequence Episodic Memory scores did not decline with age, years of education, number of known languages, or average English proficiency,  $ps = 0.20$ .

**Pattern Comparison Processing Speed.** A total of 153 participants contributed to the model for pattern comparison processing speed performance, including 22 monolinguals, 33 late-

unbalanced bilinguals, 22 early-unbalanced bilinguals, 27 late-balanced bilinguals, and 49 early-balanced bilinguals. An additional 16 participants were not included in the analysis due to missing ( $n = 8$ ) scores or due to scores greater than 2 standard deviations below the grand mean ( $n = 8$ ). Participants had a mean Pattern Comparison score of 126.22 ( $SD = 16.55$ ). Tukey-adjusted planned pairwise comparisons revealed that scores did not significantly differ among monolinguals ( $M = 125.86$ ,  $SD = 16.25$ ), late-unbalanced bilinguals ( $M = 123.39$ ,  $SD = 19.05$ ), early-unbalanced bilinguals ( $M = 122.41$ ,  $SD = 13.50$ ), late-balanced bilinguals ( $M = 128.02$ ,  $SD = 17.31$ ), and early-balanced bilinguals ( $M = 129.02$ ,  $SD = 15.67$ ),  $ps > 0.11$ . Patten Comparison Processing Speed scores declined with age ( $Estimate = -1.37$ ,  $SE = 0.56$ ,  $t(144) = -2.45$ ,  $p = 0.015$ ), but not years of education, number of known languages, or average English proficiency,  $ps > 0.70$ .



**Fig. 3** Mean List Sorting Working Memory scores by group. List sorting scores were higher among late-unbalanced bilinguals relative to early-balanced bilinguals, and marginally higher among monolinguals relative to early-balanced bilinguals. Error bars represent 95% confidence intervals. \*\* $p < 0.01$ , \* $p < 0.05$ , ~ $p < 0.07$ .

## Discussion

The goal of the present study was to investigate the impact of language experience on cognitive function by examining how individuals with distinct language profiles perform on the NIH Toolbox Cognition Battery. In particular, the utility of employing a data-driven approach for identifying subgroups of participants was examined, which may provide a more nuanced and ecologically valid model of language experience than traditional monolingual-bilingual dichotomies. A cluster analysis based on self-reported ages of language acquisition and proficiency indicated that participants with second-language experience could be broadly grouped into “early” versus “late” bilinguals, as well as “balanced” versus “unbalanced” bilinguals. Early bilinguals acquired a second language at the same time as or soon after the first language (mean difference in ages of acquisition = 4.21 years), while late bilinguals acquired a second language substantially later than the first language (mean difference = 12.84 years). Balanced bilinguals had similar levels of first and second language proficiency (mean difference = 0.63 on a scale from 0 to 10) and unbalanced bilinguals were relatively more dominant in their first than second language (mean difference = 6.63).

Overall, our results align with previous work indicating that greater bilingual language experience is associated with better performance on *executive control* tasks (e.g., Flanker task and DCCS test; see Bialystok, 2017 for a review), but weaker performance on tasks that rely on linguistic and verbal abilities within a single language (e.g., List Sorting Working Memory test; Bialystok et al. 2022; Gollan et al. 2005; Ivanova and Costa, 2008; Marian et al. 2018a). We found that balanced proficiency and early age of acquisition moderated performance on *executive function* tasks (i.e., DCCS test, Flanker task, and List Sorting Working Memory test), but not on *language-based* tasks (i.e., Picture Vocabulary test and Oral Reading Recognition test). Controlling for English proficiency, performance on picture vocabulary task (English vocabulary), picture sequence memory task (episodic memory), and pattern comparison task (processing speed) was comparable among the five cluster groups. Our findings indicate that the effects of language experience on cognitive function vary depending on both the *domain* of cognitive function (e.g., verbal

vs. nonverbal) and the *dimension* of language experience (e.g., age of acquisition vs. proficiency).

The results from the flanker test indicate that inhibitory control may be jointly impacted by dual-language proficiency and age of acquisition. Flanker scores were significantly higher among early-balanced, late-balanced, and early-unbalanced bilinguals compared to late-unbalanced bilinguals. It has been hypothesized that effects of bilingualism on inhibitory control may stem from the parallel activation of both languages and the use of domain-general executive functions to minimize interference from the non-target language (Bialystok et al. 2012). Our finding that Flanker scores were highest among those with more balanced bilingual proficiency is consistent with prior work showing that high-proficiency languages are more likely to be activated when not in use (e.g., Blumenfeld and Marian, 2007; Canseco-Gonzalez et al. 2010; Meuter and Allport, 1999; Weber and Cutler, 2004), which necessitates a greater degree of inhibition to suppress competition. In addition, according to Cummins’ Threshold Hypothesis (Cummins, 1976), the benefits associated with bilingualism may only appear once the speaker has attained a certain level of competency in their second language. This could also serve as an explanation for why late-unbalanced bilinguals had lower scores on the executive function measures than the other bilingual groups. Balanced bilingualism may be especially important for measures of selective attention because early- and late-balanced bilinguals need to manage attention between two jointly activated languages. These findings highlight the need to examine how different forms of bilingual experience influence specific cognitive functions.

Effects of language experience on the Dimensional Change Card Sort (DCCS) test performance suggest that cognitive flexibility may be jointly impacted by age of dual-language acquisition and relative proficiency. DCCS scores were higher for early-unbalanced compared to late-unbalanced bilinguals. Additionally, DCCS scores were higher for early-balanced bilinguals than monolinguals and late-unbalanced bilinguals. The effects of bilingualism on cognitive flexibility have been attributed to bilinguals’ experience alternating between multiple linguistic systems, which may facilitate task-switching performance by enhancing global monitoring skills and maintenance of

competing task goals (Barac and Bialystok, 2012; Soveri et al. 2011; Wiseheart et al. 2016). Early bilinguals have, by definition, a longer duration of experience managing multiple languages than late bilinguals. Hence, the advantages observed for cognitive flexibility may stem in part from a higher probability of switching between languages within a single context (e.g., with family). Indeed, although all subgroups in the current study reported greater L1 than L2 experience with family, the asymmetry in family language use was smaller among early bilinguals compared to late bilinguals (refer to Tables 2 and 3). Therefore, bilinguals with more experience using multiple languages for most of their lives showed greater cognitive flexibility than those with less experience using multiple languages.

On the executive control task that relied on verbal abilities, early-balanced bilinguals performed significantly worse than late-unbalanced bilinguals and marginally worse than monolinguals on the List Sorting Working Memory Test. These findings are likely driven by differences in English proficiency between groups. In our sample, the unbalanced bilinguals performed the NIH Toolbox Cognition Battery in their more dominant and proficient language (i.e., English). Late-unbalanced bilinguals ( $M = 9.73$ ,  $SD = 0.51$ ) rated their overall proficiency in English higher than early-balanced bilinguals ( $M = 9.16$ ,  $SD = 0.99$ ,  $p < 0.001$ ). Prior work with children has shown that unbalanced bilinguals generally outperform balanced bilinguals on language-based tasks in their dominant language (e.g., Persici et al. 2019). The observed effects for English verbal working memory performance may stem from relatively less extensive English experience rather than effects on memory per se. Preliminary support for this interpretation comes from the finding that non-verbal picture sequence memory did not differ among the five cluster groups, suggesting that the lower list sorting scores among early-balanced bilinguals may be primarily attributed to demands associated with English language processing. Furthermore, this explanation is supported by a significant correlation between list working memory scores and average self-reported proficiency in English,  $R = 0.22$ ,  $t(163) = 2.95$ ,  $p = 0.0037$ , 95%CI [0.07 0.37]. Considering that bilinguals have difficulties in lexical access and retrieval compared to their monolingual counterparts (e.g., Gollan et al. 2005), early-balanced bilinguals may be performing less well than monolinguals and late-unbalanced bilinguals on the working memory task because it is verbal in nature (Calvo et al. 2016).

Future research should consider performing a cluster analysis with a) more than two dimensions (e.g., x-axis as age of acquisition, y-axis as proficiency, and z-axis as language usage or exposure), b) different measures of language experience and seeing which measures provide better clustering, or c) objective measures of language proficiency because self-assessments of proficiency can vary significantly between different bilingual populations (Tomoschuk et al. 2019). The inclusion of language usage and objective measures of language proficiency will allow for more fine-grained clustering and contribute to our understanding of how variability in bilingual language experience influences cognition.

Considering there is no standard definition for “bilingualism” and language background can be characterized by multiple dimensions (e.g., proficiency, age of acquisition, and so on; Byers-Heinlein et al. 2019), the current study used machine learning to examine the complexities and nuances associated with language experience. Specifically, we gained a better understanding of how variability in bilingual language experience shapes certain aspects of cognition and not others. Our results suggest that more balanced bilinguals may have enhanced inhibitory control abilities by virtue of having to select between two jointly activated languages. Furthermore, early bilinguals may have enhanced

cognitive flexibility than late bilinguals due to their extensive experience in managing multiple languages at an early age.

In conclusion, the present study provides a model for how effects of language experience on cognitive function may be examined using a bottom-up and data-driven approach to identifying subgroups of participants. The interaction between first and second-language proficiency and age of acquisition has differential effects on nonverbal and verbal executive control, but no effect on processing speed and episodic memory among young adults. In addition to enabling more nuanced characterizations of linguistic profiles that extend beyond arbitrary dichotomies and isolated measures of cognitive function, comparisons of naturally occurring subgroups can provide a more holistic and ecologically valid perspective on how bilinguals differ among each other in and outside the lab. A multidimensional approach that considers variation in language experience and cognitive abilities as meaningful patterns in human behaviour can advance our understanding of how bilingualism shapes the mind and brain.

### Data availability

The dataset analyzed for this study is available from the corresponding author upon reasonable request.

Received: 9 March 2024; Accepted: 7 January 2025;

Published online: 18 January 2025

### References

- Adesope OO, Lavin T, Thompson T et al. (2010) A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Rev Educ Res* 80(2):207–245. <https://doi.org/10.3102/0034654310368803>
- Alladi S, Bak TH, Duggirala V et al. (2013) Bilingualism delays age at onset of dementia, independent of education and immigration status. *Neurology* 81(22):1938–1944. <https://doi.org/10.1212/01.wnl.0000436620.33155.a4>
- Anderson JAE, Chung-Fat-Yim A, Bellana B et al. (2018) Language and cognitive control networks in bilinguals and monolinguals. *Neuropsychologia* 117:352–363. <https://doi.org/10.1016/j.neuropsychologia.2018.06.023>
- Anitha P, Patil MM (2022) RFM model for customer purchase behavior using K-Means algorithm. *Journal of King Saud University-Computer and Information Sciences* 34(5):1785–1792
- Anjomshoae F, Wiebe SA, Nicoladis E (2024) Executive function's structure in monolingual and bilingual adults using confirmatory factor analysis. *Biling (Cambridge)*:1–11. <https://doi.org/10.1017/S136672892400021X>
- Antón E, Carreiras M, Duñabeitia JA (2019) The impact of bilingualism on executive functions and working memory in young adults. *PLoS One* 14(2):e0206770. <https://doi.org/10.1371/journal.pone.0206770>
- De Baene W, Duyck W, Brass M et al. (2015) Brain circuit for cognitive control is shared by task and language switching. *J Cogn Neurosci* 27(9):1752–1765. [https://doi.org/10.1162/jocn\\_a\\_00817](https://doi.org/10.1162/jocn_a_00817)
- Barac R, Bialystok E (2012) Bilingual effects on cognitive and linguistic development: Role of language, cultural background, and education. *Child Dev* 83(2):413–422. <https://doi.org/10.1111/j.1467-8624.2011.01707.x>
- von Bastian CC, Souza AS, Gade M (2016) No evidence for bilingual cognitive advantages: A test of four hypotheses. *J Exp Psychol Gen* 145(2):246–258. <https://doi.org/10.1037/xge0000120>
- Beatty-Martínez AL, Dussias PE (2017) Bilingual experience shapes language processing: Evidence from codeswitching. *J Mem Lang* 95:173–189
- Bialystok E (2017) The bilingual adaptation: How minds accommodate experience. *Psychol Bull* 143(3):233–262. <https://doi.org/10.1037/bul0000099>
- Bialystok E, Craik FIM (2022) How does bilingualism modify cognitive function? Attention to the mechanism. *Psychon Bull Rev* 29(4):1246–1269. <https://doi.org/10.3758/s13423-022-02057-5>
- Bialystok E, Craik FIM, Freedman M (2007) Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia* 45(2):459–464. <https://doi.org/10.1016/j.neuropsychologia.2006.10.009>
- Bialystok E, Craik F, Luk G (2008) Cognitive control and lexical access in younger and older bilinguals. *J Exp Psychol Learn Mem Cogn* 34(4):859–873. <https://doi.org/10.1037/0278-7393.34.4.859>
- Bialystok E, Craik FIM, Luk G (2012) Bilingualism: Consequences for mind and brain. *Trends Cogn Sci* 16(4):240–250. <https://doi.org/10.1016/j.tics.2012.03.001>

- Bialystok E, Hawrylewicz K, Grundy JG et al. (2022) The swerve: How childhood bilingualism changed from liability to benefit. *Dev Psychol* 58(8):1429–1440. <https://doi.org/10.1037/dev0001376>
- Blanco C (2020) 2020 Duolingo Language Report: Global Overview. Available via Duolingo. <https://blog.duolingo.com/global-language-report-2020/>
- Blumenfeld HK, Marian V (2007) Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eye-tracking. *Lang Cogn Process* 22(5):633–660. <https://doi.org/10.1080/01690960601000746>
- Blumenfeld HK, Schroeder SR, Bobb SC et al. (2016) Auditory word recognition across the lifespan: Links between linguistic and nonlinguistic inhibitory control in bilinguals and monolinguals. *Linguist Approaches Biling* 6(1-2):119–146. <https://doi.org/10.1075/lab.14030.bl>
- Bonifacci P, Giombini L, Bellocchi S et al. (2011) Speed of processing, anticipation, inhibition and working memory in bilinguals. *Dev Sci* 14(2):256–269. <https://doi.org/10.1111/j.1467-7687.2010.00974.x>
- Byers-Heinlein K, Esposito A, Winsler A et al. (2019) The case for measuring and reporting bilingualism in developmental research. *Collabra Psychol* 5(1):37. <https://doi.org/10.1525/collabra.233>
- Calabria M, Hernández M, Martin CD et al. (2011) When the tail counts: The advantage of bilingualism through the ex-gaussian distribution analysis. *Front Psychol* 2:250. <https://doi.org/10.3389/fpsyg.2011.00250>
- Calvo N, Ibáñez A, García AM (2016) The impact of bilingualism on working memory: A null effect on the whole may not be so on the parts. *Front Psychol* 7:265. <https://doi.org/10.3389/fpsyg.2016.00265>
- Canseco-Gonzalez E, Brehm L, Brick CA et al. (2010) Carpet or Cárcel: The effect of age of acquisition and language mode on bilingual lexical access. *Lang Cogn Process* 25(5):669–705. <https://doi.org/10.1080/01690960903474912>
- Casaletto KB, Umlauf A, Beaumont J et al. (2015) Demographically corrected normative standards for the English version of the NIH Toolbox cognition battery. *J Int Neuropsychol Soc* 21:378–391. <https://doi.org/10.1017/s1355617715000351>
- Chabal S, Hayakawa S, Marian V (2022) Language is activated by visual input regardless of memory demands or capacity. *Cogn* 222:104994. <https://doi.org/10.1016/j.cognition.2021.104994>
- Chen M, Ma F, Zhang Z et al. (2021) Language switching training modulates the neural network of non-linguistic cognitive control. *PLoS One* 16(4):e0247100. <https://doi.org/10.1371/journal.pone.0247100>
- Chen P, Bobb SC, Hoshino N et al. (2017) Neural signatures of language co-activation and control in bilingual spoken word comprehension. *Brain Res* 1665:50–64. <https://doi.org/10.1016/j.brainres.2017.03.023>
- Cockroft K, Wigdorowitz M, Liversage L (2019) A multilingual advantage in the components of working memory. *Biling (Camb Engl)* 22(1):15–29. <https://doi.org/10.1017/S1366728917000475>
- Comishen KJ, Bialystok E, Adler SA (2019) The impact of bilingual environments on selective attention in infancy. *Dev Sci* 22(4):e12797. <https://doi.org/10.1111/desc.12797>
- Cummins J (1976) The influence of bilingualism on cognitive growth: A synthesis of research findings and explanatory hypotheses. *Working Papers on Bilingualism* 9:1–43
- D'Souza AA, Moradzadeh L, Wiseheart M (2018) Musical training, bilingualism, and executive function: working memory and inhibitory control. *Cogn Res Princ Implic* 3(1):11. <https://doi.org/10.1186/s41235-018-0095-6>
- Dash T, Joannette Y, Ansaldo AI (2022) Exploring attention in the bilingualism continuum: A resting-state functional connectivity study. *Brain Lang* 224:105048. <https://doi.org/10.1016/j.bandl.2021.105048>
- Dash T, Berroir P, Joannette Y, Ansaldo AI (2019) Alerting, orienting, and executive control: The effect of bilingualism and age on the subcomponents of attention. *Front Neurol* 10:1122. <https://doi.org/10.3389/fneur.2019.01122>
- DeLuca V, Rothman J, Bialystok E, Platiškas C (2019) Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proc. Natl. Acad. Sci. USA* 116(15):7565–7574. <https://doi.org/10.1073/pnas.1811513116>
- Dhanachandra N, Mangle K, Chanu YJ (2015) Image segmentation using K-means clustering algorithm and subtractive clustering algorithm. *Procedia Comput Sci* 54:764–771
- Dick AS, Garcia NL, Pruden SM et al. (2019) No evidence for a bilingual executive function advantage in the nationally representative ABCD study. *Nat Hum Behav* 3(7):692–701. <https://doi.org/10.1038/s41562-019-0609-3>
- Donnelly S, Brooks PJ, Homer BD (2019) Is there a bilingual advantage on interference-control tasks? A multiverse meta-analysis of global reaction time and interference cost. *Psychon Bull Rev* 26(4):1122–1147. <https://doi.org/10.3758/s13423-019-01567-z>
- Dunn LM, Dunn DM (2007) Peabody Picture Vocabulary Test—Fourth Edition (PPVT-4) [Database record]. *APA PsycTests* <https://doi.org/10.1037/t15144-000>
- Ebert KD (2021) Revisiting the influences of bilingualism and developmental language disorder on children's nonverbal processing speed. *J Speech Lang Hear Res* 64(9):3564–3570. [https://doi.org/10.1044/2021\\_JSLHR-21-00156](https://doi.org/10.1044/2021_JSLHR-21-00156)
- Eriksen BA, Eriksen CW (1974) Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept Psychophys* 16:143–149
- Fernandez-Duque M, Standley M, Roy A, Marian V (in press). Language-dependent memory in bilinguals. In: A de Bruin, D Trenkic (eds) *The encyclopedia of applied linguistics*. Wiley-Blackwell, New Jersey
- Festman J, Rodriguez-Fornells A, Münte TF (2010) Individual differences in control of language interference in late bilinguals are mainly related to general executive abilities. *Behav Brain Funct* 6:5. <https://doi.org/10.1186/1744-9081-6-5>
- Friesen DC, Luo L, Luk G et al. (2015) Proficiency and control in verbal fluency performance across the lifespan for monolinguals and bilinguals. *Lang Cogn Neurosci* 30(3):238–250. <https://doi.org/10.1080/23273798.2014.918630>
- Gollan TH, Montoya RI, Fennema-Notestine C et al. (2005) Bilingualism affects picture naming but not picture classification. *Mem Cognit* 33(7):1220–1234. <https://doi.org/10.3758/bf03193224>
- Green DW (1998) Mental control of the bilingual lexico-semantic system. *Biling (Camb Engl)* 1(2):67–81. <https://doi.org/10.1017/S1366728998000133>
- Grundy JG, Timmer K (2017) Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second Lang Res* 33(3):325–340. <https://doi.org/10.1177/0267658316678286>
- Gullifer JW, Titone D (2020) Characterizing the social diversity of bilingualism using language entropy. *Biling (Camb Engl)* 23(2):283–294. <https://doi.org/10.1017/s1366728919000026>
- Gullifer JW, Koussaie S, Gilbert AC et al. (2021) Bilingual language experience as a multidimensional spectrum: Associations with objective and subjective language proficiency. *Appl Psycholinguist* 42(2):245–278
- Haft SL, Kepinska O, Caballero JN et al. (2019) Attentional fluctuations, cognitive flexibility, and bilingualism in kindergarteners. *Behavioral Sciences* 9:58. <https://doi.org/10.3390/bs9050058>
- Hamann C, Ibrahim AL (2017) Methods for identifying specific language impairment in bilingual populations in Germany. *Front Commun (Lansanne)* 2:16. <https://doi.org/10.3389/fcomm.2017.00016>
- Hansen LB, Macizo P, Duñabeitia JA et al. (2016) Emergent bilingualism and working memory development in school aged children. *Lang Learn* 66:51–75. <https://doi.org/10.1111/lang.12170>
- Hartigan JA, Wong MA (1979) Algorithm AS 136: A K-means clustering algorithm. *J R Stat Soc Ser C Appl Stat* 28:100–108. <https://doi.org/10.2307/2346830>
- Hayakawa S, Shook A, Marian V (2020) When it's harder to ignore than to ignore: Evidence of greater attentional capture from a non-dominant language. *Int J Biling* 24(5-6):999–1016. <https://doi.org/10.1177/1367006920915277>
- Van Heuven WJ, Dijkstra T, Grainger J (1998) Orthographic neighborhood effects in bilingual word recognition. *J Mem Lang* 39(3):458–483. <https://doi.org/10.1006/jmla.1998.2584>
- Ivanova I, Costa A (2008) Does bilingualism hamper lexical access in speech production? *Acta Psychol (Amst)* 127(2):277–288. <https://doi.org/10.1016/j.actpsy.2007.06.003>
- Jiao L, Liu C, Wang R et al. (2019) Working memory demand of a task modulates bilingual advantage in executive functions. *Int J Biling* 23(1):102–117. <https://doi.org/10.1177/1367006917709097>
- Jones SK, Davies-Thompson J, Tree J (2021) Can machines find the bilingual advantage? Machine learning algorithms find no evidence to differentiate between lifelong bilingual and monolingual cognitive profiles. *Front Hum Neurosci* 15:621772. <https://doi.org/10.3389/fnhum.2021.621772>
- Juang LH, Wu MN (2010) MRI brain lesion image detection based on color-converted K-means clustering segmentation. *Measurement* 43(7):941–949
- Kail R, Salthouse TA (1994) Processing speed as a mental capacity. *Acta Psychol (Amst)* 86(2-3):199–225. [https://doi.org/10.1016/0001-6918\(94\)90003-5](https://doi.org/10.1016/0001-6918(94)90003-5)
- Kałamala P, Senderecka M, Wodniecka Z (2022) On the multidimensionality of bilingualism and the unique role of language use. *Biling (Cambridge)* 25(3):471–483. <https://doi.org/10.1017/S1366728921001073>
- Kapa LL, Colombo J (2013) Attentional control in early and later bilingual children. *Cogn Dev* 28(3):233–246. <https://doi.org/10.1016/j.cogdev.2013.01.011>
- Khanmohammadi S, Adibeig N, Shanebandy S (2017) An improved overlapping k-means clustering method for medical applications. *Expert Syst Appl* 67:12–18
- Kovács AM, Mehler J (2009) Cognitive gains in 7-month-old bilingual infants. *Proc Natl Acad Sci USA* 106(16):6556–6560. <https://doi.org/10.1073/pnas.0811323106>
- Kuzyk O, Friend M, Severdija V et al. (2020) Are there cognitive benefits of code-switching in bilingual children? A longitudinal study. *Biling (Camb Engl)* 23(3):542–553. <https://doi.org/10.1017/S1366728918001207>
- Kwon YH, Yoo K, Nguyen H et al. (2021) Predicting multilingual effects on executive function and individual connectomes in children: An ABCD study. *Proc Natl Acad Sci USA* 118(49):e211081118. <https://doi.org/10.1073/pnas.2110811118>
- Lenth R, Singmann H, Love J et al. (2018) *Emmeans*: Estimated marginal means, aka least-squares means. R package version 1(1):3. <https://cran.r-project.org/web/packages/emmeans/emmeans.pdf> Retrieved from

- Ljungberg JK, Hansson P, Andrés P et al. (2013) A longitudinal study of memory advantages in bilinguals. *PLoS One* 8(9):e73029. <https://doi.org/10.1371/journal.pone.0073029>
- Ljungberg JK, Elbe P, Sörman DE (2020) The bilingual effects of linguistic distances on episodic memory and verbal fluency. *Scand J Psychol* 61:195–203. <https://doi.org/10.1111/sjop.12609>
- Luk G, de Sa E, Bialystok E (2011) Is there a relation between onset age of bilingualism and enhancement of cognitive control? *Biling (Camb Engl)* 14(4):588–595. <https://doi.org/10.1017/S1366728911000010>
- Luo L, Craik FIM, Moreno S et al. (2013) Bilingualism interacts with domain in a working memory task: Evidence from aging. *Psychol Aging* 28(1):28–34. <https://doi.org/10.1037/a0030875>
- Luque A, Morgan-Short K (2021) The relationship between cognitive control and second language proficiency. *J Neurolinguistics* 57:100956. <https://doi.org/10.1016/j.jneuroling.2020.100956>
- MacQueen J (1967) Classification and analysis of multivariate observations. In *5th Berkeley Symposium on Mathematical Statistics and Probability* (p. 281–297)
- Marian V (2023) *The Power of Language: How the Codes We Use to Think, Speak, and Live Transform our Minds*. Dutton, New York
- Marian V, Spivey M (2003) Bilingual and monolingual processing of competing lexical items. *Appl Psycholinguist* 24(2):173–193. <https://doi.org/10.1017/S0142716403000092>
- Marian V, Kaushanskaya M (2007) Language context guides memory content. *Psychon Bull Rev* 14:925–933. <https://doi.org/10.3758/BF03194123>
- Marian V, Blumenfeld HK, Kaushanskaya M (2007) The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *J Speech Hear Res* 50(4):940–967. [https://doi.org/10.1044/1092-4388\(2007\)067](https://doi.org/10.1044/1092-4388(2007)067)
- Marian V, Farooqi-Shah Y, Kaushanskaya M et al. (2009) Bilingualism: Consequences for language, cognition, development and the brain. *ASHA Lead* 14(13):10–13. <https://doi.org/10.1044/leader.FTR2.14132009.10>
- Marian V, Hayakawa S, Lam TQ et al. (2018b) Language experience changes audiovisual perception. *Brain Sci* 8(5):85. <https://doi.org/10.3390/brainsci8050085>
- Marian V, Lam TQ, Hayakawa S et al. (2018a) Top-down cognitive and linguistic influences on the suppression of spontaneous otoacoustic emissions. *Front Neurosci* 12:378. <https://doi.org/10.3389/fnins.2018.00378>
- Marton K (2016) Executive control in bilingual children: Factors that influence the outcomes. *Linguist Approaches Biling* 6(5):575–589. <https://doi.org/10.1075/lab.15038.mar>
- McVeigh C, Wylie J, Mulhern G (2019) Verbal and visuospatial working memory in immersion-educated bilingual children. *Int J Biling Educ Biling* 22:505–517. <https://doi.org/10.1080/13670050.2016.1271769>
- Meuter RFI, Allport A (1999) Bilingual language switching in naming: Asymmetrical costs of language selection. *J Mem Lang* 40(1):25–40. <https://doi.org/10.1006/jmla.1998.2602>
- Miyake A, Friedman NP, Emerson MJ et al. (2000) The unity and diversity of executive functions and their contributions to complex “Frontal Lobe” tasks: A latent variable analysis. *J Cogn Psychol (Hove)* 41(1):49–100. <https://doi.org/10.1006/cogp.1999.0734>
- Monnier C, Boiché J, Armandon P et al. (2022) Is bilingualism associated with better working memory capacity? A meta-analysis. *Int J Biling Educ Biling* 25(6):2229–2255. <https://doi.org/10.1080/13670050.2021.1908220>
- Morales J, Calvo A, Bialystok E (2013) Working memory development in monolingual and bilingual children. *J Exp Child Psychol* 114(2):187–202. <https://doi.org/10.1016/j.jecp.2012.09.002>
- Nayak S, Tarullo AR (2020) Error-related negativity (ERN) and ‘hot’ executive function in bilingual and monolingual preschoolers. *Biling (Camb Engl)* 23(4):897–908. <https://doi.org/10.1017/S1366728919000725>
- Ong G, Sewell DK, Weekes B et al. (2017) A diffusion model approach to analysing the bilingual advantage for the flanker task: The role of attentional control processes. *J Neurolinguistics* 43(Part A):28–38. <https://doi.org/10.1016/j.jneuroling.2016.08.002>
- Paap KR, Greenberg ZI (2013) There is no coherent evidence for a bilingual advantage in executive processing. *Cogn Psychol* 66(2):232–258. <https://doi.org/10.1016/j.cogpsych.2012.12.002>
- Persici V, Vihman M, Burro R et al. (2019) Lexical access and competition in bilingual children: The role of proficiency and lexical similarity of the two languages. *J Exp Child Psychol* 179:103–125. <https://doi.org/10.1016/j.jecp.2018.10.002>
- Poarch GJ, van Hell JG (2012) Cross-language activation in children’s speech production: Evidence from second language learners, bilinguals, and trilinguals. *J Exp Child Psychol* 111(3):419–438. <https://doi.org/10.1016/j.jecp.2011.09.008>
- Ratui I, Azuma T (2015) Working memory capacity: Is there a bilingual advantage? *J Cogn Psychol (Hove)* 27(1):1–11. <https://doi.org/10.1080/20445911.2014.976226>
- Roberts G (2021) 2021 Canvass of dual Language and immersion (DLI) programs in U.S. public schools. American Councils Research Center. <https://www.americancouncils.org/research-assessment/studies-and-reports>
- Rosselli M, Ardila A, Lalwani LN et al. (2016) The effect of language proficiency on executive functions in balanced and unbalanced Spanish–English bilinguals. *Biling (Camb Engl)* 19(3):489–503. <https://doi.org/10.1017/S1366728915000309>
- Sandoval TC, Gollan TH, Ferreira VS et al. (2010) What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Biling (Camb Engl)* 13(2):231–252. <https://doi.org/10.1017/S1366728909990514>
- Schroeder SR, Marian V (2012) A bilingual advantage for episodic memory in older adults. *J Cogn Psychol (Hove)* 24(5):591–601. <https://doi.org/10.1080/20445911.2012.669367>
- Schroeder SR, Marian V (2014) Bilingual episodic memory: How speaking multiple languages influences remembering. In: Heredia RR, Altarriba J (eds) *Foundations of bilingual memory*. Foundations Biling Mem. Springer, Berlin, p 111–132. [https://doi.org/10.1007/978-1-4614-9218-4\\_6](https://doi.org/10.1007/978-1-4614-9218-4_6)
- Shook A, Marian V (2019) Covert co-activation of bilinguals’ non-target language: Phonological competition from translations. *Linguist Approaches Biling* 9(2):228–252. <https://doi.org/10.1075/lab.17022.sho>
- Soares SMP, Ong G, Abutalebi J et al. (2019) A diffusion model approach to analyzing performance on the Flanker task: The role of the DLPFC. *Biling (Camb Engl)* 22(5):1194–1208. <https://doi.org/10.1017/S1366728918000974>
- Soveri A, Rodriguez-Fornells A, Laine M (2011) Is there a relationship between language switching and executive functions in bilingualism? Introducing a within group analysis approach. *Front Psychol* 2:183. <https://doi.org/10.3389/fpsyg.2011.00183>
- Sulpizio S, Del Maschio N, Fedeli D, Abutalebi J (2020) Bilingual language processing: A meta-analysis of functional neuroimaging studies. *Neurosci Biobehav Rev* 108:834–853
- Tomoschuk B, Ferreira VS, Gollan TH (2019) When a seven is not a seven: Self-ratings of bilingual language proficiency differ between and within language populations. *Biling (Cambridge)* 22(3):516–536. <https://doi.org/10.1017/S1366728918000421>
- Tse CS, Altarriba J (2014) The relationship between language proficiency and attentional control in Cantonese-English bilingual children: Evidence from Simon, Simon switching, and working memory tasks. *Front Psychol* 5:954. <https://doi.org/10.3389/fpsyg.2014.00954>
- Vaughn KA, Nguyen MVH, Ronderos J et al. (2021) Cortical thickness in bilingual and monolingual children: Relationships to language use and language skill. *Neuroimage* 243:118560. <https://doi.org/10.1016/j.neuroimage.2021.118560>
- Videsott G, Della Rosa PA, Wiater W et al. (2012) How does linguistic competence enhance cognitive functions in children? A study in multilingual children with different linguistic competences. *Biling (Camb Engl)* 15(4):884–895. <https://doi.org/10.1017/S1366728912000119>
- Wang J, Wu X, Zhang C (2005) Support vector machines based on K-means clustering for real-time business intelligence systems. *Int J Bus Intell Data Min* 1(1):54–64
- Weber A, Cutler A (2004) Lexical competition in non-native spoken-word recognition. *J Mem Lang* 50(1):1–25. [https://doi.org/10.1016/S0749-596X\(03\)00105-0](https://doi.org/10.1016/S0749-596X(03)00105-0)
- Weintraub S, Dikmen SS, Heaton RK et al. (2013) Cognition assessment using the NIH Toolbox. *Neurology* 80(11 Suppl 3):S54–S64. <https://doi.org/10.1212/WNL.0b013e3182872ded>
- Wiseheart M, Viswanathan M, Bialystok E (2016) Flexibility in task switching by monolinguals and bilinguals. *Biling (Camb Engl)* 19(1):141–146. <https://doi.org/10.1017/s1366728914000273>
- Wismüller A, Lange O, Dersch DR et al. (2002) Cluster analysis of biomedical image time-series. *Int J Comput Vis* 46:103–128. <https://doi.org/10.1023/A:1013550313321>
- Wu YJ, Thierry G (2010) Chinese-English bilinguals reading English hear Chinese. *J Neurosci* 30(22):7646–7651. <https://doi.org/10.1523/JNEUROSCI.1602-10.2010>
- Yow WQ, Li X (2015) Balanced bilingualism and early age of second language acquisition as the underlying mechanisms of a bilingual executive control advantage: Why variations in bilingual experiences matter. *Front Psychol* 6:164. <https://doi.org/10.3389/fpsyg.2015.00164>
- Zelazo PD (2006) The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nat Protoc* 1:297–301
- Zelazo PD, Anderson JE, Richler J et al. (2013) II. NIH Toolbox Cognition Battery (CB): measuring executive function and attention. *Monogr Soc Res Child Dev* 78(4):16–33. <https://doi.org/10.1111/mono.12032>
- Zhou B, Krott A (2018) Bilingualism enhances attentional control in non-verbal conflict tasks: Evidence from ex-Gaussian analyses. *Biling (Camb Engl)* 21(1):162–180. <https://doi.org/10.1017/S1366728916000869>

## Acknowledgements

This research was supported in part by the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health under Award Number R01HD059858 to Viorica Marian. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

### Author contributions

Ashley Chung-Fat-Yim: Conceptualization, Data curation, Formal analysis, Validation, Writing – original draft and review & editing. Sayuri Hayakawa: Conceptualization, Data curation, Formal analysis, Software, Validation, Visualization, Writing – original draft and review & editing. Viorica Marian: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing - review & editing.

### Competing interests

The authors declare no competing interests.

### Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The studies involving human participants were reviewed and approved by Northwestern University's Institutional Review Board (IRB # STU00023477; Date: December 15, 2009). The scope of approval covered research locations, procedures, methods, measures, statistical analyses, and data monitoring and storage plans, as well as participant-related protocols, including recruitment methods, sample size, inclusion criteria, consent procedures, compensation, and disclosure of potential risks and benefits.

### Informed consent

Written informed consent to participate in the studies was obtained from the participant by experimenters between August 2013 - November 2015. The scope of consent included participation, de-identified data use and sharing, as well as consent to publish. Participants either received course credit or monetary compensation for their time and effort. All participants were fully informed that their anonymity is assured, why the research is

being conducted, how their data will be utilised, and of any risks associated with participating in the study.

### Additional information

**Correspondence** and requests for materials should be addressed to Ashley Chung-Fat-Yim.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2025