languages

Article

The Influence of Multilingual Experience on Executive Function and Structure Learning: Effects in Young Adults in the UK and Singapore

Chrysoula Vassiliu [1,*], Victoria Leong [2,3] and Henriette Hendriks [1]

1 Faculty of Modern and Medieval Languages and Linguistics, University of Cambridge, Cambridge CB2 1TN, UK; hpjmh2@cam.ac.uk
2 Early Mental Potential and Wellbeing Research (EMPOWER) Centre, Nanyang Technological University, Singapore 639798, Singapore; victorialeong@ntu.edu.sg
3 Department of Pediatrics, University of Cambridge, Cambridge CB2 1TN, UK
* Correspondence: cv335@cam.ac.uk

Abstract: Most studies regarding the relationship between multilingualism and cognitive control reduce linguistic diversity to a dichotomous comparison, viz., monolinguals vs. bilinguals, failing to capture the multifactorial nature of multilingualism. Language research is largely restricted to the Global North, albeit most of the world's population resides in the Global South, limiting the interpretability of the existing literature. Cognitive performance is assessed using very few tasks, yielding unreliable measurements. In this study, we identify the manner in which multilingual experiences influence cognitive performance in diverse sociolinguistic contexts. Young adults from the UK (n = 51, mean age = 24.0, SD = 3.18) and Singapore (n = 36, mean age = 21.3, SD = 2.15) were tested using an extensive battery of cognitive tasks, including cognitive flexibility (CF), working memory (WM), inhibition, and structure learning (SL). Information on language proficiency, use, age of acquisition, and frequency of switching was collected. The effects of various linguistic factors on the cognitive performance of each group were assessed using multiple linear regression models. The UK and Singapore samples exhibited significantly different linguistic profiles, which in turn dissimilarly influenced their cognitive performance. Our study underscores the necessity for more research in the Global South, challenging the prevailing Northern-centric focus on the multilingualism–cognition relationship.

Keywords: multilingualism; executive function (EF); structure learning (SL); sociolinguistic context; Singapore; UK

check for updates

Citation: Vassiliu, Chrysoula, Victoria Leong, and Henriette Hendriks. 2024. The Influence of Multilingual Experience on Executive Function and Structure Learning: Effects in Young Adults in the UK and Singapore. *Languages* 9: 136. https://doi.org/10.3390/languages9040136

Academic Editor: Peter Ecke

Received: 16 November 2023
Revised: 26 February 2024
Accepted: 25 March 2024
Published: 8 April 2024

Correction Statement: This article has been republished with a minor change. The change does not affect the scientific content of the article and further details are available within the backmatter of the website version of this article.

1. Introduction

The effects of multiple-language control on non-linguistic cognitive control are highly debated. Behavioural studies have reported better performance for bilinguals across a range of cognitive tests, including inhibition tasks (Bialystok et al. 2004; Hernández et al. 2010); cognitive flexibility (CF) tasks, for example, task-switching (Prior and Macwhinney 2010; Stasenko et al. 2017); and on working memory (WM), both visuospatial (Bialystok 2009; Kerrigan et al. 2017) and verbal (Antón et al. 2019; Blom et al. 2014). Advantages in statistical learning have also been reported (Bonifacci et al. 2011; Verhagen and de Bree 2021). However, an important segment of the literature reports null or even negative results (Antón et al. 2014; Carlson and Meltzoff 2008; Duñabeitia et al. 2014; Paap et al. 2014; Paap and Greenberg 2013).

The inconsistency in the literature can be attributed to multiple reasons. Most studies simply perform cross-sectional comparisons of monolingual vs. bilingual groups. Such designs reflect a binary understanding of bilingualism. They fail to capture the diversity and multifactorial nature of the phenomenon and cannot account for differences within bi-/multilingual groups (Luk and Bialystok 2013). Factors such as proficiency, language use, age of acquisition, or frequency of language switching are not considered, yet they may be a source of meaningful differences.

Additionally, these between-group comparisons may be confounded by variables such as socio-economic status (SES), education, intelligence, or immigration status, which often correlate with general cognitive performance (Fuller-Thomson and Kuh 2014; Goldberg et al. 2008; Kirk et al. 2014; Yang et al. 2011). In other words, differences between monolingual and bilingual groups in regard to/on non-linguistic variables may be driving spurious effects or masking genuine impacts (Bak 2015), so it is very important that these variables are controlled for.

The limited scope of cognitive assessments, which usually encompass one task per cognitive component, neglect the intricate architecture of executive functions (EFs) (Jylkkä et al. 2018). While EFs encompass inhibition, WM, and CF (Miyake et al. 2000), the interrelatedness and potential subcomponents of these functions remain underexplored.

Finally, most existing literature is focused on the Global North, even though the majority of the world's population resides in the Global South (Makoni et al. 2022). This unavoidably produces a distorted image of the relationship between bi-/multilingualism and cognitive processing (Bak and Alladi 2016). Different sociolinguistic contexts are highly likely to influence this relationship, so research within different environments is crucial. For example, while countries in the Global North frequently have a single official national language, several countries in the Global South have multiple official languages. And while multilingualism may be present in certain areas of the Global North as well, societal perspectives tend to be different. Monoglossic ideologies are more dominant in the Global North (especially in Europe), whereas numerous contexts in the Global South embrace multilingualism as standard practice (Bunk and Wiese 2024).

This study examines the impact of managing multiple languages on domain-general cognitive control, addressing the aforementioned weaknesses. Instead of comparing distinct groups of monolinguals and bilinguals, we are operationalising "lingualism" as a continuum (Wigdorowitz et al. 2022), breaking it down into specific continuous linguistic variables (Gallo et al. 2021) and evaluating their effect on cognitive performance, while controlling for background variables, such as SES, intelligence, and age. The term "bilingual" is used to refer to people using two or more languages (or dialects) in their daily lives (Grosjean 2021). This view of bilingualism places more emphasis on language use rather than fluency. When the term "multilingual" is used, it is employed to emphasise that speakers use more than two languages.

Following recent studies, we have included various continuous measures of linguistic experience. These include age of onset of second-language acquisition (L2 AoA) (Luk et al. 2011; Soveri et al. 2011), balanced proficiency, and language use (Yow and Li 2015), as well as the relatively new measure of language entropy, which reflects the social diversity of language use (Gullifer and Titone 2020). High entropy indicates the use of multiple languages in an integrated fashion, and low entropy indicates compartmentalised use in distinct contexts or the use of a single language. We also evaluate the frequency and context of language switching (Jylkkä et al. 2017; Verreyt et al. 2016; Woumans et al. 2019), in line with the adaptive control hypothesis (ACH) (Green and Abutalebi 2013). The ACH proposes that there are distinct contexts of bi-/multilingual language use. In the single-language context, bilinguals use each of their languages in distinct environments. In the dual-language context, they use their languages within the same environment but with different interlocutors. In dense code-switching, they use their languages with the same interlocutor, making frequent switches between languages, even within a sentence. Each context is assumed to pose unique linguistic and cognitive demands.

To measure the influences of linguistic experience on EF and statistical learning, we employ an extensive battery of cognitive tests, which includes tasks on CF, WM, inhibition, as well as a statistical learning paradigm, here a specific instantiation of learning under uncertainty, termed structure learning (SL) (Wang et al. 2017a, 2017b). We thus hope to detect more nuanced effects and pinpoint specific aspects of cognition susceptible to multiple-language experience.

Finally, we investigate the effects of linguistic experience on cognition in two groups, one from the UK and one from Singapore. We thus broaden the scope of the existing literature by including a country from the Global South[1] (World Population Review 2023). While research in Singapore is not as scarce as it is in other countries of the Global South (e.g., Hartanto and Yang 2016; Yow and Li 2015), the Singaporean sociolinguistic context remains underexplored. Ooi et al. (2018) compared bilinguals from Singapore with monolinguals and bilinguals in Edinburgh, who differed in switching tendencies and L2 AoA. While the authors focused on specific linguistic experiences to distinguish the groups and did not employ a binary categorisation, they still used a between-group design. In any case, assessing bilingual (and multilingual) populations from different countries and sociolinguistic environments is very informative because such investigations help to determine how the relationship between multilingualism and cognitive processing can vary as a function of linguistic context.

Singapore is an extremely interesting case, as it is a highly multilingual environment, with four official languages (English, Mandarin, Malay, Tamil), along with several other dialects (Ooi et al. 2018). Not all Singaporeans speak or understand all the official languages. In fact, recent research indicates that over the past decades, there has been a shift to "English Plus"; that is, Singaporeans mostly speak English, along with one other language, depending on their ethnic group (Cavallaro and Chin 2014). This is due to language policies: English has been defined as the main medium of instruction in the educational system, and the mother tongues are taught as L2s (Dixon 2005). A recent study with young adults (university students) indicates that most of the participants exhibited bilingual or trilingual identities (Siemund et al. 2014).

While many Singaporeans may not actively use more than two languages, they are still exposed to significant linguistic diversity, as speakers of the various ethnic and linguistic groups live side by side and interact on a daily basis (Leimgruber 2013). Singlish, a creolised form of English arising due to the language contact between British colonisers and indigenous Chinese, Malay, and Indian populations, is also frequently used and is part of the Singaporean identity (Siemund et al. 2014). The use of Singlish further increases the linguistic variety to which Singaporeans are exposed.

The United Kingdom (UK), on the other hand, is predominantly monolingual, with English as its official language, reported as the main language by 91.1% (52.6 million) of usual residents, according to Census 2021 data. The most common main languages other than English (English or Welsh in Wales) are: Polish (1.1%, 612,000), Romanian (0.8%, 472,000), Panjabi (0.5%, 291,000), and Urdu (0.5%, 270,000) (Office for National Statistics 2022). Research suggests that there is significantly less linguistic diversity in the UK than in more multilingual countries (e.g., South Africa) (Wigdorowitz et al. 2022), with the majority of the population speaking only English (Gough 2023). However, there may be variation in the linguistic diversity in the UK (and Europe, more generally) in different areas (e.g., rural vs. urban). When compared to more rural areas or smaller cities, larger cities such as London, Brussels, or Berlin may exhibit more linguistic diversity as a result of frequent interaction between groups, even though societal attitudes still privilege the major languages differently than is the case in Singapore (Bunk and Wiese 2024; Siemund et al. 2014).

In summary, this study sheds light on the relationship between multilingual experience and cognitive ability, using a variety of tasks and examining two diverse sociolinguistic contexts. To this effect, the study has two main aims: (i) to assess the differences in the linguistic profiles of young adults in the UK and Singapore; (ii) to identify whether different multilingualism experience factors influence cognitive performance and determine whether the relationship between linguistic experience and cognitive performance is different in these two contexts (UK and Singapore).

2. Materials and Methods

2.1. Participants

All participants were university students in either the UK (University of Cambridge, Cambridge, UK) or Singapore (Nanyang Technological University—NTU, Singapore). The UK sample included 51 participants (33 females; aged 19–30; *M* = 24.0; *SD* = 3.18), and the Singapore sample comprised 36 participants (27 females; aged 18–27; *M* = 21.3; *SD* = 2.15). The participants received monetary compensation for the participation (GBP 8 per hour of participation for participants in the UK and SGD 20 for the participants in Singapore).

All participants exhibited normal or corrected-to-normal vision and had no history of neurological impairment or brain injury.

All participants rated themselves as highly proficient in English[2]. We did not focus on a specific language combination, hoping to increase the ecological validity of the study and make our samples more representative in terms of the language realities of each of the contexts. We encouraged participants to report all the languages they have used in their lifespan, but they were asked to indicate proficiency and daily use for each language at the time of the study. In the UK, the languages other than English reported by the participants were: Afrikaans, Arabic, Bulgarian, Danish, French, German, Greek, Hokkien, Indonesian, Italian, Korean, Mandarin, Norwegian, Portuguese, Russian, Serbian, Sinhalese, Spanish, and isiZulu. In Singapore, the languages, other than English, reported were: Cantonese, French, German, Hindi, Hokkien, Indonesian, Japanese, Korean, Malay, Mandarin, Tamil, Telugu, and Teochew. More information on self-reported language proficiency and use in the different groups is presented in Section 3.1.

2.2. Materials and Procedure

2.2.1. Procedure

Participants were tested individually following a remote guided protocol for supervised web-based testing that was developed during the pandemic [for a detailed description of the protocol, see Leong et al. (2022)]. To ensure high-quality data, we required participants to have access to a personal computer with Windows 7 or 10 or MAC OS, minimum 8GB RAM, a webcam and microphone, administrator rights to download and install software, and a reliable internet connection. We ensured that all the requirements were met by employing a screening questionnaire administered before accepting individuals for participation. Furthermore, by using the remote guided testing (RGT) protocol, we could make sure that participants were concentrating when performing the tasks, as they were connected via a video call with the experimenter, sharing their screens throughout testing.

The participants also signed a consent form and completed a demographics and language questionnaire before beginning the testing session. During the testing session, participants were assessed on a battery of cognitive tasks measuring EF and SL. They also completed a test of non-verbal intelligence, in which they were required to solve matrix puzzles of increasing difficulty. A test of verbal intelligence (the Vocabulary subtest from the Wechsler Abbreviated Scale of Intelligence, 2nd edition (WASI-II); Wechsler (2011)) was also administered. For the remaining cognitive tests, five pseudo-randomised orders of tasks were created, and each participant was assigned to one of these. The testing session lasted for about 2.5 to 3 h in total. Frequent breaks between tasks were planned by the experimenter, but we also checked with the participants throughout the process to see if they needed additional breaks.

The study was approved by the University of Cambridge and NTU Singapore review boards.

2.2.2. Background Measures

Demographics

Information on demographics, including gender, age, and SES, was collected using an extensive demographics questionnaire (Appendix A). For SES, we used a composite score ranging from 0 to 1, which comprised measures of parental and maternal education, as well as annual income. Higher scores indicate higher SES.

Intelligence[3]

Non-Verbal Intelligence (NVIQ)

To measure NVIQ, we used a seven-item progressive matrices test, inspired by the Raven's Standard Progressive Matrices (Raven et al. 2000). The task was hosted on the iABC website, a Cambridge University in-house platform (*I-ABC: A Personalised Learning Study—Adaptive Brain Lab* n.d.). Participants were presented with a 3×3 grid of images, with the final image left blank, and were asked to select which of six options best completed the grid.

The duration of the task is about 7 min.

Verbal Intelligence (VIQ)

We used the Wechsler Abbreviated Scale of Intelligence, 2nd edition (WASI-II) (Wechsler 2011) Vocabulary subtest to measure the participants' VIQ. The Vocabulary subtest requires participants to define up to 30 words. It assesses participants' word understanding and reflects language development, expressive language skills, and retrieval of information from long-term memory. Raw scores range from 3 to 59 (for this age group), but they were transformed based on the age-related standards. The adjusted scores ranged from 20 to 80.

The duration of the task is about 6 min.

The background demographics for each group are summarised in Tables 1 and 2. The groups differed significantly in regard to age, SES, NVIQ, and VIQ.

Table 1. Means and standard deviations (SD) for Age, SES, intelligence.

Background Measure	UK		Singapore			
	Mean	SD	Mean	SD	t	p
Age	24.0	3.18	21.3	2.15	−4.36	<0.0001
SES [a]	0.59	0.17	0.45	0.17	−3.76	0.0003
NVIQ [b]	0.65	0.24	0.48	0.24	−3.31	0.001
VIQ [c]	63.5	10.5	51.7	7.89	−5.68	<0.0001

[a] SES values could range from 0 to 1. [b] NVIQ values could range from 0 to 1. [c] VIQ values could range 20 to 80.

Table 2. Information on gender, dominance in English, and videogame usage.

Background Measure	UK	Singapore		
	N	N	X^2	p
Female	33 (of 51)	27 (of 36)	1.05	0.307
English-dominant (use)	42 (of 51)	31 (of 36)	0.22	0.638
Videogame usage	18 (of 51)	20 (of 36)	3.52	0.061

For this reason, as discussed in the Section 2.3, we did not perform direct comparisons between the two contexts; that is, we did not add Country as a predictor in our models. Instead, we employed separate models for each context. Given the great differences in these demographic measures, as well as other factors outside of our experimental control (such as the educational system), a difference between countries would not necessarily be meaningful or easily interpreted. It is the language-related effects present in each country that are of interest to us, rather than the effect of the country per se.

2.2.3. Language Measures

In our attempt to quantify multilingualism as a multifactorial continuum, we used an extensive questionnaire, collecting information on several language background variables.

Specifically, we required participants to provide the following information: languages they use (or have used), age of acquisition of each language, and hours (and percentage) of daily use. They also reported the context of acquisition and the context of use, but these data are not analysed in the present study, as we only focused on continuous variables.

Following the adaptive control hypothesis (ACH) (Green and Abutalebi 2013), we asked participants to rate how frequently they use their languages in a single-language, a dual-language, and in a dense code-switching context. The values ranged from 1 (never) to 5 (always). The full language questionnaire, with the exact phrasing used in the questions, can be found in Appendix B.

Finally, the questionnaire included a section in which participants had to rate their proficiency in listening, reading, spoken interaction, spoken production, and writing in each of their languages. To prevent participants from providing mechanical responses, we did not ask them to simply select a value, but provided them with different descriptions of proficiency, based on the Common European Framework of Reference for Languages (CEFR) (Council of Europe 2020). These were then converted to values ranging from 1 to 6.[4]

Based on this information and following the methods in recent literature regarding the effects of multilingualism on cognitive control (Li et al. 2021; Yow and Li 2015), we calculated the following indices: Balanced Proficiency, Balanced Usage, and Entropy. Balanced Proficiency was calculated as the difference between Proficiency_1 (the highest proficiency score of the two most proficient languages) and Proficiency_2 (the lowest proficiency score of the two languages).[5] Balanced Usage was calculated as the difference between Usage_1 (the highest usage score of the two most used languages) and Usage_2 (the lowest usage score of the two languages).[6] Higher scores in Balanced Proficiency and Balanced Usage indicated greater differences between the languages, and, therefore, greater "imbalance", whereas a value of 0 indicated perfect balance. We focused on the first two languages to make our results comparable to those from existing studies, but also because all of our participants mainly used (and were most proficient in) the first two languages they reported. Nevertheless, as we possessed information on all the languages that each participant used, we also calculated a richer metric of linguistic diversity, that is, Language Entropy (Gullifer et al. 2018; Gullifer and Titone 2020). This is particularly relevant for the Singapore context, which is highly multilingual, with four official languages (English, Mandarin, Malay, and Tamil) (Siemund et al. 2014). We used the Language Entropy R package (Gullifer and Titone 2018) for our evaluation. Unlike the Balanced Usage metric, Language Entropy allowed us to take into account the usage of all of a participant's languages. The participants were asked to list all of their languages and the percentage of use for each one. The sum of the percentages added up to 100%. The percentages were converted to proportions (e.g., 47% to 0.47), and the following equation was used to calculate Language Entropy:

$$H = -\sum_{i=1}^{n} P_i log_2(P_i)$$

Here, n represents the possible languages, and P_i is the proportion of use of language_i. Language Entropy scores range from 0 to $log_2(n)$. Lower entropy scores indicate lower diversity and the use of predominantly one language in a compartmentalised fashion. Higher entropy scores indicate higher language diversity and a more balanced use of languages, as well as language switching. Participants who only used one language had a Language Entropy score of 0.

2.2.4. EF and SL Measures

Cognitive Flexibility

A total of five tasks were administered to measure CF.

(1) Task-Set switching (TSS)

In Task-set switching (TSS), adapted from Kehagia et al. (2017), participants had to categorise the stimuli on their screens based on certain rules. For each trial, a digit (1 to

9) and a letter (A, E, I, U, C, F, T, or X) were presented in the centre of the screen. The participants were required to switch between two tasks, one concerning the number and one the letter. The target stimuli in each trial were accompanied by a cue indicating whether the participants had to focus on the number or the letter. A circular shape (a vertical or horizontal ellipse) was used as a cue for the letter task. An angular shape (square or rhomb) was used as a cue for the number task. The stimuli (letter and number) were presented one above the other. In the letter task, when the circular-shape cue appeared, participants had to decide whether the presented letter was a vowel or a consonant. For vowels, participants had to press the left-arrow key on their keyboard, and for consonants, they were to press the right-arrow key. In the number task, the task with the angular-shaped cue, participants had to decide whether the number was odd or even. They had to press the left-arrow key for odd numbers and the right-arrow key for even numbers.

There were four blocks in total, each comprising 49 trials. The first trial of each block was excluded from the analysis, as it could not be regarded as either a switch or repeat. A *switch trial* is defined as a trial in which the participant has to perform a task different from the one they performed in the previous trial, i.e., switching from performing the number task to performing the letter task, and the switch was indicated by a change in cue (from angular to circular). In *repeat trials*, participants would perform the same task as the one performed in the previous trial, (e.g., continue performing the number task), so the cue would not change (remaining angular). The ratio of switch and repeat trials in each block was 1:2. In each trial, participants initially saw the cue for 300 ms, and then the letter and number pair appeared on the screen, inside the cue. The cue and stimuli remained on screen until the participants provided a response or for a maximum 2000 ms. After that, a fixation cross appeared in the centre of the screen for a variable inter-trial interval of 1700, 1825 or 1950 ms. No feedback was offered after selection. Before the experimental part of the task began, a very detailed set of instructions was provided to the participants, and they practiced on five sample trials, during which they were given feedback for their responses. The task was administered through Gorilla (*Gorilla* 2022). The duration of the task is about 25 min (see Figure 1).

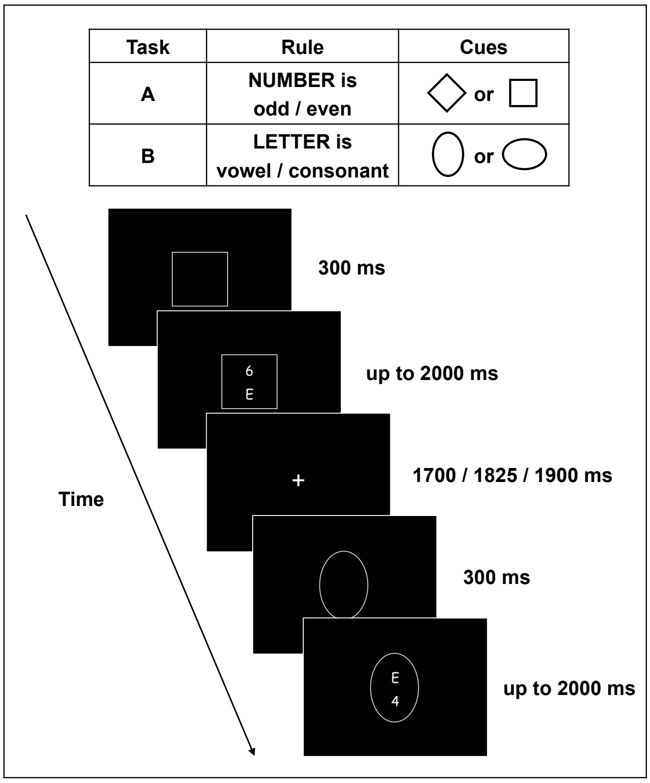

Figure 1. Example trial sequence of task-set switching (TSS).

The mean accuracy for switch and non-switch trials was calculated for each participant. The mean reaction time (RT) for switch and non-switch trials was also calculated, after error trials were removed, along with trials faster than 300 ms, and trials > 2.5 SD from the mean. The TSS switch cost was then calculated as follows:

1. TSS switch Accuracy cost = mean Accuracy in switch trials − mean Accuracy in non-switch trials
2. TSS switch RT cost = mean RT in switch trials − mean RT in non-switch trials

(2) Trail Making Test (TMT)

The Trail Making Test (TMT) (Armitage 1946) is a test of visual attention and task switching. We administered a computerised version of the task hosted on Millisecond-Inquisit (*Inquisit 6* 2021). The participants were asked to use the mouse to connect a series of nodes as quickly and accurately as possible, without interrupting the course of the line. When the participants made an error, the software automatically interrupted the line and required them to continue from the last correct node they had reached.

The task consists of two trails, Trail A and Trail B. In Trail A, participants are required to connect nodes numbered from 1 to 25, in ascending order. In Trail B, numbers alternate with letters, so participants must switch between the two, going from 1 to A to 2 to B to 3 to C and so on, up to the number 13. Participants completed two short training trials, one before Trail A and one before Trail B. The duration of the task is about 3 min (see Figure 2).

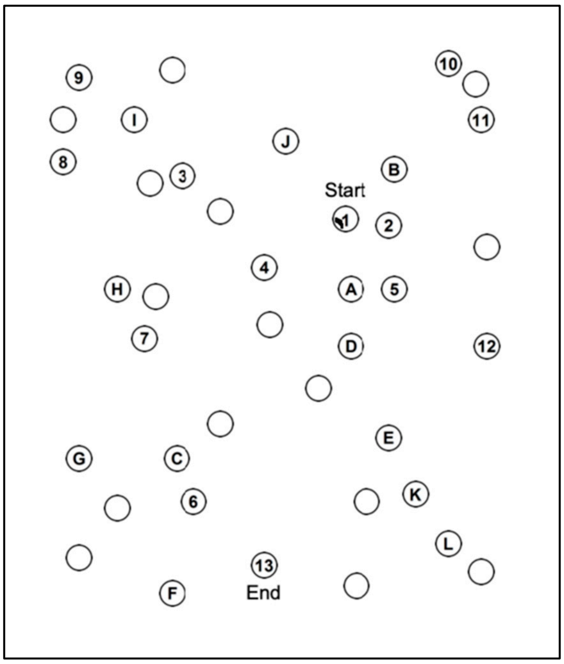

Figure 2. Trail B of the Trail Making Test (TMT).

The outcome measures yielded from TMT are latency (response time) and number of errors for Trail A and Trail B. We used latency scores to calculate the Trail B:A ratio, which was then entered in our analyses.

(3) Wisconsin Card Sorting Test (WCST)

The Wisconsin Card Sorting Test (WCST) evaluates the ability to adapt to unannounced rule changes (Grant and Berg 1948). We implemented a computerised version of the task, using the iABC platform.

The participants are required to sort cards into four different categories. No instructions are given regarding the categorisation rules, but participants receive trial-level feedback after each selection, that is, whether they sorted the card correctly. The four categories appear as four different cards, each containing spaceships of different shapes and colours. These are: one blue spaceship of shape 1, two orange spaceships of shape 2, three yellow spaceships of shape 3, and four green spaceships of shape 4. The cards that the participants are asked to sort have similar designs and vary in colour (four options), shape (four options), and number (four options). In this version of the task, the participants must sort 60 trials in total. Every 10 trials, the rule changes. Each rule (colour, shape, number) is tested two times. Participants have a maximum of 5000 ms in which to give a response. The duration of the task is about 3 min (see Figure 3).

Figure 3. Example trial of the iABC Wisconsin Card Sorting Test (WCST).

The number of perseverative errors, that is, erroneous responses showing that the participant has failed to switch to the new relevant rule, was used in analyses.

(4) Intra-Extra Dimensional Set Shift (IED) task

The Intra-Extra Dimensional Set Shift (IED) task (Robbins et al. 1994) is a test of rule acquisition and reversal. It is considered an analogue of the WCST, yet with fewer dimensions (colour and shape). The IED task involves visual discrimination, set formation, shifting, and flexibility of attention. The task was administered through the CANTAB® cognitive assessment software (Cambridge Cognition 2019).

The participants are presented with two cards in each trial and are asked to choose one of the two. They receive feedback after each selection (the words “correct” or “incorrect” appear on the screen). They must figure out the rule through trial and error. After they reach the criterion of six correct responses, the stimuli and/or rule changes. In the first stages, the task only involves simple stimuli consisting of one dimension, for example, two white lines differing in shape. As the task proceeds, the stimuli become more complex, for example, white lines and pink shapes; thus, the number of potentially relevant dimensions increases to two, i.e., shape and colour. The rule shifts are initially intra-dimensional; thus, for example, the white lines remain the only relevant dimension (colour). Later in the task, an extra-dimensional shift in the rule occurs, so the pink shapes become the pertinent dimension (shape). There are nine stages in total, and the extra-dimensional shift occurs in Stage 8. If the participants fail to reach the criterion of six correct responses after 50 trials, the task terminates. Explicit instructions are offered through voice-over technology in the first few trials of the task until the participant is familiarised with the process. The duration of the task is about 7 min (see Figure 4).

The outcome measures included the total number of errors (which included an adjustment of 25 errors per stage, for participants who did not complete all stages) and the number of errors in Stage 8, which was the extra-dimensional shift stage. Again, for participants who did not reach Stage 8, 25 errors were assumed. The former is an index of the participants’ overall efficiency in attempting the test, whereas the latter is a measure of their ability to shift attentional set.

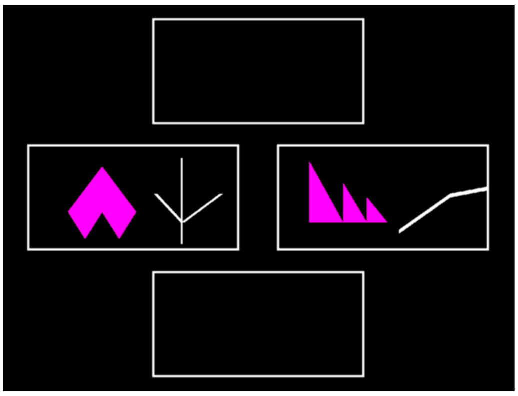

Figure 4. Example trial of the Intra-Extra Dimensional (IED) Set Shift task.

(5) Probabilistic Reversal (PR) Learning task

In the Probabilistic Reversal (PR) Learning task (Cools et al. 2002), participants were presented with two stimuli (two different spaceships), each with a specific probability of being correct (80% vs. 20%). The participants were asked to select one of the two spaceships, and they would receive feedback after each selection. The rule defining the most probably correct stimulus changed at some point without warning, and the probabilities were then reversed. The participants had to learn and adapt to the new rule. There were 80 trials in total, 40 of each probabilistic structure. The task was administered through the iABC platform. The duration of the task is about 4 min (see Figure 5).

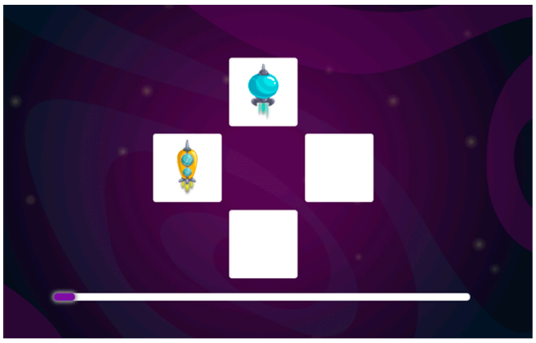

Figure 5. Example trial of the Probabilistic Reversal (PR) Learning task.

The outcome measure used was the number of perseverative errors participants made after the probabilities had been switched.

Inhibition

(1) Stroop task

The Stroop task (Stroop 1935) measures the interference of word meaning with the naming of the colour in which the words are written, as indicated by RT and Accuracy differences between colour-meaning incongruent and congruent conditions. It is a task of inhibitory control. We implemented a computerised version of the task requiring keyboard responding, hosted on the Millisecond-Inquisit platform (*Inquisit 6* 2021).

Participants see names of colours, written in colour, and are asked to indicate the colour of the word (and not its meaning) through key press, as quickly and accurately as they can. Specifically, they were required to press "D" for red, "F" for green, "J" for blue, and "K" for yellow. The colour of the presented word may or may not match the meaning (e.g., "RED" printed in red or green). In congruent trials, the meaning of the word and its colour are the same, whereas in incongruent trials, they are not. The stimuli stay on the screen until a response is provided, and latencies are measured from the onset of the stimuli.

If the participants give an incorrect response, an X is shown on the screen for 400 ms, and then a blank screen appears for 200 ms. If their response is correct, no feedback is given. There were 180 trials in total, 75% of which were congruent and 25% incongruent. The different keys and the colours they corresponded to were present on the screen throughout the task. The duration of the task is about 8 min (see Figure 6).

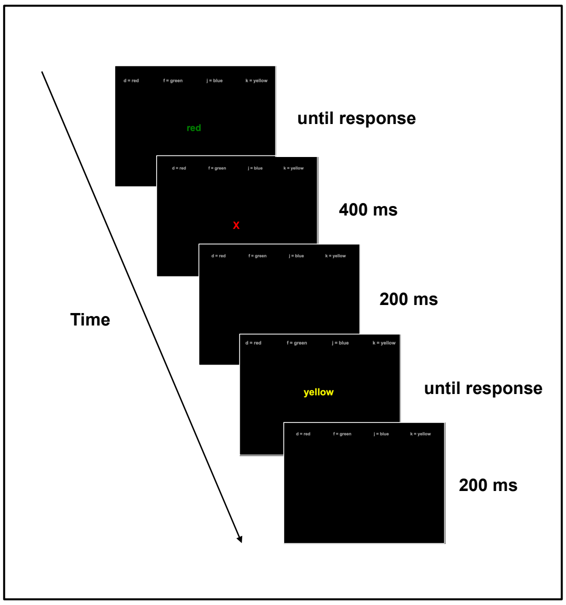

Figure 6. Example sequence of the Stroop task.

The mean accuracy for the congruent and incongruent trials was calculated for each participant. The mean RT for the congruent and incongruent trials was also calculated, after error trials were removed, along with trials faster than 300 ms, and trials > 2.5 SD, from the mean. The Stroop cost was then calculated as follows:

1. Stroop Accuracy cost = mean Accuracy in incongruent trials − mean Accuracy in congruent trials
2. Stroop RT cost = mean RT in incongruent trials − mean RT in congruent trials

Working Memory

Two tasks were used to measure WM, one for visuospatial and one for verbal WM.

(1) Spatial Working Memory (SWM) task

We used the Spatial Working Memory (SWM) task from the CANTAB® (Cambridge, UK) cognitive assessment battery (Cambridge Cognition 2019; Robbins et al. 1994). This is a self-ordered EF task requiring the retention and manipulation of visuospatial information.

The participants view some coloured boxes on the screen. They are asked to find all the yellow tokens and are told there is one in each box, so once they have found it, they need not revisit that box again. In the first stage, there are 4 boxes, and this number gradually increases to 6, 8, and 12. The colour and position of the boxes change as the participants progress through the stages. Explicit instructions are offered through voice-over technology in the first few trials of the task until the participant is familiarised with the process. The duration of the task is about 8 min (see Figure 7).

Outcome measures include the number of errors in each stage (selecting the same box twice in a run, even though it has been found to be empty, and revisiting boxes in which a token has already been found), as well as strategy. Strategy reflects the number of times a participant begins a new search pattern at the same box where they started their previous search. Starting one's search consistently from the same box demonstrates planned strategy rather than random selection. Therefore, a lower score indicates better strategy, with a score of 1 indicating that the participant always began their searches at the same box.

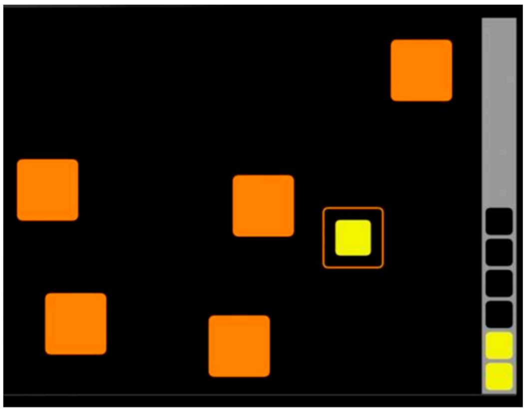

Figure 7. Example trial of the Spatial Working Memory (SWM) task.

(2) Backward Digit Span (BDS) task

We measured executive verbal Working Memory with the Backward Digit Span (BDS) task (Blackburn and Benton 1957). The stimuli (random numbers from 1 to 9) were presented auditorily, and participants were asked to repeat them in reverse order. There were eight stages in total, each with two trials. The participants started at Stage 1 with two digits and moved to the next stage if they answered at least one trial correctly. In Stage 8, the final stage, participants had to recall nine numbers backwards. The administration stopped once a participant answered both trials in a stage incorrectly. A practice trial using two numbers was offered in the beginning. The duration of the task is about 8 min.

The outcome measure was the number of correct responses, ranging from 0 to 16.

Statistical Learning

Statistical learning was measured through a specific instantiation of learning under uncertainty, termed the structure learning (SL) task.

(1) Structure Learning (SL) task

To measure statistical learning ability, we used a novel methodological framework, coined Structure Learning (SL) (Karlaftis et al. 2019; Wang et al. 2017a, 2017b). A detailed description of the SL task can be found in (Wang et al. 2017a, 2017b). Briefly, it is a visual statistical learning paradigm, in which participants are required to learn complex context-based probabilistic contingencies and not just to memorise simple frequency statistics. Participants are exposed to sequences of visual stimuli and are periodically asked to predict the upcoming symbol, without receiving trial-level feedback (see Figure 8).

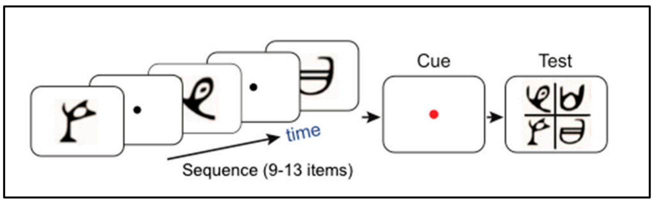

Figure 8. Example sequence of the Structure Learning (SL) task; adapted from R. Wang et al. (2017a).

In the version of the task employed in this study, four highly discriminable visual symbols were used. The sequences of symbols were generated by a 1st-level Markov model, which means that each new symbol depends on the symbol that appeared immediately before it. However, at each time point, the generated symbol is determined probabilistically. For each given symbol, only one of two others are allowed to follow, one with high

(80%) and one with low (20%) probability. This makes the sequence stochastic and not deterministic. In other words, symbol A is most likely (80%) followed by symbol B, but 20% of the time, it will be followed by symbol C. The model is summarised below:

Level 1		Target			
		A	B	C	D
Context	A		0.8	0.2	
	B			0.8	0.2
	C	0.2			0.8
	D	0.8	0.2		

Participants were familiarised with the process over five practice trials. The duration of the task is about 55 min.

The outcome measures used for the analyses were a performance index (PI) and a strategy index [strategy integral curve difference (ICD)]. The former indicates how closely the participants' responses reflected the presented probabilistic structure, with higher scores denoting better performance. The latter indicates the strategy they used. Strategy ICD values close to 0 reflected a matching strategy, and larger positive values reflected a maximising strategy. Negative values indicated random performance. Participants employ a matching strategy when they select symbols by trying to follow the presented structure as closely as possible, while with the maximisation strategy, participants consistently select the most probable symbol.

The performance index (PI) was calculated as the difference in performance between the final two blocks (Blocks 6 and 7) and the initial two blocks (Blocks 1 and 2). This estimate is preferrable to an average PI score, as it more accurately depicts improvement in performance and hence, learning. Strategy ICD was computed across the seven blocks. Details on the computation of PI and strategy ICD can be found in the original papers by Wang et al. (2017a, 2017b).

2.3. Analysis

In this study, we aimed to assess the influence of multiple multilingual experience factors, attempting to identify those that best capture the effects of multilingualism on cognitive performance. We also tried to evaluate cognitive performance in a thorough and extensive manner, using a variety of tasks of CF, inhibition, WM, and SL.

Multiple linear regression models were used separately for each of the cognitive measures (described in Section 2.2.4), with the language background variables entered as predictors (described in Section 2.2.3), following the methods of recent papers employing a similar approach (Li et al. 2021; Soveri et al. 2011; Yow and Li 2015). We employed separate models for each context; that is, we did not add Country as a predictor in the models but conducted distinct analyses for the UK and Singapore. This is because: (1) the two samples differed significantly regarding most background variables (Age, SES, NVIQ, VIQ); (2) the UK and Singapore are distinct environments with different education systems, which may also have dissimilar impacts on EF development. This means that even if a significant Country effect were to arise, it would not necessarily be meaningful, and it could have been driven by factors outside our control and experimental design. It is the language-related effects present in each country that are of interest to us, rather than the effect of country per se.

Before our main analysis, we performed a data-cleaning procedure on the tasks that included RT indices. Specifically, in TSS, the first trial of each block was removed of the analysis, as it could not be evaluated as either a switch or repeat. When calculating RT costs, we had to remove the error and post-error trials because the type of trial could not be determined after an error had occurred. We then removed responses faster than 300 ms, as well as those that were 2.5 SD slower than the mean. For the Stroop task, we excluded trials with responses faster than 300 ms, as well as those that were 2.5 SD above the mean in the incongruent and congruent condition, respectively.

In the UK group, one participant was excluded from the TSS analysis due to too many excessively slow or fast responses, which together resulted in more than 35% data loss for that participant. One participant was not included in the TMT analysis because due to technical issues, their responses were not recorded. We decided to not exclude the participants completely (i.e., we kept their scores for the tasks they had completed without issues), as removing them would reduce our sample size and because generalised linear models (GLM), which were used for our analysis, can handle missing values.

To simplify the interpretation of the analyses and graphs, we switched the signs in the Accuracy cost measures. Specifically, the Accuracy cost in TSS was calculated as the difference in accuracy between the switch and non-switch trials (or repeat and pure trials, for the mixing cost). As participants generally perform worse on switch trials, this difference (or Accuracy cost) was usually negative, with values closer to (or above) zero, signifying better performance or more flexibility, and more negative values signifying worse performance. However, for RT costs, the opposite pattern emerges: higher RTs (slower responses) are associated with switch trials, so the difference between switch and non-switch trials is usually positive, with values closer to (or below) zero indicating better performance, and more positive values indicating worse performance. Therefore, to avoid confusion, we switched the signs for Accuracy cost so that lower values indicate better performance for this measure as well. The same logic applies for the Stroop task, this time regarding the difference between the incongruent and congruent trials.

The multiple regression models for our main analysis were run in R version 4.2.2 (R Core Team 2021), using the glm() (generalised linear models) function. The best combination of factors was determined using the stepAIC() function of the MASS package in R (Venables and Ripley 2002). The direction of the search was set to "both" (default), so a forward–backward search was performed, which, at each step, decides whether to include or exclude a predictor. A predictor that was included/excluded previously can later be included/excluded, until the lowest Akaike information criterion (AIC) is reached. After identifying the best structure including the language-related predictors of interest, we then performed warranted inclusion and gradually added standardised VIQ and NVIQ scores, age, and SES to the models to determine whether their inclusion would improve the model fit (χ^2, AIC) (Cunnings 2012). Importantly, there were no issues with collinearity between the background variables and language variables in the models reported. Correlation matrices, with all background and language measures, are included in Appendix C. The plots were generated using the ggplot2 package (Wickham 2016).

3. Results

3.1. Language Profiles in the UK and Singapore

3.1.1. Differences between Groups

The mean scores and standard deviations (SDs) of each group regarding the language background measures of interest are presented in Table 3. Pairwise comparisons revealed that the groups differed significantly in: (a) L2 AoA (Mann–Whitney U = 178, $p < 0.0001$), with the Singapore group acquiring their L2 significantly earlier; (b) Balanced Usage (Mann–Whitney U = 688, $p = 0.046$), with the Singapore group being more balanced in the use of their two languages; (c) Dual-Language context use (Mann–Whitney U = 1205.5, $p = 0.011$), with the Singapore group using their languages in a dual-language context more frequently; (d) Code-Switching (Mann–Whitney U = 1327.5, $p = 0.0003$), with the Singapore group using their languages in a dense code-switching context more frequently.

Table 3. Means and standard deviations (SD) for language background measures.

	UK		Singapore			
	Mean	**SD**	**Mean**	**SD**	***MW U***	***p***
Age of L2 Acquisition (L2 AoA) (in years)	8.47	4.24	2.98	1.95	178	<0.0001
Usage$_1$	0.87	0.15	0.81	0.16		
Usage$_2$	0.13	0.15	0.19	0.16		
Proficiency$_1$	5.98	0.07	5.97	0.14		
Proficiency$_2$	4.24	2.11	4.39	1.11		
Balanced Usage (Usage$_1$ − Usage$_2$) [a]	0.75	0.30	0.62	0.32	688	0.046
Balanced Proficiency (Proficiency$_1$ − Proficiency$_2$) [b]	1.74	2.12	1.57	1.11	1038.5	0.295
Language Entropy [c]	0.50	0.44	0.64	0.39	1068	0.196
Single-Language [d]	3.86	1.22	3.89	1.01	883.5	0.756
Dual-Language [d]	2.67	1.28	3.36	0.90	1205.5	0.011
Dense Code-Switching [d]	2.31	1.14	3.19	0.86	1327.5	0.0003

[a] Difference between Usage$_1$ and Usage$_2$. Usage$_1$ indicates use of most frequent language (irrespective of whether it was L1 or L2 or two L1s, in certain cases). Usage$_2$ indicates use of second most frequent language (irrespective of whether it was L1 or L2). Usage of each language was rated on a scale from 0 to 1. Higher values in Balanced Usage indicate a greater difference between languages, thus, more unbalanced profiles. [b] Difference between Proficiency$_1$ and Proficiency$_2$. Proficency$_1$ indicates the most proficient language. Proficency$_2$ indicates the second most proficient language. Values of Proficiency in each language ranged from 0 (no proficiency) to 6. Higher values in Balanced Proficiency indicate a greater difference between languages, thus more unbalanced profiles. [c] Higher Entropy scores indicate higher language diversity and more balanced use of languages. Entropy scores are based on all of a participants' languages. [d] Based on the ACH, participants were asked to rate how frequently they use their languages in a Single-Language, a Dual-Language, and a Dense Code-Switching context. Values range from 1 (never) to 5 (always).

3.1.2. Differences between Groups

The different language variables of interest showed different correlations in the two groups. These can be found in Table 4 for the UK sample and Table 5 for the Singapore sample.

Table 4. Correlation matrix for the language background measures in the UK sample.

	1	**2**	**3**	**4**	**5**	**6**	**7**
Balanced Usage	–						
Language Entropy	−0.85 ***	–					
Balanced Proficiency	0.54 ***	−0.68 ***	–				
L2 AoA	0.11	−0.20	0.45 *	–			
Single-Language	0.03	0.02	0.02	−0.02	–		
Dual-Language	−0.34	0.34	−0.32	−0.35	0.05	–	
Code-Switching	−0.33	0.28	−0.27	−0.35	−0.10	0.80 ***	–

Asterisks indicate significance after Holm–Bonferroni correction (Holm 1979). * $p < 0.05$. *** $p < 0.001$.

Table 5. Correlation matrix for the language background measures in the SG sample.

	1	**2**	**3**	**4**	**5**	**6**	**7**
Balanced Usage	–						
Language Entropy	−0.84 ***	–					
Balanced Proficiency	0.13	−0.17	–				
L2 AoA	−0.08	−0.03	0.18	–			
Single-Language	0.03	0.14	−0.12	−0.07	–		
Dual-Language	−0.36	0.35	−0.16	−0.05	−0.24	–	
Code-Switching	−0.41	0.26	−0.17	0.02	−0.14	0.69 ***	–

Asterisks indicate significance ($p < 0.05$) after Holm–Bonferroni correction. ** $p < 0.01$. *** $p < 0.001$.

UK Sample

Balanced Usage significantly correlated with Language Entropy ($r = -0.85$, $p < 0.001$), even after applying a Holm–Bonferroni correction for multiple comparisons. This is not unexpected, as both indices reflect balance of use, but the Language Entropy index is a slightly "richer" metric, as it encompasses all languages an individual uses and not just two. Both Balanced Usage and Language Entropy significantly correlated with Balanced Proficiency ($r = 0.54$ and $r = -0.68$ respectively, ps < 0.001). Balanced Proficiency also correlated with L2 AoA ($r = -0.85$, $p = 0.032$), with more unbalanced participants indicating a later L2 AoA. Dual-Language context use significantly correlated with Dense Code-Switching ($r = 0.80$, $p < 0.001$). Finally, Dual-Language context use and Code-Switching showed high correlations with Balanced Usage, Entropy, and Balanced Proficiency, but these were not significant after applying the Holm–Bonferroni correction ($p > 0.05$). All correlations for the UK sample can be seen in Table 4.

Singapore Sample

Similar to the results for the UK sample, Balanced Usage significantly correlated with Language Entropy ($r = -0.83$, $p < 0.001$) in the Singapore sample as well, even after applying a Holm–Bonferroni correction for multiple comparisons. Dual-Language context use significantly correlated with Dense Code-Switching ($r = 0.69$, $p < 0.001$). Finally, Dual-Language context use and Code-Switching showed high correlations with Balanced Usage and Entropy and Balanced Proficiency, but these were not significant after applying the Holm–Bonferroni correction ($p > 0.05$). No other correlations were found to be significant in the Singapore sample. All correlations for the Singapore sample can be seen in Table 5.

3.2. Language Effects on Cognitive Performance

As some of our language background variables correlated significantly with each other, potentially leading to issues of multicollinearity (Dormann et al. 2013), we performed a generalised linear regression, using only specific combinations of factors, to produce feasible and interpretable models (Kuhn and Johnson 2013; Yow and Li 2015).

The combinations for the UK sample were the following:

i. Language Entropy and L2 AoA;
ii. Balanced Proficiency, Single-Language context use, Code-Switching;
iii. Balanced Proficiency, Single-Language context use, Dual-Language context use.

Language Entropy significantly correlated with Balanced Usage, and as Entropy is a richer metric, we decided to use it in our models. In addition, Entropy significantly correlated with Balanced Proficiency, so they were not entered in the same models. Correlations with Dual-Language and Code-Switching did not reach significance after performing the Holm–Bonferroni correction, but they were still high, so we chose not to include these indices in the same model. Models including combinations ii and iii did not differ significantly, so we will only report models with Code-Switching as a predictor, where relevant.

The combinations for the Singapore sample were the following:

i. Language Entropy, L2 AoA, Balanced Proficiency, Single-Language context use, Code-Switching;
ii. Language Entropy, L2 AoA, Balanced Proficiency, Single-Language context use, Dual-Language context use.

Again, as Language Entropy significantly correlated with Balanced Usage, we only used Entropy in our models. The correlations present in the UK sample did not apply in the Singapore sample, so we could add all the predictors in one model, with the exception of Code-Switching and Dual-Language context, which again correlated significantly. We therefore ran separate models with Code-Switching and Dual-Language Switching, but again, these yielded virtually identical results, so we will only report models with Code-Switching as a predictor. As suggested by Ooi et al. (2018), dual-language and dense-code switching contexts are difficult to distinguish in the linguistic environment of Singapore, as it is very likely that individuals engage in both contexts with similar frequency.

As described in Section 2.3, the best combination of predictors was identified using stepwise model selection with AIC values. Importantly, after entering the language-related predictors of interest, we performed warranted inclusion of age, SES, NVIQ, and VIQ to assess whether any language effects remained after controlling for potential confounding variables. We report the best models for each cognitive measure (lowest AIC). Only models that were significant and improved the base model (intercept only) are reported. Full model results can be found in Appendix D.

3.2.1. UK Sample

1. *Cognitive Flexibility (CF)*

Task-Set Switching (TSS)

Accuracy cost

For the TSS task, the best model for Accuracy cost included VIQ as a predictor [AIC = −104.46, $F(1,48) = 7.84$, $p = 0.007$, $R^2 = 0.14$, $R^2_{\text{adjusted}} = 0.12$], and the VIQ coefficient was significant (β = 0.003, $t = 2870$; $p = 0.007$).

2. *Working Memory (WM)*

Spatial Working Memory (SWM)

Strategy score

The best model for SWM Strategy included L2 AoA, Age, SES, NVIQ, and VIQ [AIC = 251.78, $F(5,39) = 4.03$, $p = 0.005$, $R^2 = 0.34$, $R^2_{\text{adjusted}} = 0.26$]. The effect of SES (β = −9.861, $t = -3.081$; $p = 0.004$) and VIQ (β = −0.113, $t = -2.105$; $p = 0.042$) were significant, with higher SES and VIQ leading to better Strategy. None of the language background measures had a significant effect.

However, as there was a correlation between Age and VIQ (which did not remain significant after correcting for multiple comparisons), we also ran separate models. The model including VIQ was better than the model including Age [AIC = 252.97, $F(4,40) = 4.13$, $p = 0.007$, $R^2 = 0.29$, $R^2_{\text{adjusted}} = 0.22$]. The remaining outputs were virtually the same as those for the previously reported model. Both can be found in Appendix D.

3. *Structure Learning (SL)*

SL Performance Index (PI)

The best model for SL PI improvement included Balanced Proficiency, Code-Switching, and VIQ [AIC = −68.21, $F(5,39) = 6.26$, $p = 0.001$, $R^2 = 0.29$, $R^2_{\text{adjusted}} = 0.24$]. The effects of Balanced Proficiency and VIQ were significant (β = −0.022, $t = 2.635$, $p = 0.011$; β = 0.006, $t = 3.486$, $p = 0.001$). Higher Balanced Proficiency scores indicate less balanced profiles, so participants with less balanced proficiency showed reduced learning compared to the results for the more balanced profiles. Higher VIQ also correlated with increased performance.

However, as there were correlations between Balanced Proficiency and VIQ, and Code-Switching and VIQ (which did not remain significant after correcting for multiple comparisons), we also ran separate models. The model including VIQ only was the best [AIC = −64.52, $F(1,49) = 9.98$, $p = 0.003$, $R^2 = 0.17$, $R^2_{\text{adjusted}} = 0.15$]. The effect of VIQ was significant (β = 0.005, $t = 3.158$, $p = 0.003$). The outputs for both models can be found in Appendix D.

The significant models for the UK sample are summarised in Table 6.

Table 6. Summary of the significant multiple regression models for the UK sample: background variables as predictors of performance in EF and SL tasks.

Variable	TSS Accuracy Cost	SWM Strategy	SL PI
	β	β	β
(Intercept)	−0.082	26.193 ***	−0.184
L2 AoA	–	−0.095	–
Balanced Proficiency	–	–	−0.022 *
Code-Switching	–	–	−0.021
VIQ	0.003 **	−0.081	0.006 **
NVIQ	–	−3.867	–
SES	–	−9.915 **	–
R^2	0.14	0.29	0.29
F	7.84 **	4.13 **	6.26 **

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

3.2.2. Singapore Sample

1. *Cognitive Flexibility (CF)*

Trail Making Test (TMT) B:A Ratio

The best model for TMT B:A Ratio included L2 AoA and Balanced Proficiency [AIC = 27.80, $F(2,33) = 5.90$, $p = 0.006$, $R^2 = 0.26$, $R^2_{adjusted} = 0.22$]. The effect of Balanced Proficiency was significant (β = 0.123, $t = 2.377$, $p = 0.002$), and the effect of L2 AoA approached significance (β = 0.059, $t = 2.009$, $p = 0.053$). Participants with more balanced proficiency showed a reduced B:A ratio and thus, greater flexibility. Earlier L2 AoA also resulted in increased flexibility (but only marginally significantly).

Intra-Extra Dimensional (IED) Set Shift

Total Errors

The best model for Total Errors in the IED task included Code-Switching, Single-Language context use, and Age [AIC = 349.23, $F(3,32) = 3.632$, $p = 0.023$, $R^2 = 0.25$, $R^2_{adjusted} = 0.18$]. Only the effect of Code-Switching was significant (β = 15.705, $t = 2.753$, $p = 0.010$). Importantly, our model suggests that individuals who engage in more Code-Switching make more errors in the IED task (see Figure 9).

After visual observation of the effect of Code-Switching on total errors, we thought that the effect may be driven by extreme outliers. We therefore conducted the analysis again, after removing the outliers (i.e., scores that fell into the 1.5 inter-quartile ranges (IQR) below/above the 1st/3rd quartile). The effect of Code-Switching disappeared. The best model [AIC =177.47, $F(1,29) = 6.024$, $p = 0.020$, $R^2 = 0.17$, $R^2_{adjusted} = 0.14$], after removing the outliers, included VIQ only (β = 0.224, $t = 2.454$, $p = 0.02$).

Extra-Dimensional Shift Errors

The best model for Extra-Dimensional Shift Errors included Entropy, Age, and NVIQ [AIC = 238.48, $F(3,32) = 8.618$, $p = 0.0002$, $R^2 = 0.45$, $R^2_{adjusted} = 0.40$]. All coefficients were found significant: Entropy (β = 11.366, $t = 4.090$, $p = 0.0002$), Age (β = 1.362, $t = 2.823$, $p = 0.008$), and NVIQ (β = −11.077, $t = -2.432$, $p = 0.021$). Higher Entropy scores led to more errors, as did higher age. Higher NVIQ led to reduced number of Extra-Dimensional Shift Errors (see Figure 10).

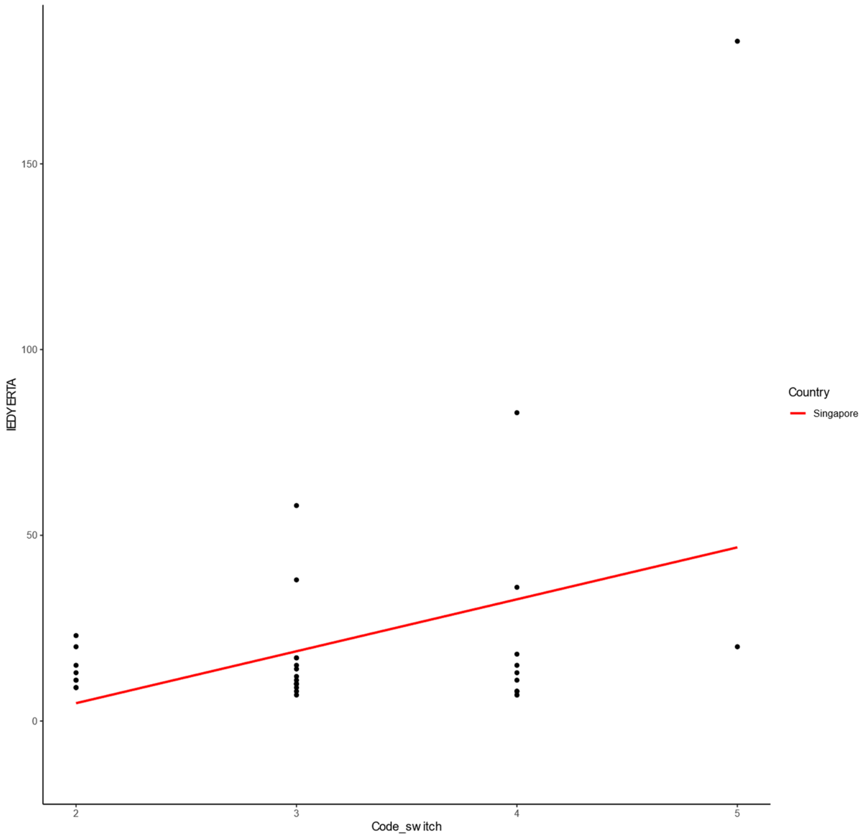

Figure 9. Effect of Code-Switching on number of Total Errors in the IED task.

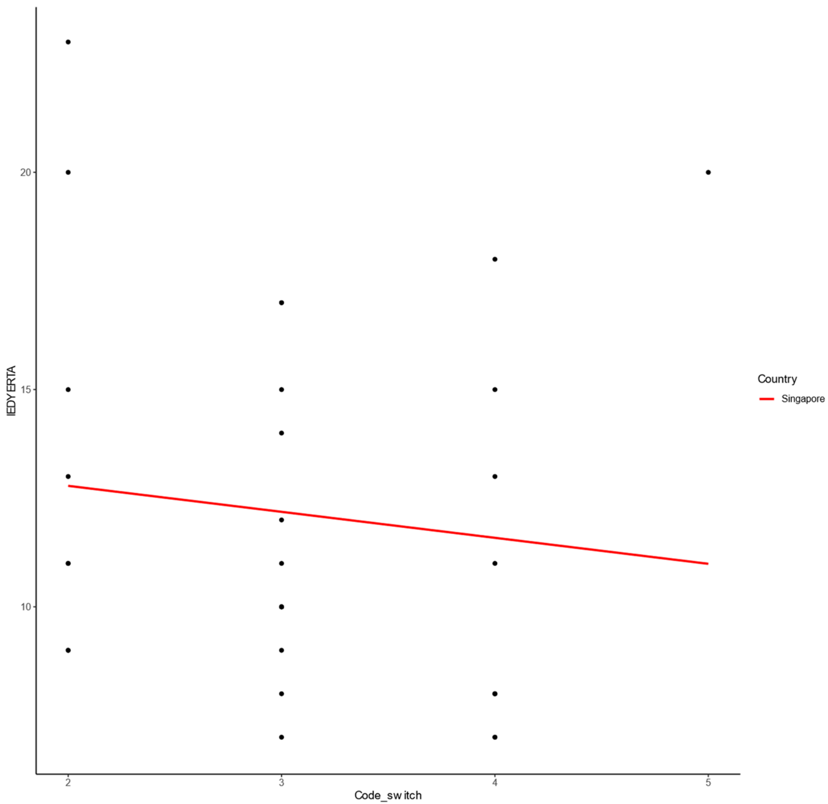

Figure 10. The effect of Code-Switching on number of total errors in the IED task (after removing outliers) was not significant.

2. *Inhibition*

Stroop Task

RT cost

For the Stroop RT cost, the best model included Balanced Proficiency, Single-Language context use, and Code-Switching [AIC = 428.86, $F(3,32) = 6.748$, $p = 0.001$, $R^2 = 0.39$, $R^2_{adjusted} = 0.33$]. The effect of Balanced Proficiency was significant ($\beta = 29.57$, $t = 2.192$, $p = 0.036$), with an increase in Stroop RT cost for participants with less balanced proficiency. The effects of Code-Switching ($\beta = 43.94$, $t = 2.505$, $p = 0.018$) and Single-Language context ($\beta = 56.83$, $t = 3.850$, $p = 0.0005$) were both significant. Interestingly, despite the two types of contexts negatively correlating with each other, they both led to increased RT costs in the Stroop task. In other words, both engaging in dense code-switching and using language(s) in a single-language mode seemed to negatively impact inhibitory control, as measured by the Stroop task, yet the effect of single-language use was stronger (see Figure 11).

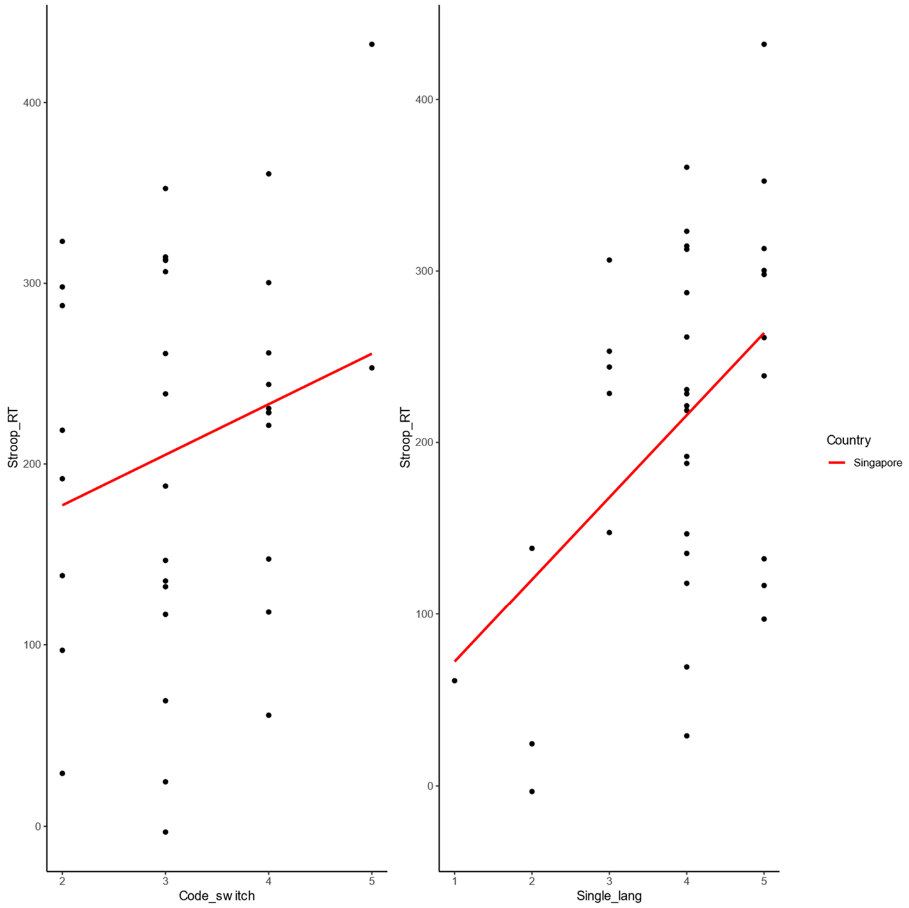

Figure 11. Effects of Code-Switching and Single-Language context use on Stroop RT cost.

3. *Working Memory (WM)*

Spatial Working Memory (SWM)

Total Errors

The best model for SWM total errors included Entropy and SES [AIC = 318.22, $F(2,33) = 3.21$, $p = 0.053$, $R^2 = 0.16$, $R^2_{adjusted} = 0.11$]. In fact, the model only approached significance ($p = 0.053$). The coefficients were not significant, but Entropy negatively correlated with the number of errors; that is, there was a tendency for participants with more linguistically diverse profiles to make fewer errors, but the effect did not reach significance ($\beta = -13.456$, $t = -1.638$, $p = 0.11$).

Strategy score

The best model for SWM Strategy also included Entropy and SES [AIC = 205.07, $F(2,33) = 6.005$, $p = 0.006$, $R^2 = 0.27$, $R^2_{adjusted} = 0.22$]. The model was significant, as was the coefficient for SES (β = 9.473, $t = 2.461$, $p = 0.019$), with a higher SES correlating with a higher Strategy score, that is, worse strategy. The effect of Entropy observed on total errors was clearer here, as higher Entropy scores almost significantly correlated with better SWM Strategy (β = −3.426, $t = -2.007$, $p = 0.053$).

Verbal Working Memory (VWM)

Backward Digit Span (BDS)

For the BDS task, the best model included Code-Switching, SES, and NVIQ [AIC = 181.81, $F(3,32) = 4.56$, $p = 0.009$, $R^2 = 0.30$, $R^2_{adjusted} = 0.23$]. Only the intercept for NVIQ was significant (β = 6.677, $t = 3.233$, $p = 0.003$). NVIQ significantly predicted length of backward digit span.

Structure Learning (SL)

SL Performance Index (PI)

The best model for SL PI in the Singapore sample included Entropy and SES [AIC = −51.40, $F(2,33) = 12.75$, $p < 0.0001$, $R^2 = 0.44$, $R^2_{adjusted} = 0.40$]. Both factors had a significant and positive effect on SL PI improvement (Entropy: β = 0.179, $t = 3.693$, $p = 0.0008$; SES: β = 0.437, $t = 3.999$, $p = 0.0003$). This means that participants with more diverse linguistic profiles showed greater improvement in the SL task.

SL Strategy ICD

The model for SL Strategy ICD including Entropy approached significance [AIC = −34.66, $F(1,34) = 3.589$, $p = 0.067$, $R^2 = 0.10$, $R^2_{adjusted} = 0.07$]. Higher Entropy scores correlated with higher Strategy ICD scores, reflecting a maximisation strategy (Entropy: β = 0.116, $t = 1.894$, $p = 0.0667$).

The significant models for the Singapore sample are summarised in Table 7.

Table 7. Summary of the significant multiple regression models for the Singapore sample: background variables as predictors of performance in EF and SL tasks.

Variable	TMT B:A Ratio	IED Total Errors	IED ED Errors	Stroop RT Cost	SWM Strategy	BDS	SL PI
	β	*β*	*β*	*β*	*β*	*β*	*β*
(Intercept)	0.86 ***	0.39	−24.93 *	−35.41	11.84 ***	10.66 ***	−0.23 ***
L2 AoA	0.06	–	–	–	–	–	–
Balanced Proficiency	0.12 *	–	–	29.57 *	–	–	–
Language Entropy	–	–	11.37 ***	–	−3.43	–	0.18 ***
Single-Language Context	–	–	–	56.83 ***	–	–	–
Code-Switching	–	–	–	43.94 *	–	−0.94	–
VIQ	–	0.22 **	–	–	–	–	–
NVIQ	–	–	−11.08 *	–	–	6.68 **	–
SES	–	–	–	–	9.47 *	−4.50	0.44 ***
Age	–	–	1.36 **	–	–	–	–
R^2	0.26	0.17	0.45	0.39	0.27	0.20	0.44
F	5.90 **	6.02 *	8.62 **	6.75 **	6.01 **	4.56 **	12.75 ***

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

4. Discussion

The aim of this study was twofold. First, after operationalising linguistic experience as a multifactorial continuum, we assessed differences in the linguistic profiles of participants from the UK and Singapore. Second, we explored whether these multilingualism experience factors correlate with cognitive performance in several tasks, and if the relationship varies in the different contexts.

4.1. Linguistic Profiles

In terms of linguistic profiles, the UK and Singapore participants differed in three important aspects. The Singapore group acquired their L2[7] at 2.98 years of age on average, significantly earlier than did the UK group, which, on average, acquired an L2 at 8.46 years. Participants from Singapore also exhibited a more balanced language usage, measured as the difference between an individual's first and second most used languages (weighted against each other). Higher scores on the Balanced Usage metric reflect a greater difference between a participant's first and second most used languages, hence less balance, and the average score for UK participants was 0.74 vs. 0.62 for Singapore participants. Finally, the participants from Singapore reported a higher engagement in dual-language use and dense code-switching contexts compared to the UK participants.

The above patterns align with the sociolinguistic realities of the UK and Singapore and the expectations that come with them. The environment of Singapore is highly multilingual, with the various ethnic and linguistic groups interacting frequently.[8] Due to Singapore's bilingual education system, people are at least bilingual, but often trilingual or even multilingual. Studies on the linguistic environment of Singapore suggest that languages are acquired very early (if not from birth), partly through immersion and partly through instruction (Leimgruber 2013; Ooi et al. 2018; Siemund et al. 2014). This is in line with our findings, as most participants reported acquiring two languages from birth, while the rest report acquiring a second language before the age of 8. Therefore, the significantly earlier L2 AoA for the Singapore group is expected. Notably, the UK group included participants who were not UK nationals but who resided in the UK to attend university. These individuals learned English at a young age, usually at the beginning of formal education. If our sample was confined to UK nationals, the difference in L2 AoA would likely be more pronounced.

The difference in Balanced Usage was also expected. In recent years, English has become more prominent in Singapore and is endorsed as the main working, educational, administrative, and governmental language of the country. Still, three additional languages are listed as official in the Constitution, providing recognition to the three major ethnic groups (Chinese, Malays, and Indians) (Leimgruber 2013). It is therefore not surprising that our Singapore sample exhibited more balanced language use (i.e., greater use of a second language) than our UK sample. Despite our inclusion of non-UK nationals, who were usually highly proficient English-as-an-L2 speakers, the monolingual environment of the UK does not provide many occasions in which these speakers could use their L1. These were restricted to communications with family and friends from home (as many stated in the questionnaire). Interestingly, while the Singapore sample also had higher Entropy scores than did the UK sample, the difference between the two groups did not reach significance. Entropy was calculated based on the active use of all an individual's languages (not just the two most used, as in the Balanced Usage measure). This suggests that despite exposure to diverse languages, Singaporeans might primarily engage in bilingual rather than multilingual active language use. This is line in with the data from recent literature (Cavallaro and Chin 2014; Siemund et al. 2014) and underscores the need for more refined instruments for capturing both active use and passive exposure.

Greater engagement in dual-language and dense code-switching contexts among Singapore participants is consistent with the linguistic diversity and code-switching prevalence in the country (Ong and Zhang 2010; Xie and Cavallaro 2016). In addition, the use of Singlish is widespread. While our questionnaire did not include a question explicitly regarding Singlish, Singaporeans could consider speaking Singlish as speaking one language, with frequent code-switching. This could further explain the significant difference between Singapore and the UK participants. In any case, opportunities for code-switching for bi-/multilinguals residing in the UK are far more limited.

4.2. Relationship between Language Experience and Cognitive Control in Different Contexts

Our second aim was to assess the relationship between different multilingualism experience factors (L2 AoA, Balanced Proficiency, Balanced Usage, Language Entropy, Context of Use) and cognitive performance (on EF and SL tasks), taking the sociolinguistic context (UK vs. Singapore) into account.

In the UK sample, multilingualism factors had a limited impact on cognitive control, with significance observed only in the Structure Learning (SL) task. Specifically, Balanced Proficiency significantly predicted SL Performance Index (PI) improvement. This means that participants with more balanced proficiency had a greater capacity for learning the stochastic sequences presented in the SL task. Prior literature supports the link between bilingual proficiency and statistical learning (Bartolotti et al. 2011; Onnis et al. 2018; Potter et al. 2017). The interesting element of our SL task is that it involves learning stochastic, that is, non-deterministic, sequences of stimuli, without explicit feedback. Participants were, in some sense, required to impose order on chaos and to use subtle cues in their input to form rules. This process is very similar to natural language learning, so it is not surprising that people who have had extensive experience in mastering two languages perform better in such a task. The SL task was the only one to show sensitivity to multiple-language experience in the UK sample, indicating its potential to unveil the so-far undeciphered relationship between multilingualism and cognition. This task has been shown to correlate with the executive and memory-related networks of the brain (Giorgio et al. 2018; Karlaftis et al. 2018), and our analysis shows that it may also relate to linguistic experience. It is therefore a very interesting avenue to explore in future research.[9]

The image emerging in the Singapore sample is quite different, as more tasks were influenced by multilingualism-related factors. Out of the EF tasks, two measuring CF, the Trail Making Test (TMT) and the Intra-Extra Dimensional (IED) Shift task, revealed significant effects, as did the Stroop task, a task of inhibition, and the SWM task. Significant effects were found in the SL task as well.

In the TMT, a significant effect of Balanced Proficiency emerged, revealing that individuals with greater balance between their two most proficient languages exhibited a reduced B:A ratio, indicative of enhanced flexibility. Literature examining the performance of bilinguals on the TMT is limited, but studies usually do not report an effect, owing to the fact that the switching and visual attention involved in the TMT are arguably not trained through multiple language control (Filippi et al. 2022; Stasenko et al. 2017; Torres et al. 2022). Nevertheless, there are a few studies reporting benefits for bilinguals in different age groups (Bialystok 2010; Estanga et al. 2017), but these usually evaluate the performance in Trail A and Trail B separately, while the B:A ratio is in fact a more accurate measure of CF (Kopp 2011). Notably, our study revealed a significant effect on B:A ratio in the Singapore sample, warranting further exploration to determine generalisability across bi-/multilingual populations. An intriguing hypothesis would be that people in Singapore are often literate in more than one language, and that these languages also have distinct writing systems, thus, these participants may have benefitted from exposure to multiple scripts, positively impacting the visual attention and switching skills involved in the TMT (Yin et al. 2022). Research on the effects of biliteracy and different scripts (alphabetic vs. logographic) on EF is, however, extremely scarce, so this is only speculation.

The effects noted from the results of the IED task lean in a different direction. In the two measures we evaluated, total number of errors and extra-dimensional errors (the latter considered a more accurate measure of CF), increased Code-Switching and higher Entropy (i.e., frequent use of more than two languages) had a negative impact, leading to more errors. Such an effect is not unprecedented. A number of studies suggest that higher contextual or everyday language-switching predicts higher switch costs (Jylkkä et al. 2017, 2021), while other research that found reduced switch costs for bilinguals failed to link them to language-switching frequency or fluency (Woumans et al. 2019). It could be argued that frequent switching or increased exposure to multiple languages are not sufficient, in their own right, to lead to cognitive benefits. In fact, there is evidence that the type of switching (mandatory vs. voluntary) poses different cognitive demands (Jevtović et al. 2020), so freely code-switching between languages, without restrictions (e.g., when both interlocutors speak both languages), may not require significant cognitive effort, thus not leading to cognitive benefits. Nevertheless, inspection of the data revealed a few extreme outliers, so we reran our analysis without them. This made the effect of Code-Switching disappear. We therefore argue that frequent code-switching is not necessarily detrimental, but that this may be the case for a few individuals. This calls for a more refined evaluation of code-switching, as it may not in fact be a uniform practice.

The model for the Stroop RT cost in the Singapore sample was significant, with Balanced Proficiency, Single-Language use, and Code-Switching all reaching levels of significance. Regarding the balanced proficiency effect, greater balance between languages, equivalent to greater proficiency in L2, led to a reduced Stroop effect. This finding is interesting, as literature regarding the effects of L2 proficiency on inhibition is highly controversial. In a study on elderly bilinguals, no effect of L2 proficiency was found, either on a numerical or on a verbal Stroop task (Antón et al. 2016). Neuroimaging evidence suggests that higher L2 proficiency is associated with more automatic and efficient inhibitory control on a Simon task (Jia 2022). A positive effect on ERPs has also been reported in a non-verbal Stroop task, but this effect did not stand up when behavioural measures were considered (Jiao et al. 2019). Crucially, the two studies most similar to ours, employing comparable methodology and focusing on young adults in Singapore, show dissimilar results. In the study using a verbal Stroop task, the same as that used in our study, the effects of L2 AoA and Balanced Usage were found to be significant. The effect of Balanced Proficiency was in the same direction as that noted in our study but did not reach significance (Yow and Li 2015). On the other hand, the study using a numerical Stroop task reported no effect of Balanced Proficiency whatsoever (Li et al. 2021). It is difficult to extract a common thread from this plethora of results, but a trend that seems to emerge is that Balanced Proficiency may be more relevant for versions of the Stroop task with a verbal or linguistic component, rather than for the numerical versions. While not entirely corroborated by the reported literature, this hypothesis is worth exploring further, as it is highly likely that the effects may be, to some extent, task-specific. It is interesting, and curious at the same time, that our results do not fully align with those of other studies assessing very similar samples. We should, however, bear in mind that all our participants reported quite high proficiency in both their languages; hence, there was not much variation in their profiles. In addition, the studies by Li et al. (2021) and Yow and Li (2015) used English–Mandarin bilinguals only, whereas we did not apply such restrictions to our sample. This allows for greater variation in the combinations of languages, which may be closer or further from each other, potentially influencing the magnitude of the effects.

The Single-Language and Dense Code-Switching contexts both led to detrimental effects in the Stroop task; that is, using one's languages in a single-language mode, as well as frequent code-switching, were both correlated with a greater Stroop effect. Both effects are, in principle, plausible. Operating in a single-language context means that the need to inhibit one's other language is not regular, so inhibitory control is not exercised frequently. Similarly, unrestricted code-switching, as mentioned, may not pose great cognitive demands. Nevertheless, the fact that the effect of both contexts reached significance is interesting and poses certain challenges for the adaptive control hypothesis (ACH) (Green and Abutalebi 2013). Conceptually, these two contexts are opposing, as frequently operating in single-language mode should equate to infrequent code-switching. While the two measures negatively correlated in our sample, the correlation did not reach significance. In addition, careful observation of our data showed that the individuals who drove the negative effect of code-switching also drove the negative effect of single-language context use. This hints towards two possible explanations: (1) these two contexts are not as easily distinguishable as the ACH posits; (2) our operationalisation of the different contexts was not entirely successful, as participants may not have understood their difference, or may have been unable to place themselves in the most appropriate context, perhaps influenced by their thinking thinking of Singlish. A recent study conducted with young English–Mandarin bilinguals in Singapore reports results very similar to ours, as Single-Language and Dual-Language contexts were not found to be diametrically opposed, and Dual-Language and Dense-Code-Switching contexts correlated highly (Lai and O'Brien 2020). Evidently, the contexts proposed by the ACH may not be completely distinguishable, especially in the case of multilingual populations. The authors posit that three clearly distinct interactional contexts may lack ecological validity in multilingual populations due to the fluidity of communicative environments and the general overlap of language experiences. The ACH may provide a more nuanced approach than a simple comparison of monolinguals and bilinguals, but a categorical operationalisation of linguistic contexts can still obscure intricacies that arise in the natural linguistic ecology of bilinguals (and multilinguals).

Moving to SWM Strategy, the best significant model included SES and Entropy, with Entropy nearly reaching significance ($p = 0.053$). While there is abundant literature reporting a positive relationship between bilingualism and SWM in multiple age groups (Antón et al. 2019; Blom et al. 2014; Kerrigan et al. 2017; Luo et al. 2013; Morales et al. 2013; Sullivan et al. 2016), the Entropy measure used here more broadly encompasses linguistic diversity, taking all of a participant's languages into account. A study with young multilingual adults from South Africa found that the addition of a third language had detrimental effects on an extensive set of WM tasks (Cockcroft 2022). This finding is not in line with our results, but it highlights the importance of taking sociolinguistic context into account when examining the relationship between language experience and cognitive processing. Differences in the linguistic realities of Singapore and South Africa may generate different cognitive demands, hence training different cognitive aspects. Importantly, the effect of Entropy in our sample only approached significance, and the Singapore sample was relatively small.

Finally, in the Singapore group, Entropy reached significance in the SL PI model and approached significance in the SL Strategy ICD model ($p = 0.067$). We have already discussed the importance of the SL task, and our results here further demonstrate its potential in uncovering the relationship between language and cognitive control. Interestingly, in the Singapore sample, the model including Entropy rather than Balanced Proficiency reached significance, unlike the results for the UK sample. This is noteworthy because Entropy encompasses the use of all a participant's languages, while Balanced Proficiency accounts for proficiency in only the two most proficient languages. The role of sociolinguistic context, therefore, once again becomes apparent, as different multilingual experience factors seem to be pertinent in different contexts. We do not overestimate the effect of Entropy on SL Strategy, as it only approached significance. However, we must highlight that different strategies in the SL task have been associated with the activation of different areas of the brain (Giorgio et al. 2018; Karlaftis et al. 2018), so it is intriguing to consider that linguistic diversity could correlate with distinct brain activation patterns.

5. Conclusions

This study resonates with previous work suggesting that cognitive control is influenced by an interplay of multiple language experience factors (Ooi et al. 2018). We add that this interplay varies as a function of sociolinguistic context. The effects of multilingual experience on cognitive performance were much more evident in our Singapore sample. This underscores the significance of diversifying language research, particularly in the Global South, to avoid a skewed representation based primarily on Western samples (Bak and Alladi 2016; Makoni et al. 2022). In our study, the effects of multilingual experiences on cognitive performance were very limited in the UK sample, but were quite dispersed in the Singapore sample. We therefore support the suggestion of Ooi et al. (2018) that we should be talking of "bilingualisms" (or even "multilingualisms") rather than simply bilingualism. Further investigations in Singapore, an exceptionally rich multilingual environment, may reveal relationships between language and cognition that have so far been masked.

Another important insight is the great potential of the Structure Learning (SL) task for research on the potential influence of multilingualism on cognitive functioning. Importantly, this was the only task in which language-related effects surfaced for both the UK and Singapore samples. As illustrated, this task relates to both executive function (EF) and language learning. The prospect of SL acting as an intermediary link should be further investigated.

Of course, our study includes certain limitations. Firstly, our UK and Singapore samples were not restricted in terms of their linguistic profiles. This means that our participants varied in the combinations of languages they spoke, which led to less homogeneous samples. While this increased ecological validity, future studies targeting specific linguistic profiles could eliminate the confounding effects of typological divergence in the participants' languages.

Secondly, the cross-sectional design, though mitigated by controlling for confounding variables, such as SES, age, and intelligence, remains subject to inherent limitations and restricts causal inferences (Laine and Lehtonen 2018). Finally, between-group differences in background variables (e.g., SES, age, intelligence) restricted our ability to perform direct comparisons between Singapore and the UK. It would be interesting to directly compare participants from the UK and Singapore, but there were too many parameters which were out of our control to allow for the performance of this comparison.

Finally, challenges with the operationalisation of the adaptive control hypothesis (ACH) and the measurement of multilingual experience highlight areas for refinement, such as the use of more detailed and continuous measures like the Contextual Linguistic Profile Questionnaire (CLiP-Q), which takes both individual and contextual diversity into account (Wigdorowitz et al. 2020).

Author Contributions: Conceptualisation, C.V., V.L. and H.H.; methodology, C.V.; software, C.V.; formal analysis, C.V., V.L. and H.H.; investigation, C.V.; resources, C.V., V.L. and H.H.; writing—original draft preparation, C.V.; writing—review and editing, C.V., V.L. and H.H; visualisation, C.V.; supervision, V.L. and H.H.; project administration, C.V.; funding acquisition, V.L. and H.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was conducted as part of C.V.'s PhD. C.V. was funded by the Economic and Social Research Council (ESRC) (ES/P000738/1), the Onassis Foundation (F ZQ 006-1/2020-2021), and the A.G. Leventis Foundation. V.L. is supported by the Ministry of Education, Singapore, under its Science of Learning grant (MOESOL2021-0001) and by a Social Science & Humanities Research Fellowship (MOE2020-SSHR-008). Part of this research was conducted in the context of the Centre for Lifelong Learning and Individualised Cognition (CLIC) project. CLIC is supported by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of Nanyang Technological University, Singapore (NTU) (IRB-2020-02-001-01, 25 April 2020) and the Cambridge Psychology Research Ethics Committee of the University of Cambridge, UK (PRE.2017.088, 5 May 2020).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on reasonable request from corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A Demographics Questionnaire

1. When is your birthday (Enter x if you would prefer not to say)?

   Enter in the form of: dd-mm-yyyy

2. What is your gender?

   Male

Female
Other
Prefer not to say

3. Are you now married, widowed, divorced, separated or never married?

Married
Widowed
Divorced
Separated
Never Married
Prefer not to say

4. Have you ever been diagnosed with any neurological or psychiatric disorders or suffered from brain injury?

Yes: Please give details (optional)
No
Prefer not to say

5. Are you colour-blind?

Yes
No
Prefer not to say

6. Describe your handedness:

Right-handed
Left-handed
Ambidextrous
Prefer not to say

7. Choose one or more ethnicities that you consider yourself to be:

White or White American/British
Black or African American/British
Central/South American
Asian: Indian
Asian: Pakistani
Asian: Chinese
Asian: Malay
Asian: Other
Other (please specify)
Prefer not to say

8. In which country do you currently reside?

 Country
Pick a country

9. In which country were you born?

 Country
Pick a country

10. In which country did you spend the majority of your youth?

 Country
Pick a country

11. Have you ever lived in a country other than the one in which you were born?

    Yes
    No
    Prefer not to say

12. What is your current Residence Status

    Permanent Resident
    Non-permanent resident?
    Citizen
    Prefer not to say

13. What is your Postcode/District? (Enter x if you would prefer not to say)

    (Enter your postcode with no spaces; e.g., cb28pq)

14. What is the highest level of education you have completed?

    Primary school
    Secondary school
    University Degree (BA or equivalent)
    Master's degree
    Doctoral degree
    Professional degree (JD, MD)
    Prefer not to say

15. Which statement best describes your current employment status?

    Working (paid employee)
    Working (self-employed)
    Student
    Homemaker
    Not working (temporary layoff from a job)
    Not working (looking for work)
    Not working (retired)
    Not working (unable to work)
    Not working (other; please specify)
    Prefer not to say

16. Indicate your best guess as to your household's earnings in the last year (before tax)? *(*in Singapore $ and UK £ - depending on the sample)*

    Less than 10,000
    10,000 to 19,999
    20,000 to 29,999
    30,000 to 39,999
    40,000 to 49,999
    50,000 to 59,999
    60,000 to 69,999
    70,000 to 79,999
    80,000 to 89,999
    90,000 to 99,999
    100,000 to 149,999
    150,000 or more
    I don't know
    Prefer not to say

17. What type of dwelling does your family live in? *(*Singapore Question)*

1- and 2-room flat
3-room flat
4-room flat
5-room and executive flat
Condominium or other Apartments
Landed property
Other
Prefer not to say

18. Is your family dwelling privately owned or rented? *(*UK question)*

Privately owned
Rented
Prefer not to say

19. Follow-up: What type of dwelling does your family reside in?

Detached
Semi-detached
Flat
Prefer not to say

20. How many children are there in your household?

1
2
3
4
5 or more
Prefer not to say

21. Which members compose your household? (select all that apply)

Parent(s)
Child(ren)
Grandparent(s)
Domestic Helper
Prefer not to say

22. What is the highest level of school your father has completed or the highest degree they have received?

Primary school
Secondary school
University Degree (BA or equivalent)
Master's degree
Doctoral degree
Professional degree (JD, MD)
Prefer not to say

23. What is the highest level of school your mother has completed or the highest degree they have received?

Primary school
Secondary school
University Degree (BA or equivalent)
Master's degree
Doctoral degree
Professional degree (JD, MD)
Prefer not to say

24. Which statement best describes your father's current employment status?

Working (paid employee)
Working (self-employed)
Homemaker
Not working (temporary layoff from a job)
Not working (looking for job)
Not working (retired)
Not working (unable to work)
Not working (other; please specify)
Prefer not to say

25. Which statement best describes your mother's current employment status?

Working (paid employee)
Working (self-employed)
Homemaker
Not working (temporary layoff from a job)
Not working (looking for job)
Not working (retired)
Not working (unable to work)
Not working (other; please specify)
Prefer not to say

26. Do you play video games (of any sort)?

Yes
No
Prefer not to say

Appendix B Language Questionnaire & CEFR Grid

Q1. Participant ID

Q2. Please answer the following questions for all the languages/dialects you know (native and non-native)

- Which language(s)
- Age of Acquisition (For languages you speak from birth, put 0).
- Context of Acquisition (e.g., home, environment, school).
- Hours of current Usage (in a day) (0–18 h).
- Percentage of current Usage (in a day). Put 0% if you do not use this particular language (your answers should add up to 100%).
- Context of Usage (e.g., home, education, community, work, language school, etc.).

Q3. How often are the following statements true for your everyday use of language?

- I only/mainly use one of the languages I know.
  Never ☐ Rarely ☐ Sometimes ☐ Most of the time ☐ Always ☐
- I use the languages I know in distinct contexts (e.g., Language 1 at home, Language 2 at school/work, etc.).
  Never ☐ Rarely ☐ Sometimes ☐ Most of the time ☐ Always ☐
- I switch between the languages I know even within the same context (e.g., home or school), but with different interlocutors.
  Never ☐ Rarely ☐ Sometimes ☐ Most of the time ☐ Always ☐
- I switch between the languages I know even within the same conversation/interaction (with the same interlocutor).
  Never ☐ Rarely ☐ Sometimes ☐ Most of the time ☐ Always ☐

Q4. Select the number of non-native languages/dialects that you know.

Q5. You will now be presented with some statements describing levels of proficiency in a language. Please write each of your additional languages/dialects under the box that most accurately describes your competence in that language/dialect.

For additional (non-native) languages/ dialects: Please write each language/dialect under the box that most accurately describes your competence in that language/dialect.

		1	2	3	4	5	6
UNDERSTANDING	Listening	I can recognise familiar words and very basic phrases concerning myself, my family and immediate concrete surroundings when people speak slowly and clearly.	I can understand phrases and the highest frequency vocabulary related to areas of most immediate personal relevance (e.g. very basic personal and family information, shopping, local area, employment). I can catch the main point in short, clear, simple messages and announcements.	I can understand the main points of clear standard speech on familiar matters regularly encountered in work, school, leisure, etc. I can understand the main point of many radio or TV programmes on current affairs or topics of personal or professional interest when the delivery is relatively slow and clear.	I can understand extended speech and lectures and follow even complex lines of argument provided the topic is reasonably familiar. I can understand most TV news and current affairs programmes. I can understand the majority of films in standard dialect.	I can understand extended speech even when it is not clearly structured and when relationships are only implied and not signalled explicitly. I can understand television programmes and films without too much effort.	I have no difficulty in understanding any kind of spoken language, whether live or broadcast, even when delivered at fast native speed, provided I have some time to get familiar with the accent.
	Reading	I can understand familiar names, words and very simple sentences, for example on notices and posters or in catalogues.	I can read very short, simple texts. I can find specific, predictable information in simple everyday material such as advertisements, prospectuses, menus and timetables and I can understand short simple personal letters.	I can understand texts that consist mainly of high frequency everyday or job-related language. I can understand the description of events, feelings and wishes in personal letters.	I can read articles and reports concerned with contemporary problems in which the writers adopt particular attitudes or viewpoints. I can understand contemporary literary prose.	I can understand long and complex factual and literary texts, appreciating distinctions of style. I can understand specialised articles and longer technical instructions, even when they do not relate to my field.	I can read with ease virtually all forms of the written language, including abstract, structurally or linguistically complex texts such as manuals, specialised articles and literary works.

		1	2	3	4	5	6
SPEAKING	Spoken Interaction	I can interact in a simple way provided the other person is prepared to repeat or rephrase things at a slower rate of speech and help me formulate what I'm trying to say. I can ask and answer simple questions in areas of immediate need or on very familiar topics.	I can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar topics and activities. I can handle very short social exchanges, even though I can't usually understand enough to keep the conversation going myself	I can deal with most situations likely to arise whilst travelling in an area where the language is spoken. I can enter unprepared into conversation on topics that are familiar, of personal interest or pertinent to everyday life (e.g. family, hobbies, work, travel and current events).	I can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible. I can take an active part in discussion in familiar contexts, accounting for and sustaining my views.	I can express myself fluently and spontaneously without much obvious searching for expressions. I can use language flexibly and effectively for social and professional purposes. I can formulate ideas and opinions with precision and relate my contribution skilfully to those of other speakers.	I can take part effortlessly in any conversation or discussion and have a good familiarity with idiomatic expressions and colloquialisms. I can express myself fluently and convey finer shades of meaning precisely. If I do have a problem I can backtrack and restructure around the difficulty so smoothly that other people are hardly aware of it.
	Spoken Production	I can use simple phrases and sentences to describe where I live and people I know.	I can use a series of phrases and sentences to describe in simple terms my family and other people, living conditions, my educational background and my present or most recent job.	I can connect phrases in a simple way in order to describe experiences and events, my dreams, hopes and ambitions. I can briefly give reasons and explanations for opinions and plans. I can narrate a story or relate the plot of a book or film and describe my reactions.	I can present clear, detailed descriptions on a wide range of subjects related to my field of interest. I can explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.	I can present clear, detailed descriptions of complex subjects integrating sub-themes, developing particular points and rounding off with an appropriate conclusion.	I can present a clear, smoothly-flowing description or argument in a style appropriate to the context and with an effective logical structure which helps the recipient to notice and remember significant points.

		1	2	3	4	5	6
W R I T I N G	**Writing**	I can write a short, simple postcard, for example sending holiday greetings. I can fill in forms with personal details, for example entering my name, nationality and address on a hotel registration form.	I can write short, simple notes and messages relating to matters in areas of immediate needs. I can write a very simple personal letter, for example thanking someone for something.	I can write simple connected text on topics which are familiar or of personal interest. I can write personal letters describing experiences and impressions.	I can write clear, detailed text on a wide range of subjects related to my interests. I can write an essay or report, passing on information or giving reasons in support of or against a particular point of view. I can write letters highlighting the personal significance of events and experiences.	I can express myself in clear, well- structured text, expressing points of view at some length. I can write about complex subjects in a letter, an essay or a report, underlining what I consider to be the salient issues. I can select style appropriate to the reader in mind.	I can write clear, smoothly-flowing text in an appropriate style. I can write complex letters, reports or articles which present a case with an effective logical structure which helps the recipient to notice and remember significant points. I can write summaries and reviews of professional or literary works.

Appendix C Correlation Matrices

Appendix C.1 Correlation Matrices of Background Measures

UK Sample

	1	2	3	4
Age	–			
SES	0.06	–		
NVIQ	−0.009	0.12	–	
VIQ	−0.34 *	−0.01	−0.03	–

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

After applying a Holm–Bonferroni correction for multiple comparisons, the correlation between Age and VIQ was no longer significant. Still, for extra caution, we ran separate models as well.

Singapore Sample

	1	2	3	4
Age	–			
SES	0.11	–		
NVIQ	−0.05	0.30	–	
VIQ	0.08	0.08	0.11	–

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Appendix C.2 Correlation Matrices of Background and Language Measures

UK Sample

	Balanced Usage	
Balanced Usage	–	
Age	−0.23	
SES	0.02	
NVIQ	0.05	
VIQ	0.38 **	p = 0.006

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

The correlation between Balanced Usage and VIQ remained significant even after applying a Holm–Bonferroni correction. In any case, none of the significant models included a combination of the two variables, so it did not affect our analysis.

	Language Entropy	
Language Entropy	–	
Age	0.44 **	$p = 0.001$
SES	−0.10	
NVIQ	-0.10	
VIQ	−0.39 **	$p = 0.005$

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

The correlations remained significant even after applying a Holm–Bonferroni correction for multiple comparisons. We therefore ran separate models for these measures and reported the best one, where relevant.

	Balanced Proficiency	
Balanced Proficiency	–	
Age	−0.23	
SES	0.02	
NVIQ	0.05	
VIQ	0.38 **	$p = 0.006$

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

The correlation between Age and Balanced Proficiency remained significant even after applying a Holm–Bonferroni correction. In any case, none of the significant models included a combination of the two variables, so it did not affect our analysis.

The correlation between VIQ and Balanced Proficiency was not significant after applying a Holm–Bonferroni correction. Still, for extra caution, we ran separate models, in addition to the combined model for these variables.

	L2 AoA
L2 AoA	–
Age	−0.03
SES	−0.006
NVIQ	0.19
VIQ	0.08

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

	Single-Language
Single-Language	–
Age	0.20
SES	0.20
NVIQ	−0.15
VIQ	0.23

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

	Dual-Language
Dual-Language	–
Age	0.14
SES	0.06
NVIQ	−0.18
VIQ	−0.17

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

	Code-Switching	
Code-Switching	–	
Age	0.04	
SES	0.02	
NVIQ	−0.08	
VIQ	−0.31 *	$p = 0.030$

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

The correlation between VIQ and Code-Switching was not significant after applying a Holm–Bonferroni correction for multiple comparisons. Still, for extra caution, we ran separate models, in addition to the combined model for these variables.

Singapore Sample

	Balanced Usage	
Balanced Usage	–	
Age	0.07	
SES	0.06	
NVIQ	−0.38 *	$p = 0.021$
VIQ	0.0006	

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

After applying a Holm–Bonferroni correction for multiple comparisons, the correlation between NVIQ and Balanced Usage was no longer significant. In any case, none of the significant models included a combination of the two variables, so it did not affect our analysis.

	Language Entropy
Language Entropy	–
Age	−0.05
SES	−0.16
NVIQ	0.30
VIQ	−0.16

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

	Balanced Proficiency	
Balanced Proficiency	–	
Age	−0.13	
SES	−0.04	
NVIQ	−0.03	
VIQ	0.43 **	$p = 0.009$

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

The correlation between VIQ and Balanced Proficiency remained significant even after applying a Holm–Bonferroni correction for multiple comparisons. We therefore ran separate models for these measures and report the best one, where relevant.

	L2 AoA	
L2 AoA	–	
Age	0.31	
SES	−0.06	
NVIQ	0.02	
VIQ	0.44 **	$p = 0.007$

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

The correlation between VIQ and L2 AoA remained significant even after applying a Holm–Bonferroni correction for multiple comparisons. Nevertheless, none of the significant models included a combination of the two variables, so the correlation did not affect our analysis.

	Single-Language
Single-Language	–
Age	−0.13
SES	0.04
NVIQ	−0.15
VIQ	−0.10

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

	Dual-Language
Dual-Language	–
Age	0.006
SES	−0.07
NVIQ	0.20
VIQ	−0.16

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

	Code-Switching
Code-Switching	–
Age	−0.04
SES	0.003
NVIQ	0.01
VIQ	−0.13

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Appendix D Regression Models

UK Sample

Task-Set Switching (TSS) Accuracy

AIC	BIC
−104.46	−98.72

	Estimate	β	**S.E.**	t	p
TSS Acc cost	(Intercept)	−0.082	0.071	−1.154	0.254
	VIQ	0.003	0.001	2.800	0.007 **

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. $F(1,48) = 7.84$, $p = 0.007$, $R^2 = 0.14$, $R^2_{\text{adjusted}} = 0.12$.

Spatial Working Memory (SWM) Strategy

AIC	BIC
251.78	264.42

	Estimate	β	S.E.	t	p
SWM Strategy	(Intercept)	35.933	7.030	5.111	<0.0001 ***
	L2_AoA	−0.095	0.132	−0.716	0.478
	Age	−0.317	0.187	−1.693	0.098
	SES	−9.861	3.201	−3.081	0.004 **
	NVIQ	−3.896	2.362	−1.649	0.107
	VIQ	−0.113	0.054	−2.105	0.042

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. $F(5,39) = 4.03$, $p = 0.005$, $R^2 = 0.34$, $R^2_{adjusted} = 0.26$.

Due to a correlation between Age and VIQ (which did not remain significant after multiple-comparison correction), we also ran separate models. The model with VIQ (excluding Age) was better, so we report it below:

AIC	BIC
252.97	263.81

	Estimate	β	S.E.	t	p
SWM Strategy	(Intercept)	26.193	4.133	6.338	<0.0001 ***
	L2_AoA	−0.095	0.136	−0.703	0.486
	SES	−9.915	3.275	−3.028	0.004 **
	NVIQ	−3.867	2.417	−1.600	0.117
	VIQ	−0.081	0.052	−1.578	0.122

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. $F(4,40) = 4.13$, $p = 0.007$, $R^2 = 0.29$, $R^2_{adjusted} = 0.22$.

Structure Learning (SL) Performance Index (PI)

AIC	BIC
−68.21	−58.55

	Estimate	β	S.E.	t	p
SL PI improvement	(Intercept)	−0.184	0.122	−1.510	0.138
	Balanced Proficiency	−0.022	0.008	−2.635	0.011 *
	Code-Switching	−0.021	0.016	−1.324	0.192
	VIQ	0.006	0.002	3.486	0.001 **

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. $F(5,39) = 6.26$, $p = 0.001$, $R^2 = 0.29$, $R^2_{adjusted} = 0.24$.

Due to correlations between Balanced Proficiency and VIQ, and Code-Switching and VIQ (which did not remain significant after multiple-comparison correction), we alsoran separate models. The model with VIQ (excluding Balanced Proficiency and Code-Switching) was better, so we report it below:

AIC	BIC
−64.52	−58.72

	Estimate	β	S.E.	t	p
SL PI improvement	(Intercept)	−0.225	0.107	−2.108	0.040 *
	VIQ	0.005	0.002	3.158	0.003 **

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. $F(1{,}49) = 9.975$, $p = 0.003$, $R^2 = 0.17$, $R^2_{\text{adjusted}} = 0.15$.

Singapore Sample

Trail Making Test (TMT) Ratio

AIC	BIC
27.80	34.14

	Estimate	β	S.E.	t	p
TMT B:A Ratio	(Intercept)	0.855	0.121	7.051	<0.0001 ***
	L2 AoA	0.059	0.029	2.009	0.053
	Balanced Proficiency	0.123	0.052	2.377	0.0234 *

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. $F(2{,}33) = 5.90$, $p = 0.006$, $R^2 = 0.26$, $R^2_{\text{adjusted}} = 0.22$.

Intra-Extra Dimensional (IED) Set Shift—Total Errors (with Outliers Included)

AIC	BIC
349.23	357.15

	Estimate	β	S.E.	t	p
IED Total Errors (incl. outliers)	(Intercept)	−138.007	59.271	−2.328	0.026 *
	Single-Language	8.232	4.882	1.686	0.102
	Code-Switching	15.705	5.704	2.753	0.010 **
	Age	3.627	2.267	1.600	.119

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. $F(3{,}32) = 3.632$, $p = 0.023$, $R^2 = 0.25$, $R^2_{\text{adjusted}} = 0.18$.

(Without Outliers)

AIC	BIC
177.47	181.78

	Estimate	β	S.E.	t	p
IED Total Errors (w/o Outliers)	(Intercept)	0.392	4.835	0.081	0.936
	VIQ	0.224	0.091	2.454	0.020 **

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. $F(1{,}29) = 6.024$, $p = 0.020$, $R^2 = 0.17$, $R^2_{\text{adjusted}} = 0.14$.

Intra-Extra Dimensional (IED) Set Shift—Extra-Dimensional Shift Errors

AIC	BIC
238.48	246.40

	Estimate	***β***	**S.E.**	***t***	***p***
IED Extra-Dimensional Shift Errors	(Intercept)	−24.925	10.748	−2.319	0.027 *
	Language Entropy	11.366	2.779	4.090	0.0002 ***
	Age	1.362	0.483	2.823	0.008 **
	NVIQ	−11.077	4.5547	−2.432	0.021 *

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. $F(3,32) = 8.618$, $p = 0.0002$, $R^2 = 0.45$, $R^2_{adjusted} = 0.40$.

Stroop Task Reaction Time (RT)

AIC	BIC
428.86	436.78

	Estimate	***β***	**S.E.**	***t***	***p***
Stroop RT Cost	(Intercept)	−35.414	94.63	−2.086	0.045
	Balanced Proficiency	29.570	13.49	2.192	0.036 *
	Single-Language	56.830	14.76	3.850	0.0005 ***
	Code-Switching	43.940	17.54	2.505	0.018 *

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. $F(3,32) = 6.748$, $p = 0.001$, $R^2 = 0.39$, $R^2_{adjusted} = 0.33$.

Spatial Working Memory (SWM) Total Errors

AIC	BIC
318.22	324.55

	Estimate	***β***	**S.E.**	***t***	***p***
SWM Errors	(Intercept)	26.592	10.987	2.420	0.021 *
	Language Entropy	−13.456	8.216	−1.638	0.111
	SES	30.415	18.528	1.642	0.110

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. $F(2,33) = 3.21$, $p = 0.053$, $R^2 = 0.16$, $R^2_{adjusted} = 0.11$.

Spatial Working Memory (SWM) Strategy

AIC	BIC
205.07	211.41

	Estimate	***β***	**S.E.**	***t***	***p***
SWM Strategy	(Intercept)	11.843	2.282	5.189	<0.0001 ***
	Language Entropy	−3.426	1.707	−2.007	0.0530
	SES	9.473	3.849	2.461	0.0192 *

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. $F(2,33) = 6.005$, $p = 0.006$, $R^2 = 0.27$, $R^2_{adjusted} = 0.22$.

Backward Digit Span (BDS)

AIC	BIC
181.81	189.72

	Estimate	β	S.E.	t	p
BDS	(Intercept)	10.656	2.264	4.708	<0.0001 ***
	Code-Switching	−0.944	0.551	−1.714	0.096
	SES	−4.495	2.842	−1.582	0.124
	NVIQ	6.677	2.066	3.233	0.003 **

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. $F(3,32) = 4.56$, $p = 0.009$, $R^2 = 0.30$, $R^2_{\text{adjusted}} = 0.23$.

Structure Learning (SL) Performance Index (PI)

AIC	BIC
−51.40	−45.07

	Estimate	β	S.E.	t	p
SL PI improvement	(Intercept)	−0.234	0.065	−3.614	<0.001 ***
	Language Entropy	0.179	0.048	3.693	0.0008 ***
	SES	0.437	0.109	3.999	0.0003 ***

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. $F(2,33) = 12.75$, $p < 0.0001$, $R^2 = 0.44$, $R^2_{\text{adjusted}} = 0.40$.

Structure Learning (SL) Strategy ICD

AIC	BIC
−34.66	−29.90

	Estimate	β	S.E.	t	p
SL Strategy ICD	(Intercept)	−0.236	0.046	−5.154	<0.0001 ***
	Language Entropy	0.116	0.061	1.894	0.0667

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. $F(1,34) = 3.589$, $p = 0.067$, $R^2 = 0.10$, $R^2_{\text{adjusted}} = 0.07$.

Notes

1 The "Global South" is a term that has been traditionally used to refer to economically disadvantaged nations. A broader definition incorporates countries that have historically frequently faced colonisation by Global North countries (especially European), have unstable democracies, and/or are in the process of industrialising (Finance Center for South-South Cooperation n.d.; World Population Review 2023). We acknowledge that our sample of young University students in Singapore comprise a relatively privileged group compared to other areas of the Global South.

2 Mean self-rated English Proficiency for the UK sample was 5.9 out of 6, and for the Singapore sample was 5.82 out of 6 (the groups did not differ significantly $t = 1.12$, $p = 0.27$).

3 We tried to use well-validated tasks, whose psychometric properties have been thoroughly analysed, and which have been used as indicators of general intelligence in multiple areas of the world (Kaufman et al. 2006; Raven 2000). However, we recognise that

"culture-free" measures of intelligence suffer certain weaknesses (Gonthier 2022; Walker et al. 2009), and this should be taken into account in future studies.

4 For participants who listed a single language, a score of 0 was added to L2 Proficiency in order to calculate Balanced Proficiency score (i.e., difference in proficiency between most proficient and second most proficient language).

5 As most of our participants reported 6 out of 6 in L1 Proficiency, the Balanced Proficiency score mostly depends on L2 Proficiency.

6 For the calculation of Balanced Usage, we only included the first two most used languages, weighted against each other. For example, if a participant reported 60% use of English, 30% use of French, and 10% use of German, we excluded German and converted the percentages for English and French to 67% and 33% respectively. The outputs of the models, however, were not different when we used unconverted percentages (unweighted Balanced Usage).

7 L2 here can technically be another L1 for participants that acquired both languages at birth.

8 As mentioned earlier, this interaction may also be present in the UK (and Europe more generally) in big centres with greater cultural and linguistic variation, but societal attitudes still privilege the major languages (Bunk and Wiese 2024).

9 We must point out that there was an apparent correlation between Balanced Proficiency and VIQ in the UK sample, but this stopped being significant when we corrected for multiple comparisons. Still, when we ran separate models, the model with VIQ was better than the one with Balanced Proficiency.

References

Antón, Eneko, Jon A. Duñabeitia, Adelina Estévez, Juan A. Hernández, Alejandro Castillo, Luis J. Fuentes, Douglas J. Davidson, and Manuel Carreiras. 2014. Is There a Bilingual Advantage in the ANT Task? Evidence from Children. *Frontiers in Psychology* 5: 1–12. [CrossRef]

Antón, Eneko, Manuel Carreiras, and Jon Andoni Duñabeitia. 2019. The Impact of Bilingualism on Executive Functions and Working Memory in Young Adults. *PLoS ONE* 14: e0206770. [CrossRef] [PubMed]

Antón, Eneko, Yuriem Fernández García, Manuel Carreiras, and Jon Andoni Duñabeitia. 2016. Does Bilingualism Shape Inhibitory Control in the Elderly? *Journal of Memory and Language* 90: 147–60. [CrossRef]

Armitage, Stewart G. 1946. An Analysis of Certain Psychological Tests Used for the Evaluation of Brain Injury. *Psychological Monographs* 60: i. [CrossRef]

Bak, Thomas H. 2015. Beyond a Simple 'Yes' and 'No'. *Cortex* 73: 332–33. [CrossRef] [PubMed]

Bak, Thomas H., and Suvarna Alladi. 2016. Bilingualism, Dementia and the Tale of Many Variables: Why We Need to Move beyond the Western World. Commentary on Lawton et Al. (2015) and Fuller-Thomson (2015). *Cortex* 74: 315–17. [CrossRef]

Bartolotti, James, Viorica Marian, Scott R. Schroeder, and Anthony Shook. 2011. Bilingualism and Inhibitory Control Influence Statistical Learning of Novel Word Forms. *Frontiers in Psychology* 2: 1–10. [CrossRef] [PubMed]

Bialystok, Ellen. 2009. Bilingualism: The Good, the Bad, and the Indifferent. *Bilingualism: Language and Cognition* 12: 3–11. [CrossRef]

Bialystok, Ellen. 2010. Global-Local and Trail-Making Tasks by Monolingual and Bilingual Children: Beyond Inhibition. *Developmental Psychology* 46: 93–105. [CrossRef]

Bialystok, Ellen, Raymond Klein, Fergus I.M. Craik, and Mythili Viswanathan. 2004. Bilingualism, Aging, and Cognitive Control: Evidence from the Simon Task. *Psychology and Aging* 19: 290–303. [CrossRef]

Blackburn, Harold L., and Arthur L. Benton. 1957. Revised Administration and Scoring of the Digit Span Test. *Journal of Consulting Psychology* 21: 139–43. [CrossRef] [PubMed]

Blom, Elma, Aylin C. Küntay, Marielle Messer, Josje Verhagen, and Paul Leseman. 2014. The Benefits of Being Bilingual: Working Memory in Bilingual Turkish–Dutch Children. *Journal of Experimental Child Psychology* 128: 105–19. [CrossRef] [PubMed]

Bonifacci, Paola, Lucia Giombini, Stéphanie Bellocchi, and Silvana Contento. 2011. Speed of Processing, Anticipation, Inhibition and Working Memory in Bilinguals. *Developmental Science* 14: 256–69. [CrossRef] [PubMed]

Bunk, Oliver, and Heike Wiese. 2024. Contact Dialects in Urban Youth Culture and Beyond. In *The Routledge Handbook of Language and Youth Culture*, 1st ed. Edited by Bente A. Svendsen and Rickard Jonsson. New York: Routledge, pp. 333–46. [CrossRef]

Cambridge Cognition. 2019. CANTAB® [Cognitive Assessment Software]. Available online: www.cantab.com (accessed on 1 March 2020).

Carlson, Stephanie M., and Andrew N. Meltzoff. 2008. Bilingual Experience and Executive Functioning in Young Children. *Developmental Science* 11: 282–98. [CrossRef] [PubMed]

Cavallaro, Francesco, and Ng Bee Chin. 2014. Language in Singapore: From Multilingualism to English Plus. In *Challenging the Monolingual Mindset*. Edited by John Hajek and Yvette Slaughter. Bristol: Channel View Publications, pp. 33–48. [CrossRef]

Cockcroft, Kate. 2022. Working Memory in Multilinguals: Factors That Assist and Impede. *International Journal of Multilingualism* 21: 997–1013. [CrossRef]

Cools, Roshan, Luke Clark, Adrian M. Owen, and Trevor W. Robbins. 2002. Defining the Neural Mechanisms of Probabilistic Reversal Learning Using Event-Related Functional Magnetic Resonance Imaging. *The Journal of Neuroscience* 22: 4563–67. [CrossRef] [PubMed]

Council of Europe. 2020. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment—Companion Volume*. Edited by Council of Europe Publishing. Available online: www.coe.int/lang-cefr (accessed on 10 March 2020).

Cunnings, Ian. 2012. An Overview of Mixed-Effects Statistical Models for Second Language Researchers. *Second Language Research* 28: 369–82. [CrossRef]

Dixon, L. Quentin. 2005. Bilingual Education Policy in Singapore: An Analysis of Its Sociohistorical Roots and Current Academic Outcomes. *International Journal of Bilingual Education and Bilingualism* 8: 25–47. [CrossRef]

Dormann, Carsten F., Jane Elith, Sven Bacher, Carsten Buchmann, Gudrun Carl, Gabriel Carré, Jaime R. García Marquéz, Bernd Gruber, Bruno Lafourcade, Pedro J. Leitão, and et al. 2013. Collinearity: A Review of Methods to Deal with It and a Simulation Study Evaluating Their Performance. *Ecography* 36: 27–46. [CrossRef]

Duñabeitia, Jon Andoni, Juan Andrés Hernández, Eneko Antón, Pedro Macizo, Adelina Estévez, Luis J. Fuentes, and Manuel Carreiras. 2014. The Inhibitory Advantage in Bilingual Children Revisited. *Experimental Psychology* 61: 234–51. [CrossRef]

Estanga, Ainara, Mirian Ecay-Torres, Almudena Ibañez, Andrea Izagirre, Jorge Villanua, Maite Garcia-Sebastian, M. Teresa Iglesias Gaspar, Ane Otaegui-Arrazola, Ane Iriondo, Monserrat Clerigue, and et al. 2017. Beneficial Effect of Bilingualism on Alzheimer's Disease CSF Biomarkers and Cognition. *Neurobiology of Aging* 50: 144–51. [CrossRef] [PubMed]

Filippi, Roberto, Andrea Ceccolini, Elizabeth Booth, Chen Shen, Michael S. C. Thomas, Mireille B. Toledano, and Iroise Dumontheil. 2022. Modulatory Effects of SES and Multilinguistic Experience on Cognitive Development: A Longitudinal Data Analysis of Multilingual and Monolingual Adolescents from the SCAMP Cohort. *International Journal of Bilingual Education and Bilingualism* 25: 3489–506. [CrossRef] [PubMed]

Finance Center for South-South Cooperation. n.d. Global South Countries (Group of 77 and China)—Partnership Program—The Finance Center for South-South Cooperation. Available online: http://www.fc-ssc.org/en/partnership_program/south_south_countries (accessed on 1 October 2023).

Fuller-Thomson, Esme, and Diana Kuh. 2014. The Healthy Migrant Effect May Confound the Link between Bilingualism and Delayed Onset of Alzheimer's Disease. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* 52: 128–30. [CrossRef] [PubMed]

Gallo, Federico, Nikolay Novitskiy, Andriy Myachykov, and Yury Shtyrov. 2021. Individual Differences in Bilingual Experience Modulate Executive Control Network and Performance: Behavioral and Structural Neuroimaging Evidence. *Bilingualism: Language and Cognition* 24: 293–304. [CrossRef]

Giorgio, Joseph, Vasilis M. Karlaftis, Rui Wang, Yuan Shen, Peter Tino, Andrew Welchman, and Zoe Kourtzi. 2018. Functional Brain Networks for Learning Predictive Statistics. *Cortex* 107: 204–19. [CrossRef]

Goldberg, Wendy A., Jo Ann Prause, Rachel Lucas-Thompson, and Amy Himsel. 2008. Maternal Employment and Children's Achievement in Context: A Meta-Analysis of Four Decades of Research. *Psychological Bulletin* 134: 77–108. [CrossRef] [PubMed]

Gonthier, Corentin. 2022. Cross-Cultural Differences in Visuo-Spatial Processing and the Culture-Fairness of Visuo-Spatial Intelligence Tests: An Integrative Review and a Model for Matrices Tasks. *Cognitive Research: Principles and Implications* 7: 11. [CrossRef] [PubMed]

Gorilla. 2022. Available online: https://www.gorilla.sc (accessed on 1 January 2022).

Gough, Vicky. 2023. The British and Their Secret Language Regrets | British Council. British Council. Available online: https://www.britishcouncil.org/voices-magazine/british-and-their-secret-language-regret (accessed on 15 February 2024).

Grant, David A., and Esta Berg. 1948. A Behavioral Analysis of Degree of Reinforcement and Ease of Shifting to New Responses in a Weigl-Type Card-Sorting Problem. *Journal of Experimental Psychology* 38: 404–11. [CrossRef]

Green, David W., and Jubin Abutalebi. 2013. The Adaptive Control Hypothesis. *Journal of Cognitive Psychology* 25: 515–30. [CrossRef]

Grosjean, François. 2021. The Extent of Bilingualism. In *Life as a Bilingual*. Cambridge: Cambridge University Press, pp. 27–39. [CrossRef]

Gullifer, Jason W., and Debra Titone. 2018. Compute Language Entropy with {languageEntropy}. Available online: https://github.com/jasongullifer/languageEntropy (accessed on 5 May 2021).

Gullifer, Jason W., and Debra Titone. 2020. Characterizing the Social Diversity of Bilingualism Using Language Entropy. *Bilingualism: Language and Cognition* 23: 283–94. [CrossRef]

Gullifer, Jason W., Xiaoqian J. Chai, Veronica Whitford, Irina Pivneva, Shari Baum, Denise Klein, and Debra Titone. 2018. Bilingual Experience and Resting-State Brain Connectivity: Impacts of L2 Age of Acquisition and Social Diversity of Language Use on Control Networks. *Neuropsychologia* 117: 123–34. [CrossRef] [PubMed]

Hartanto, Andree, and Hwajin Yang. 2016. Disparate Bilingual Experiences Modulate Task-Switching Advantages: A Diffusion-Model Analysis of the Effects of Interactional Context on Switch Costs. *Cognition* 150: 10–19. [CrossRef] [PubMed]

Hernández, Mireia, Albert Costa, Luis J. Fuentes, Ana B. Vivas, and Núria Sebastián-Gallés. 2010. The Impact of Bilingualism on the Executive Control and Orienting Networks of Attention. *Bilingualism: Language and Cognition* 13: 315–25. [CrossRef]

Holm, Sture. 1979. A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics* 6: 65–70.

I-ABC: A Personalised Learning Study—Adaptive Brain Lab. n.d. Available online: https://www.abg.psychol.cam.ac.uk/participation/iabc (accessed on 1 March 2020).

Inquisit 6. 2021. Available online: https://www.millisecond.com (accessed on 1 January 2021).

Jevtović, Mina, Jon Andoni Duñabeitia, and Angela de Bruin. 2020. How Do Bilinguals Switch between Languages in Different Interactional Contexts? A Comparison between Voluntary and Mandatory Language Switching. *Bilingualism: Language and Cognition* 23: 401–13. [CrossRef]

Jia, Fanlu. 2022. Effect of Second Language Proficiency on Inhibitory Control in the Simon Task: An FMRI Study. *Frontiers in Psychology* 13: 812322. [CrossRef]

Jiao, Lu, Yuanyue Zhang, Patrick Plummer, Cong Liu, and Baoguo Chen. 2019. The Influence of Bilingual Language Experience on Executive Control: An ERPs Study. *Journal of Neurolinguistics* 51: 42–52. [CrossRef]

Jylkkä, Jussi, Anna Soveri, Jenny Wahlström, Minna Lehtonen, Antoni Rodríguez-Fornells, and Matti Laine. 2017. Relationship between Language Switching Experience and Executive Functions in Bilinguals: An Internet-Based Study. *Journal of Cognitive Psychology* 29: 404–19. [CrossRef]

Jylkkä, Jussi, Matti Laine, and Minna Lehtonen. 2021. Does Language Switching Behavior Rely on General Executive Functions? *Bilingualism: Language and Cognition* 24: 583–95. [CrossRef]

Jylkkä, Jussi, Minna Lehtonen, Fred Lindholm, Anna Kuusakoski, and Matti Laine. 2018. The relationship between general executive functions and bilingual switching and monitoring in language production. *Bilingualism: Language and Cognition* 21: 505–22. [CrossRef]

Karlaftis, Vasilis M., Joseph Giorgio, Petra E. Vértes, Rui Wang, Yuan Shen, Peter Tino, Andrew E. Welchman, and Zoe Kourtzi. 2019. Multimodal Imaging of Brain Connectivity Reveals Predictors of Individual Decision Strategy in Statistical Learning. *Nature Human Behaviour* 3: 297–307. [CrossRef] [PubMed]

Karlaftis, Vasilis M., Rui Wang, Yuan Shen, Peter Tino, Guy Williams, Andrew E. Welchman, and Zoe Kourtzi. 2018. White-Matter Pathways for Statistical Learning of Temporal Structures. *ENeuro* 5: 382–99. [CrossRef]

Kaufman, Alan S., Dawn P. Flanagan, Vincent C. Alfonso, and Jennifer T. Mascolo. 2006. Test Review: Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV). *Journal of Psychoeducational Assessment* 24: 278–95. [CrossRef]

Kehagia, Angie A., Rong Ye, Dan W. Joyce, Orla M. Doyle, James B. Rowe, and Trevor W. Robbins. 2017. Parsing the Roles of the Frontal Lobes and Basal Ganglia in Task Control Using Multivoxel Pattern Analysis. *Journal of Cognitive Neuroscience* 29: 1390–1401. [CrossRef] [PubMed]

Kerrigan, Lucy, Michael S. C. Thomas, Peter Bright, and Roberto Filippi. 2017. Evidence of an Advantage in Visuo-Spatial Memory for Bilingual Compared to Monolingual Speakers. *Bilingualism: Language and Cognition* 20: 602–12. [CrossRef]

Kirk, Neil W., Linda Fiala, Kenneth C. Scott-Brown, and Vera Kempe. 2014. No Evidence for Reduced Simon Cost in Elderly Bilinguals and Bidialectals. *Journal of Cognitive Psychology* 26: 640–48. [CrossRef] [PubMed]

Kopp, Bruno. 2011. Neuropsychologists Must Keep Their Eyes on the Reliability of Difference Measures. *Journal of the International Neuropsychological Society* 17: 562–63. [CrossRef] [PubMed]

Kuhn, Max, and Kjell Johnson. 2013. *Applied Predictive Modeling*. New York, Dordrecht, London and Berlin/Heidelberg: Dordrecht: London: Berlin/Heidelberg: Springer. [CrossRef]

Lai, Gabrielle, and Beth A. O'Brien. 2020. Examining Language Switching and Cognitive Control Through the Adaptive Control Hypothesis. *Frontiers in Psychology* 11: 525661. [CrossRef]

Laine, Matti, and Minna Lehtonen. 2018. Cognitive Consequences of Bilingualism: Where to Go from Here? *Language, Cognition and Neuroscience* 33: 1205–12. [CrossRef]

Leimgruber, Jakob R. E. 2013. The Management of Multilingualism in a City-State: Language Policy in Singapore. In *Multilingualism and Language Diversity in Urban Areas: Acquisition, Identities, Space, Education*. Edited by Peter Siemund, Ingrid Gogolin, Monika Schulz and Julia Davydova. Amsterdam: John Benjamins, pp. 227–56.

Leong, Victoria, Kausar Raheel, Jia Yi Sim, Kriti Kacker, Vasilis M. Karlaftis, Chrysoula Vassiliu, Kastoori Kalaivanan, S. H. Annabel Chen, Trevor W. Robbins, Barbara J. Sahakian, and et al. 2022. A New Remote Guided Method for Supervised Web-Based Cognitive Testing to Ensure High-Quality Data: Development and Usability Study. *Journal of Medical Internet Research* 24: e28368. [CrossRef]

Li, Xiaoqian, Kwun Kei Ng, Joey Ju Yu Wong, Jia Wen Lee, Juan Helen Zhou, and W. Quin Yow. 2021. Bilingual Language Entropy Influences Executive Functions through Functional Connectivity and Signal Variability. *Brain and Language* 222: 105026. [CrossRef] [PubMed]

Luk, Gigi, and Ellen Bialystok. 2013. Bilingualism Is Not a Categorical Variable: Interaction between Language Proficiency and Usage. *Journal of Cognitive Psychology* 25: 605–21. [CrossRef]

Luk, Gigi, Eric De Sa, and Ellen Bialystok. 2011. Is There a Relation between Onset Age of Bilingualism and Enhancement of Cognitive Control? *Bilingualism: Language and Cognition* 14: 588–95. [CrossRef]

Luo, Lin, Fergus I. M. Craik, Sylvain Moreno, and Ellen Bialystok. 2013. Bilingualism Interacts with Domain in a Working Memory Task: Evidence from Aging. *Psychology and Aging* 28: 28–34. [CrossRef]

Makoni, Sinfree, Anna Kaiper-Marquez, and Lorato Mokwena. 2022. *The Routledge Handbook of Language and the Global South-s*. London: Routledge.

Miyake, Akira, Naomi P. Friedman, Michael J. Emerson, Alexander H. Witzki, Amy Howerter, and Tor D. Wager. 2000. The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cognitive Psychology* 41: 49–100. [CrossRef] [PubMed]

Morales, Julia, Alejandra Calvo, and Ellen Bialystok. 2013. Working Memory Development in Monolingual and Bilingual Children. *Journal of Experimental Child Psychology* 114: 187–202. [CrossRef]

Office for National Statistics. 2022. Language, England and Wales: Census 2021. November. Available online: https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/language/bulletins/languageenglandandwales/census2021 (accessed on 15 February 2023).

Ong, Kenneth Keng Wee, and Lawrence Jun Zhang. 2010. Metalinguistic Filters within the Bilingual Language Faculty: A Study of Young English-Chinese Bilinguals. *Journal of Psycholinguistic Research* 39: 243–72. [CrossRef]

Onnis, Luca, Win Ee Chun, and Matthew Lou-Magnuson. 2018. Improved Statistical Learning Abilities in Adult Bilinguals. *Bilingualism: Language and Cognition* 21: 427–33. [CrossRef]

Ooi, Seok Hui, Winston D. Goh, Antonella Sorace, and Thomas H. Bak. 2018. From Bilingualism to Bilingualisms: Bilingual Experience in Edinburgh and Singapore Affects Attentional Control Differently. *Bilingualism* 21: 867–79. [CrossRef]

Paap, Kenneth R., and Zachary I. Greenberg. 2013. There Is No Coherent Evidence for a Bilingual Advantage in Executive Processing. *Cognitive Psychology* 66: 232–58. [CrossRef]

Paap, Kenneth R., Hunter A. Johnson, and Oliver Sawi. 2014. Are Bilingual Advantages Dependent upon Specific Tasks or Specific Bilingual Experiences? *Journal of Cognitive Psychology* 26: 615–39. [CrossRef]

Potter, Christine E., Tianlin Wang, and Jenny R. Saffran. 2017. Second Language Experience Facilitates Statistical Learning of Novel Linguistic Materials. *Cognitive Science* 41: 913–27. [CrossRef] [PubMed]

Prior, Anat, and Brian Macwhinney. 2010. A Bilingual Advantage in Task Switching. *Bilingualism: Language and Cognition* 13: 253–62. [CrossRef] [PubMed]

R Core Team. 2021. *R: A Language and Environment for Statistical Computing. Foundation for Statistical Computing*. Vienna: R Core Team.

Raven, John. 2000. The Raven's Progressive Matrices: Change and Stability over Culture and Time. *Cognitive Psychology* 41: 1–48. [CrossRef] [PubMed]

Raven, John C., John Hugh Court, and John Earle Raven. 2000. *Standard Progressive Matrices*. London: Psychology Press.

Robbins, Trevor W., Merle James, Adrian M. Owen, Barbara J. Sahakian, Lynn McInnes, and Patrick Rabbitt. 1994. Cambridge Neuropsychological Test Automated Battery (CANTAB): A Factor Analytic Study of a Large Sample of Normal Elderly Volunteers. *Dementia* 5: 266–81. [CrossRef] [PubMed]

Siemund, Peter, Monika Edith Schulz, and Martin Schweinberger. 2014. Studying the Linguistic Ecology of Singapore: A Comparison of College and University Students. *World Englishes* 33: 340–62. [CrossRef]

Soveri, Anna, Antoni Rodriguez-Fornells, and Matti Laine. 2011. Is There a Relationship between Language Switching and Executive Functions in Bilingualism? Introducing a within Group Analysis Approach. *Frontiers in Psychology* 2: 10197. [CrossRef]

Stasenko, Alena, Georg E. Matt, and Tamar H. Gollan. 2017. A Relative Bilingual Advantage in Switching with Preparation: Nuanced Explorations of the Proposed Association Between Bilingualism and Task Switching. *Journal of Experimental Psychology. General* 146: 1527. [CrossRef]

Stroop, J. Ridley. 1935. Studies of Interference in Serial Verbal Reactions. *Journal of Experimental Psychology* 18: 643. [CrossRef]

Sullivan, Margot D., Yolanda Prescott, Devora Goldberg, and Ellen Bialystok. 2016. Executive Control Processes in Verbal and Nonverbal Working Memory: The Role of Aging and Bilingualism. *Linguistic Approaches to Bilingualism* 6: 147–70. [CrossRef]

Torres, Valeria L., Mónica Rosselli, David A. Loewenstein, Merike Lang, Idaly Vélez-Uribe, Fernanda Arruda, Joshua Conniff, Rosie E. Curiel, Maria T. Greig, Warren W. Barker, and et al. 2022. The Contribution of Bilingualism to Cognitive Functioning and Regional Brain Volume in Normal and Abnormal Aging. *Bilingualism: Language and Cognition* 25: 337–56. [CrossRef]

Venables, William N., and Brian D. Ripley. 2002. *Modern Applied Statistics with S*. New York: Springer.

Verhagen, Josje, and Elise de Bree. 2021. Effects of Bilingualism on Statistical Learning in Preschoolers. *Linguistic Approaches to Bilingualism* 11: 611–39. [CrossRef]

Verreyt, Nele, Evy Woumans, Davy Vandelanotte, Arnaud Szmalec, and Wouter Duyck. 2016. The Influence of Language-Switching Experience on the Bilingual Executive Control Advantage. *Bilingualism: Language and Cognition* 19: 181–90. [CrossRef]

Walker, Alexandra J., Jennifer Batchelor, and Arthur Shores. 2009. Effects of Education and Cultural Background on Performance on WAIS-III, WMS-III, WAIS-R and WMS-R Measures: Systematic Review. *Australian Psychologist* 44: 216–23. [CrossRef]

Wang, Rui, Yuan Shen, Peter Tino, Andrew E. Welchman, and Zoe Kourtzi. 2017a. Learning Predictive Statistics from Temporal Sequences: Dynamics and Strategies. *Journal of Vision* 17: 1–16. [CrossRef] [PubMed]

Wang, Rui, Yuan Shen, Peter Tino, Andrew E. Welchman, and Zoe Kourtzi. 2017b. Learning Predictive Statistics: Strategies and Brain Mechanisms. *The Journal of Neuroscience* 37: 8412. [CrossRef] [PubMed]

Wechsler, David. 2011. *Wechsler Abbreviated Scale of Intelligence–Second Edition (WASI-II)*. San Antonio: NCS Pearson.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. New York: Springer.

Wigdorowitz, Mandy, Ana I. Pérez, and Ianthi M. Tsimpli. 2020. A Holistic Measure of Contextual and Individual Linguistic Diversity. *International Journal of Multilingualism* 20: 469–87. [CrossRef]

Wigdorowitz, Mandy, Ana I. Pérez, and Ianthi M. Tsimpli. 2022. Sociolinguistic Context Matters: Exploring Differences in Contextual Linguistic Diversity in South Africa and England. *International Multilingual Research Journal* 16: 345–64. [CrossRef]

World Population Review. 2023. Global South Countries 2023. Available online: https://worldpopulationreview.com/country-rankings/global-south-countries (accessed on 15 March 2023).

Woumans, Evy, Shauni Van Herck, and Esli Struys. 2019. Shifting Gear in the Study of the Bilingual Advantage: Language Switching Examined as a Possible Moderator. *Behavioral Sciences* 9: 86. [CrossRef]

Xie, Wenhan, and Francesco Cavallaro. 2016. Attitudes towards Mandarin–English Bilingualism: A Study of Chinese Youths in Singapore. *Journal of Multilingual and Multicultural Development* 37: 628–41. [CrossRef]

Yang, Sujin, Hwajin Yang, and Barbara Lust. 2011. Early Childhood Bilingualism Leads to Advances in Executive Attention: Dissociating Culture and Language. *Bilingualism: Language and Cognition* 14: 412–22. [CrossRef]

Yin, Jing, Connie Qun Guan, Elaine R. Smolen, Esther Geva, and Wanjin Meng. 2022. Does Early Exposure to Chinese–English Biliteracy Enhance Cognitive Skills? *Frontiers in Psychology* 13: 2074. [CrossRef] [PubMed]

Yow, W. Quin, and Xiaoqian Li. 2015. Balanced Bilingualism and Early Age of Second Language Acquisition as the Underlying Mechanisms of a Bilingual Executive Control Advantage: Why Variations in Bilingual Experiences Matter. *Frontiers in Psychology* 6: 123975. [CrossRef] [PubMed]