Bilingualism: Language and Cognition

cambridge.org/bil

Research Article

Cite this article: Privitera, A.J. (2025). Bilingual effects on cognitive control: Are we looking in the right place? *Bilingualism: Language and Cognition* 1–14. https://doi.org/10.1017/S1366728925100576

Received: 22 January 2025
Revised: 15 June 2025
Accepted: 26 August 2025

Keywords:
bilingualism; cognitive control; bilingual effects; Bayesian modeling; performance trajectory

Corresponding author:
Adam John Privitera;
Email: aprivite@connect.hku.hk

This research article was awarded Open Data badge for transparent practices. See the Data Availability Statement for details.

Bilingual effects on cognitive control: Are we looking in the right place?

Adam John Privitera[1,2]

[1]Science of Learning in Education Centre, National Institute of Education, Nanyang Technological University, Singapore, Singapore and [2]Centre for Research and Development in Learning, Nanyang Technological University, Singapore, Singapore

Abstract

Whether bilingualism confers non-linguistic cognitive advantages continues to generate both interest and debate in the psychological sciences. In response to mixed reports and methodological critiques, researchers have embraced more rigorous practices when investigating bilingual effects, including those in the domain of cognitive control. Despite considerable advances, one significant issue persists: the assumption that task performance remains stable over time. To address this, the present study investigated the relationship between bilingual language experience and Simon task performance modeled as a continuous function of time. In a sample of Mandarin-English bilingual young adults, we identified distinct patterns of results across both conventional and time-sensitive performance trajectory measures with each supporting a different relationship between language experience and cognitive control. Results suggest that reliance on conventional performance measures may be partially responsible for mixed results, necessitating reevaluation of how bilingual effects on cognitive control manifest and which analysis methods best support their identification.

Highlights

- Bilingualism's cognitive benefits remain debated despite rigorous research.
- Existing work on bilingual effects assumes that task performance is stable.
- We investigate bilingual effects on time-sensitive performance measures.
- Different measures reveal distinct links between language experience and cognition.
- Conventional methods may contribute to inconsistent findings in bilingual research.

1. Introduction

Growing interest and heated debate surround the assertion that bilingualism, the use of more than one language, confers non-linguistic benefits on cognitive function (Antoniou, 2019; Bialystok, 2024; Paap, 2019). While non-linguistic bilingual "advantages," referred to more generally as bilingual effects (Barac & Bialystok, 2012; Privitera, Momenian, & Weekes, 2023), have been documented since at least the 1960s (Peal & Lambert, 1962), it was not until the early 1990s that attention shifted toward cognitive control (Bialystok, 1992). Cognitive control, sometimes referred to as executive function, is a set of goal directed attentional functions including updating, shifting, and inhibitory control, which are essential in controlling thought and behavior (Friedman & Miyake, 2017; Miyake et al., 2000). Proponents of the view that bilingualism impacts cognitive control argue that the experience of managing two languages, which are thought to be simultaneously active (Abutalebi & Green, 2007; Costa et al., 1999; Kroll et al., 2012; Marian & Spivey, 2003), places higher demands on this system, resulting in more efficient cognitive control with higher levels of bilingual experience (Bialystok & Craik, 2022).

Earlier theoretical accounts framed bilingual effects on cognitive control as a form of transfer. Specifically, because the cognitive demands associated with successful communication in bilinguals, such as the need to inhibit an active but unneeded language, rely on domain general inhibitory control (Green, 1998), bilingual language experience was thought to transfer to the nonverbal domain. As if follows, bilinguals were expected to demonstrate improved inhibitory control relative to monolinguals who do not experience these cognitive demands (e.g., Bialystok et al., 2004). Arguments that bilingual language experience transfers to domain-general inhibitory control soon became contentious as results across studies supported that bilinguals outperformed monolinguals on a range of measures that did not rely exclusively on inhibitory control (Hilchey & Klein, 2011), a finding initially attributed to the overlap across dimensions of cognitive control (Friedman & Miyake, 2017). Reported bilingual effects in preverbal infants raised in bilingual environments (e.g., D'Souza & D'Souza, 2021; D'Souza et al., 2020) further called into question the validity of the transfer perspective as enhanced cognitive control was observed in the absence of the very language control experience thought to be essential for its development. Recently, in light of conflicting findings but also strong evidence supporting bilingual effects on cognitive control (Grundy, 2020), Ellen Bialystok proposed a new theoretical framework arguing that bilingual experience modifies cognitive control not through transfer, but through adaptation of the underlying attention system (Bialystok, 2024). A central feature of this adaptationist account is that bilingual experience modifies the efficiency of cognitive control in a graded manner, with effects observed in both linguistic and nonlinguistic domains.

Investigations of bilingual effects on cognitive control generally administer behavioral tasks including the Simon task (Simon & Rudell, 1967), Flanker task (Eriksen & Eriksen, 1974), and Attention Network Test (ANT; Fan et al., 2002), relying on between-groups designs which compare task performance between monolingual and bilingual samples. Performance on these tasks is typically quantified by averaging reaction times (RT) or accuracy rates to reflect performance overall and under different task conditions (e.g., congruent trials), with scores for each participant and group analyzed using analysis of variance or similar statistical tests (Privitera & Weekes, 2023). Systematic reviews and meta-analyses of these previous studies have reported mixed findings, calling into question the veracity of bilingual effects on cognitive control (Donnelly et al., 2019; Giovannoli et al., 2020; Grundy, 2020; Van den Noort et al., 2019; Ware et al., 2020).

Commonplace methodological practices across previous studies may, in part, explain mixed findings reported across evidence syntheses. An almost exclusive reliance on monolingual-bilingual comparisons may create conditions under which authentic bilingual effects on cognitive control are difficult to identify. The shift toward using exclusively bilingual between-group (e.g., Grundy, 2020; Xie & Pisano, 2019) and within-group designs (e.g., Privitera et al., 2022; Privitera, Momenian, & Weekes, 2023; Xie et al., 2024) has partially addressed these issues, aligning with recent critiques that monolinguals are not an appropriate control against which to compare the performance of bilinguals (Rothman et al., 2023). Additionally, the practice of assigning categorical language status labels to participants (Luk & Bialystok, 2013), reducing monolingualism and bilingualism to uninformative, unidimensional constructs, ignores the heterogeneity inherent in language experience (De Bruin, 2019; Gullifer et al., 2021). Operationalizing bilingualism as a multidimensional, continuous variable and more comprehensively assessing and modeling differences can support identification of meaningful associations between separable dimensions of language experience and cognitive control.

Despite the progress made in addressing pervasive methodological concerns, one significant issue persists across nearly all previous studies: the assumption that task performance is stable through time. Put another way, almost everything we know about bilingual effects on cognitive control is based on evidence that presupposes that performance on a given task does not change. The statistical impact of this assumption is evidenced in the commonplace use of aggregated data during analysis, reducing the influence of differences in performance across trials within participants and of individual differences between participants (Speelman & McGann, 2013). While some studies ask participants to complete multiple blocks of trials to model changes over time (e.g., Abutalebi et al., 2012; Costa et al., 2009), data within each block are aggregated during analysis, preventing the identification of changes occurring within a single block. To address the limits associated with these methods, there is a need to identify additional measures that best reflect nontrivial differences in performance on cognitive control tasks.

The adoption of alternative measures of task performance in the investigation of bilingual effects on cognitive control is not without precedent, although their use remains extremely uncommon. Using data collected from a Flanker task, previous studies have reported significant bilingual effects on non-decision time extracted through diffusion modeling (Ong et al., 2017), incongruent trial τ component and congruent trial μ component after ex-Gaussian analysis (Abutalebi et al., 2015), and sequential congruency effects (Grundy et al., 2017). While these approaches utilized theoretically sound indices capturing underexplored dimensions of task performance, none account for changes in performance occurring over time. Accordingly, there is a need to identify measures of performance that address the limitations of aggregated data and challenge the false assumption that performance on cognitive control tasks remains stable over time.

A potential solution can be found in a recent paper from Cochrane and Green (2021) investigating the relationship between fluid intelligence and time-sensitive measures derived from a cognitive control (i.e., working memory/updating) task. Unlike conventional measures of average performance, these time-sensitive performance trajectory measures included separate indices reflecting initial and final (asymptotic) task performance and the rate of change (time taken to change) during the task (Kattner et al., 2017). The authors identified novel links between fluid intelligence and performance trajectory measures with higher fluid intelligence associated with better final task performance and faster rate of change. Additionally, while conventional performance measures were almost perfectly correlated with final task performance, no such relationship with rate of change was identified, supporting an independent link with fluid intelligence. Crucially, findings from this study support that performance on cognitive control tasks is not stable over time, and that time-sensitive measures of performance trajectory capture the dynamic nature of task performance in manner that is not confounded by aggregation of performance over time.

Extending the work of Cochrane and Green (2021), the present study explored the temporal dynamics of bilingual effects on cognitive control. Through the use of a widely-adopted cognitive control task (i.e., Simon task), we investigated the relationship between separable dimensions of language experience and both conventional measures of aggregated performance and time-sensitive measures of performance trajectory. If bilingual effects manifest as graded differences in the efficiency and deployment of cognitive control (Bialystok, 2024; Bialystok & Craik, 2022), then higher levels of bilingual experience should be associated with faster initial task performance or rate of change. However, as the impact of improved cognitive control efficiency may be diluted over successive trials, these associations may not emerge on measures of overall or final performance. Due to limited previous work investigating the impact of separable dimensions of language experience on cognitive control (e.g., Anthony & Blumenfeld, 2019; Novitskiy et al., 2019; Privitera et al., 2022; Privitera, Momenian, & Weekes, 2023; Tao et al., 2011; Xie & Pisano, 2019; Yow & Li, 2015) including the absence of any study utilizing the performance trajectory measures used in the present study, we make no specific predictions about the nature of association between language experience and performance measures. However, previous work does suggest that separate dimensions of language experience may demonstrate opposing relationships with a given measure of cognitive control (Privitera, Momenian, & Weekes, 2023), resulting in a negative association between bilingual experience and cognitive control. Although previous work has provided some guidance to inform predictions, considering that the present study is the first investigation of bilingual effects on cognitive control utilizing time-

https://doi.org/10.1017/S1366728925100576 Published online by Cambridge University Press

sensitive performance trajectory measures, the nature of our analyses should be considered exploratory.

2. Materials and methods

2.1. Participants

Fifty-seven Mandarin-English speaking bilingual young adults (39 females; M_{age} = 19.74 years, SD_{age} = 0.99 years) were recruited from a public university in a central province of Mainland China. All participants were native Mandarin (L1) speakers who had generally began learning English (L2) in primary school in accordance with the national education language policy. Written informed consent was collected from all participants before data collection. Approval for this study was granted by Huazhong University of Science and Technology. Participants received ¥20 (~$3 USD) for completing the study.

2.2. Language experience and background measures

Language experience data for Mandarin and English were collected using a bilingual version of the Language History Questionnaire (LHQ-3; Li et al., 2020). This tool consisted of a series of self-report questions designed to measure proficiency, immersion, and dominance in all languages a participant uses. Proficiency was calculated based on participants' self-reported ability to use a language, rated on a scale from 1 ("very poor") to 7 ("excellent"). Immersion was assessed using both age of acquisition (AoA) and years of use for each language. Dominance reflected the number of hours participants reported engaging in specific activities in a given language. Reported data were used to generate aggregate scores ranging from 0 to 1 for proficiency, immersion, and dominance in both Mandarin and English. Equal weight was given to the dimensions of listening, speaking, reading, and writing in the calculation of each aggregate score, and dominance scores were further weighted by reported proficiency. Aggregate scores for Mandarin and English dominance were also used to calculate an L2/L1 dominance ratio which, depending on a participant's pattern of language usage, may exceed 1. Objective English proficiency was assessed using a 25-item multiple-choice grammar and vocabulary test taken from Transparent Language (https://www.transparent.com/) with scores calculated as the proportion of correct responses (ranging from 0 to 1). The decision to use this test was based on prior validation work and its established use in assessing objective proficiency in heterogeneous samples of bilingual participants (e.g., Branzi et al., 2016; Del Maschio et al., 2022; Privitera, 2024b; Privitera, Li, et al., 2023). In total, this resulted in three measures of *L1 experience*: subjective L1 proficiency, L1 dominance, and L1 immersion, and six measures of *L2 experience*: L2 AoA, subjective L2 proficiency, objective L2 proficiency, L2 dominance, L2 immersion, and L2/L1 dominance ratio. Finally, participants reported on basic demographic details, language switching frequency, other languages used beyond Mandarin and English, and parental education level, computed as the mean of the mother's and father's education levels (ranging from 1 = elementary school to 6 = doctorate) as a proxy for socioeconomic status (SES; Wermelinger et al., 2017).

2.3. Nonverbal intelligence measure

Differences in nonverbal intelligence were assessed using an abbreviated form of Raven's Advanced Progressive Matrices (Raven, 1962). A total of 36 items from Set I were administered online using the Chinese survey website SoJump, and presented in a manner such that the difficulty progressively increased with each subsequent item. Prior to the start of the test, participants were told that they had 10 minutes to complete as many items as possible. In the event a particular item was too difficult, participants were told to make a guess and skip to the next item. The total number of items out of 36 answered correctly in 10 minutes was taken as a participant's measure of nonverbal intelligence. While these administrative procedures deviate considerably from those generally followed, abbreviated versions of Raven's Advanced Progressive Matrices have reported similar psychometric properties to the full task (e.g., Arthur et al., 1999; Hamel & Schmittmann, 2006).

2.4. Simon task

Participants completed a two-color Simon task administered online (Privitera et al., 2022). Prior to the start of the task, participants were instructed to place their left index finger on the "Q" key and their right index finger on the "P" key of their computer's keyboard. Each trial began with the presentation of a black fixation cross in the center of a white background for 300 ms. Based on the trial condition, a blue or brown square target stimulus (2.54 × 2.54 cm) was presented to the left, right, or directly over the previous location of the fixation cross. Participants were instructed to press either the "Q" or "P" button based only on the color of the target stimulus (button and color mapping counterbalanced). The combination of target stimulus color and location resulted in three trial conditions: *congruent*, where a target stimulus appeared on the same side as the correct response key, *incongruent*, where the target stimulus appeared on the opposite side, and *neutral*, where the target stimulus was presented in the center of the screen. Target stimuli remained on screen until a response was given after which a blank screen was presented for 500 ms. In total, participants completed six practice trials with feedback and 150 experimental trials without feedback. Experimental trials contained equal numbers of each possible condition and were presented in the same pseudorandomized order to ensure that participants' task experience was comparable and that any observed changes over time could not be attributed to differences in trial order.

2.5. General administration procedures

All aspects of the experiment were carried out using the Gorilla online experiment builder (Anwyl-Irvine et al., 2020) with all written content including directions and consent forms provided simultaneously in both Chinese and English. Interested participants were sent an access link through either email or WeChat along with brief instructions asking them to find a quiet place free of distractions in which to complete the experiment on a desktop computer or laptop. The experimental portal was configured in such a way that accessing the study link using either a tablet or smartphone would result in automatic rejection of the participant. After successfully entering the experimental portal, participants were asked to maximize the size of their browser screen and were again reminded not to use their phone or engage in any other distracting behavior. After providing consent, participants completed the Simon task, followed by the LHQ-3, the objective English test, and then the nonverbal intelligence assessment (link to SoJump embedded in Gorilla experiment) before debriefing. Completion of all aspects of the experiment took around 40 minutes for each participant and breaks were available after each task.

2.6. Statistical analysis

Prior to analysis, incorrect trials, and trials with RTs shorter than 150 ms and longer than 2,000 ms were trimmed in alignment with previous recommendations when investigating bilingual effects on cognition (Zhou & Krott, 2016), including cognitive control specifically (Privitera, 2024a). Initially, average overall RT and separate average RTs for each task condition were computed for participants to generate conventional measures of task performance. Performance on the Simon task was then modeled in R (Version 4.3.1; R Core Team, 2023) using the *TEbrm* function from the *TEfits* package (Cochrane, 2020), which used nonlinear Bayesian modeling in Stan using the *brms* package (Bürkner, 2017). The *TEbrm* function allows for the modeling of task performance as a continuous function of time using a maximum-likelihood model that nonlinearly relates time (e.g., trial number) to the outcome variable (e.g., RT). Participant-level task performance is summarized with three mean parameter estimates: *initial RT*, the estimated value at trial one; *final (asymptotic) RT*, the estimated value at a theoretical infinite amount of elapsed time; and *rate of change*, the base-2 log of the time taken to change halfway from initial to asymptotic values, with higher rates corresponding to longer periods of time needed to reach asymptotic performance, reflecting differences in task learning (A. Cochrane, personal communication, February 4, 2025). These three parameter estimates, which we collectively refer to as measures of *performance trajectory* (Cochrane & Green, 2021), were extracted from models for further analysis using the *coef* function in *brms*. More details about the *TEbrm* function *TEfits* package can be found on the package creator's website (https://github.com/akcochrane/TEfits). Two separate models were run to generate participant-level performance trajectory parameters for overall task performance and performance for the three different task conditions. Both models were estimated following an ex-Gaussian distribution with priors set to default, and were run with two chains for 7,000 iterations for each chain with 2,000 iterations per chain discarded as warmup, resulting in a total of 10,000 post-warmup draws. R-hat values for both models were all equal to 1, indicating that models had converged appropriately across all chains.

All additional analyses were performed in JASP (JASP Team, 2024). Pearson or Spearman two-tailed correlations were run between conventional and performance trajectory measures, language experience, and background measures. To obtain more robust estimates, 95% confidence intervals for correlation coefficients were calculated using 1,000 bootstrap replicates. Finally, the variance in task performance accounted for by language experience was tested using generalized linear models for each performance trajectory or conventional task performance measure that was significantly or marginally significantly correlated with a language experience measure. All continuous predictor variables were standardized (z-score) prior to modeling. Models each contained a single performance trajectory or conventional task performance measure as the outcome and correlated measure(s) of language experience and control background variables as predictors. Collinearity for each model was assessed using variance inflation factor (VIF), and variables with values above five were removed (Craney & Surles, 2002). The appropriateness of final models was confirmed via inspection of residual plots and fit indices including Akaike information criterion (AIC; Stoica & Selen, 2004). Robust estimates of regression coefficients and their associated 95% confidence intervals were calculated using 5,000 bootstrap replicates.

3. Results

Participant language experience and other background variables are summarized in Table 1. As expected, considerable heterogeneity in language experience was observed across participants. Of note is the observation that some participants reported learning English either much earlier or later than expected based on national educational language policy. Taking all assessed dimensions of language experience into consideration, our sample can be most accurately described as bilingual, but Mandarin dominant. Due to a technical issue in which data from completed assessments were not recorded by SoJump for unknown reasons, nonverbal intelligence data were not available for 10 participants. To minimize the impact, analyses utilized the full dataset where possible, while analyses incorporating nonverbal intelligence as a variable were conducted on the reduced dataset. Full data and analyses can be accessed on Open Science Framework (https://doi.org/10.17605/OSF.IO/H5XVG).

3.1. Simon task results

Prior to analyses, all data (150 trials) were removed for one participant who had an overall accuracy of less than 5% on the task. Data trimming resulted in the removal of 420 incorrect trials and 40 trials with RTs shorter than 150 ms and longer than 2,000 ms. Final analyses included a total of 7,940 trials from 56 participants (38 females; M_{age} = 19.73 years, SD_{age} = 1.00 year). Simon task conventional performance and performance trajectory measure results are summarized in Table 2. No analysis of accuracy data was performed due to very low error rates. Replication of the classic Simon effect (i.e., slower RTs on incongruent relative to congruent trials) was confirmed via one-tailed paired samples t-tests for conventional performance and initial and final performance trajectory measures (*ps* < .001). Finally, expected performance improvements (i.e., faster RTs on final relative to initial performance trajectory measures) were tested using one-tailed paired samples t-tests, with significant improvements observed for overall,

Table 1. Demographic and language history data of final sample (n = 56)

	M	SD	Range
Age (years)	19.73	1,00	18–22
Socioeconomic status (1–6 points)	2.87	1.01	1–5
Nonverbal intelligence (36 points)	20.11	6.41	4–30
Number of languages used	2.16	0.63	1–4
Frequency of language switching (1–7 points)	3.16	1.65	1–7
L2 AoA (years)	8.52	2.92	0–17
L1 proficiency SUB (0–1 point)	.76	.10	.54–.96
L1 immersion (0–1 point)	.90	.07	.64–.97
L1 dominance (0–1 point)	.55	.18	.33–1.00
L2 proficiency SUB (0–1 point)	.50	.11	.25–.71
L2 proficiency OBJ (0–1 point)	.82	.12	.20–1.00
L2 immersion (0–1 point)	.57	.12	.35–.86
L2 dominance (0–1 point)	.29	.07	.15–.47
L2/L1 dominance ratio	0.55	0.16	0.26–0.86

Table 2. Summary of Simon task performance by item congruency condition

		Item congruency		
	Overall	Congruent	Incongruent	Neutral
RT (ms)	478 (71)	458 (73)	506 (77)	474 (72)
Accuracy (%)	.95 (.04)	.98 (.02)	.90 (.09)	.97 (.04)
Initial RT (ms)	505 (80)	462 (72)	560 (79)	505 (73)
Final RT (ms)	477 (35)	459 (34)	499 (36)	469 (35)
Rate of change	3.00 (1.25)	3.39 (2.24)	3.09 (2.34)	2.93 (2.26)

Note: Rate of change corresponds to the [base-2] log of the time taken to change halfway from the initial to the final (asymptotic) RT. A higher rate of change corresponds to a longer period of time needed to reach asymptotic performance.

neutral, and incongruent trial performance (*ps* ≤ .008), but not congruent trial performance (p = .211).

3.1.1. Correlation results

Conventional and performance trajectory measures. Correlation analysis results are presented in Table 3. All conventional performance measures, representing average RT, were significantly positively correlated with their respective initial and final RT performance trajectory measures. Final RT measures were all significantly positively correlated with their respective rate of change measures, indicating that slower learners tended to have worse asymptotic performance. Significant positive correlations were also identified between average incongruent and neutral trial RT and their respective rate of change measures, indicating that slower learners also tended to have worse overall performance on incongruent and neutral trials. Finally, no initial or final RT performance trajectory measures were significantly correlated with each other, supporting that these measures are independent of one another.

Table 3. Correlations between conventional and performance trajectory measures

Variable	1	2	3	4
1. OA AVG	–			
2. OA initial	**.53*** [.27, .70]**	–		
3. OA final	**.72*** [.59, .85]**	.14 [−.14, .40]	–	
4. OA RoC[a]	.20 [−.11, .47]	−.25 [−.55, .06]	**.31* [.07, .52]**	–
Variable	**1**	**2**	**3**	**4**
1. CON AVG	–			
2. CON initial	**.57*** [.33, .72]**	–		
3. CON final	**.71*** [.57, .83]**	.20 [−.07, .44]	–	
4. CON RoC[a]	.21 [−.07, .50]	.01 [−.24, .27]	**.32* [.03, .40]**	–
Variable	**1**	**2**	**3**	**4**
1. INC AVG	–			
2. INC initial	**.43*** [.20, .64]**	–		
3. INC final	**.76*** [.64, .85]**	.16 [−.07, .40]	–	
4. INC RoC[a]	**.46*** [.16, .68]**	.06 [−.22, .31]	**.37** [.10, .56]**	–
Variable	**1**	**2**	**3**	**4**
1. NEU AVG	–			
2. NEU initial	**.59*** [.37, .74]**	–		
3. NEU final	**.68*** [.56, .80]**	.20 [−.02, .42]	–	
4. NEU RoC[a]	**.39** [.12, .61]**	.03 [−.25, .28]	**.34* [.08, .56]**	–

Note. Values in square brackets represent 95% confidence intervals for each correlation based on 1,000 bootstrapped replicates. All correlations were calculated with df = 54. OA: Overall, AVG: Average, CON: Congruent, RoC: Rate of change, INC: Incongruent, NEU: Neutral. Bold values indicate statistically significant results: * = $p \leq .05$, ** = $p \leq .01$, *** = $p \leq .001$.
[a]Symbol identifies variables that were non-normally distributed, with results generated using Spearman correlations.

Language experience and background measures. Correlation analysis results are presented in Table 4. Age was significantly positively correlated with L2/L1 dominance ratio and marginally significantly correlated with L2 subjective proficiency and L2 dominance, supporting that older participants tended to use more English relative to Mandarin and demonstrate higher proficiency. This relationship was further evidenced in significant positive correlations between L2/L1 dominance ratio and L2 subjective and objective proficiency and L2 immersion, as well as L2 dominance and L2 subjective and objective proficiency. L2 AoA was significantly negatively correlated with nonverbal intelligence and marginally significantly negatively correlated with SES, indicating that those with began learning English earlier tended to be of higher status with higher nonverbal intelligence. There was a significant positive correlation between the number of languages used and frequency of language switching, supporting that those using a higher number of languages were more likely to engage in switching. Language switching was also significantly positively correlated with all measures of L2 experience and negatively correlated with L2 AoA, supporting that participants with higher levels of bilingualism tended to engage in more switching. While a significant positive correlation was identified between L2 subjective and objective proficiency, L1 subjective proficiency was only significantly positively correlated with L2 subjective proficiency. Finally, no significant correlation was identified between SES and nonverbal intelligence.

Performance measures and language experience and background measures. Full correlation analysis results are presented in Table 5. L2 AoA was significantly negatively associated with average overall and congruent trial RT, congruent, incongruent, and neutral trial initial RT, and marginally significantly negatively correlated with overall initial and average neutral trial RT. A nearly identical pattern of results, albeit with positive correlations, was observed for L2 immersion, an aggregate score that gives equal weight to both AoA and years of use. Specifically, L2 immersion was significantly positively correlated with average overall and congruent trial RT, overall, congruent, incongruent, and neutral initial trial RT, and marginally significantly positively correlated with average incongruent and neutral trial RT. Across all rate of change measurements only one significant correlation was observed: a negative correlation between overall rate of change and SES, supporting that those with higher status tended to demonstrate faster learning on the Simon task. Overall rate of change was also marginally significantly negatively correlated with L2/L1 dominance ratio, indicating that participants who used more English relative to Mandarin tended to demonstrate faster learning. L1 subjective proficiency was significantly positively correlated with congruent trial final RT, and marginally significantly positively correlated with incongruent and neutral trial final RT, indicating that those with higher Mandarin proficiency tended to reach lower levels of performance across the three Simon task trial conditions. Finally, the number of languages used was marginally significantly negatively correlated with overall, incongruent, and neutral trial average RT, but the

https://doi.org/10.1017/S1366728925100576 Published online by Cambridge University Press

Table 4. Correlations between language experience and background measures

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Age[a]	–													
2. SES	−.15 [−.43, .14]	–												
3. NV-IQ	.07 [−.22, .37]	.15 [−.11, .39]	–											
4. NUM Lang[a]	.18 [−.11, .46]	−.01 [−.26, .25]	.02 [−.22, .26]	–										
5. Switching[a]	.20 [−.05, .44]	.07 [−.20, .32]	−.01 [−.29, .29]	**.48*** [.24, .67]**	–									
6. L2 AoA	−.10 [−.36, .19]	−.26$^{\wedge}$ [−.48, .03]	**−.31* [−.55, −.02]**	−.11 [−.38, .15]	**−.32* [−.55, −.07]**	–								
7. L1 PRO$_S$	.02 [−.22, .25]	.08 [−.13, .30]	−.02 [−.30, .27]	−.06 [−.33, .23]	.04 [−.23, .31]	**−.27* [−.52, .00]**	–							
8. L1 IMM[a]	−.07 [−.35, .19]	.23$^{\wedge}$ [−.01, .45]	−.01 [−.32, .30]	−.07 [−.35, .24]	.13 [−.13, .37]	−.10 [−.37, .19]	**.53*** [.30, .71]**	–						
9. L1 DOM[a]	−.14 [−.39, .13]	.07 [−.21, .33]	−.10 [−.41, .24]	−.03 [−.29, .22]	.08 [−.18, .32]	.04 [−.24, .32]	**.46*** [.23, .65]**	**.45*** [.20, .67]**	–					
10. L2 PRO$_S$	.22$^{\wedge}$ [−.02, .46]	.15 [−.07, .36]	.09 [−.21, .40]	**.31* [.03, .54]**	**.46*** [.20, .68]**	−.18 [−.44, .07]	**.45*** [.23, .62]**	.24$^{\wedge}$ [−.01, .45]	**.28* [.05, .49]**	–				
11. L2 PRO$_O$[a]	.16 [−.10, .40]	.14 [−.11, .39]	−.09 [−.38, .20]	**.30* [.04, .52]**	**.35** [.09, .57]**	−.25$^{\wedge}$ [−.48, .00]	.01 [−.26, .27]	.08 [−.19, .35]	−.07 [−.31, .19]	**.31* [.03, .56]**	–			
12. L2 IMM	.32 [.06, .54]	.19 [−.09, .42]	**.30* [.02, .56]**	.12 [−.13, .35]	**.26* [.02, .49]**	**−.84*** [−.95, −.68]**	.18 [−.10, .44]	.06 [−.23, .33]	−.08 [−.34, .20]	.26$^{\wedge}$ [.01, .47]	.20 [−.07, .44]	–		
13. L2 DOM	.25$^{\wedge}$ [.02, .45]	.13 [−.10, .35]	−.06 [−.34, .24]	**.27* [−.02, .50]**	**.60*** [.36, .77]**	−.16 [−.44, .10]	**.34* [.13, .52]**	.24$^{\wedge}$ [−.00, .47]	**.39** [.17, .56]**	**.90*** [.83, .95]**	**.28* [.01, .54]**	.25$^{\wedge}$ [.01, .48]	–	
14. L2/L1 DOM	**.34** [.06, .58]**	.07 [−.23, .40]	.06 [−.28, .35]	**.32* [.06, .55]**	**.43*** [.20, .63]**	−.06 [−.44, .27]	−.11 [−.36, .12]	−.15 [−.41, .12]	**−.56*** [−.74, −.32]**	**.52*** [.32, .69]**	**.35** [.10, .59]**	**.27* [−.01, .52]**	**.48*** [.26, .67]**	–

Note: Values in square brackets represent 95% confidence intervals for each correlation based on 1,000 bootstrapped replicates. All correlations were calculated with df = 54 except correlations with the variable NV-IQ which were calculated with df = 45 due to missing data. SES: socioeconomic status, NV-IQ: nonverbal intelligence, NUM Lang: number of languages used, L2: second language, AoA: age of acquisition, PRO$_S$: subjective proficiency, L1: native language, IMM: immersion, DOM: Dominance, PRO$_O$: objective proficiency. Bold values indicate statistically significant results: ^ = $p \le .10$ * = $p \le .05$, ** = $p \le .01$, *** = $p \le .001$.

[a]Identifies variables that were non-normally distributed, with results generated using Spearman correlations.

https://doi.org/10.1017/S1366728925100576 Published online by Cambridge University Press

Table 5. Correlations between performance, language experience, and background measures

Variable	Age[a]	SES	NV-IQ	NUM lang[a]	Switching[a]	L2 AoA	L1 PRO_S	L1 IMM[a]	L1 DOM[a]	L2 PRO_S	L2 PRO_O[a]	L2 IMM	L2 DOM	L2/L1 DOM
OA AVG	−.03 [−.25, .21]	−.07 [−.30, .16]	.05 [−.28, .35]	−.23$^{\wedge}$ [−.43, −.01]	−.17 [−.43, .09]	**−.28* [−.50, −.06]**	.20 [−.04, .42]	.02 [−.26, .29]	.03 [−.27, .30]	.06 [−.20, .30]	.00 [−.25, .25]	**.27* [−.01, .51]**	−.02 [−.29, .24]	−.02 [−.28, .26]
OA initial	−.05 [−.31, .22]	−.02 [−.26, .23]	−.02 [−.45, .34]	−.08 [−.31, .18]	.00 [−.23, .25]	−.23$^{\wedge}$ [−.47, .02]	−.03 [−.28, .19]	−.12 [−.37, .14]	−.04 [−.31, .22]	−.02 [−.29, .23]	.16 [−.09, .39]	**.36** [.10, .57]**	−.06 [−.31, .19]	.07 [−.22, .34]
OA final	.15 [−.09, .40]	−.17 [−.41, .10]	−.11 [−.36, .15]	−.10 [−.36, .16]	−.15 [−.42, .13]	−.16 [−.44, .13]	.22 [−.04, .47]	.01 [−.27, .29]	.05 [−.23, .32]	−.01 [−.28, .27]	.00 [−.23, .25]	.12 [−.19, .39]	−.07 [−.34, .22]	−.02 [−.28, .27]
OA RoC[a]	−.07 [−.35, .19]	**−.40** [−.60, −.15]**	.05 [−.24, .33]	−.05 [−.30, .20]	−.15 [−.39, .14]	.17 [−.09, .44]	.04 [−.21, .27]	.08 [−.16, .33]	.18 [−.09, .40]	.01 [−.26, .32]	−.21 [−.43, .05]	−.16 [−.44, .12]	−.05 [−.32, .23]	−.23$^{\wedge}$ [−.46, .05]
CON AVG	−.07 [−.32, .20]	−.06 [−.28, .17]	.05 [−.30, .36]	−.20 [−.42, .03]	−.15 [−.42, .14]	**−.34* [−.55, −.10]**	.26 [.02, .47]	.03 [−.25, .30]	.04 [−.25, .31]	.11 [−.15, .34]	−.01 [−.24, .24]	**.30* [.01, .53]**	.00 [−.25, .25]	−.01 [−.25, .27]
CON initial	−.10 [−.34, .16]	−.07 [−.32, .17]	.01 [−.44, .38]	−.03 [−.26, .21]	.07 [−.20, .31]	**−.31* [−.50, −.13]**	.01 [−.24, .23]	−.15 [−.39, .09]	−.08 [−.35, .20]	.04 [−.21, .28]	.13 [−.12, .37]	**.40** [.15, .58]**	−.03 [−.27, .20]	.07 [−.20, .34]
CON final	−.20 [−.04, .44]	−.04 [−.28, .21]	−.06 [−.31, .19]	−.08 [−.35, .20]	−.17 [−.42, .11]	−.18 [−.50, .12]	**.27* [−.01, .52]**	.02 [−.26, .31]	.05 [−.22, .32]	−.03 [−.30, .25]	−.02 [−.26, .22]	.16 [−.17, .47]	−.10 [−.37, .21]	−.06 [−.31, .24]
CON RoC[a]	.01 [−.26, .29]	−.03 [−.30, .24]	−.17 [−.45, .14]	−.22 [−.45, .06]	−.12 [−.40, .15]	.14 [−.11, .39]	−.11 [−.35, .14]	−.06 [−.32, .19]	.11 [−.23, .38]	−.08 [−.33, .17]	−.01 [−.27, .25]	−.05 [−.32, .21]	−.05 [−.32, .21]	−.07 [−.37, .20]
INC AVG	.04 [−.20, .27]	−.05 [−.30, .20]	.04 [−.28, .33]	−.25$^{\wedge}$ [−.44, −.04]	−.17 [−.42, .10]	−.21 [−.46, .02]	.15 [−.09, .38]	.02 [−.24, .29]	.00 [−.28, .28]	.01 [−.24, .27]	−.01 [−.28, .24]	.23$^{\wedge}$ [−.03, .47]	−.04 [−.32, .25]	.00 [−.25, .26]
INC initial	−.09 [−.34, .18]	−.05 [−.29, .19]	−.01 [−.45, .36]	−.06 [−.28, .18]	.06 [−.22, .31]	**−.31* [−.50, −.13]**	.02 [−.23, .23]	−.14 [−.39, .11]	−.10 [−.37, .17]	.03 [−.23, .29]	.13 [−.12, .38]	**.40** [.16, .59]**	−.04 [−.28, .19]	.07 [−.20, .35]
INC final	−22$^{\wedge}$ [−.01 .45]	−.03 [−.27, .22]	−.03 [−.30, .22]	−.11 [−.37, .18]	−.17 [−.41, .13]	−.16 [−.48, .14]	.24$^{\wedge}$ [−.03, .50]	.00 [−.27, .29]	.05 [−.24, .32]	−.02 [−.29, .27]	−.02 [−.25, .23]	.15 [−.17, .44]	−.09 [−.36, .23]	−.03 [−.28, .26]
INC RoC[a]	−01 [−.27, .29]	−.03 [−.31, .26]	−.19 [−.44, .11]	−.21 [−.45, .08]	−.11 [−.39, .15]	.16 [−.12, .40]	−.12 [−.35, .13]	−.04 [−.30, .21]	.11 [−.23, .39]	−.06 [−.31, .20]	.01 [−.25, .27]	−.06 [−.30, .20]	−.02 [−.28, .23]	−.04 [−.32, .26]
NEU AVG	−02 [−.25, .21]	−.11 [−.34, .14]	.04 [−.29, .33]	−.26$^{\wedge}$ [−.45, −.05]	−.18 [−.42, .08]	−.23$^{\wedge}$ [−.47, −.01]	.16 [−.08, .37]	.03 [−.25, .30]	.06 [−.23, .34]	.05 [−.22, .29]	.01 [−.25, .27]	.24$^{\wedge}$ [−.04, .48]	−.03 [−.29, .22]	−.06 [−.31, .24]
NEU initial	−.10 [−.33, .16]	−.09 [−.33, .16]	.01 [−.44, .39]	−.05 [−.28, .18]	.05 [−.21, .30]	**−.29* [−.49, −.10]**	.00 [−.24, .22]	−.19 [−.42, .06]	−.08 [−.35, .19]	.03 [−.23, .27]	.11 [−.15, .35]	**.38** [.13, .57]**	−.05 [−.28, .19]	.07 [−.21, .35]
NEU final	.20 [−.05, .45]	−.05 [−.29, .20]	−.05 [−.31, .20]	−.09 [−.35, .19]	−.17 [−.42, .11]	−.18 [−.50, .12]	.26$^{\wedge}$ [−.03, .51]	.04 [−.25, .33]	.08 [−.19, .34]	−.03 [−.29, .26]	−.01 [−.25, .24]	.17 [−.18, .47]	−.09 [−.36, .22]	−.06 [−.30, .24]
NEU RoC[a]	.01 [−.26, .29]	−.04 [−.31, .24]	−.17 [−.44, .14]	−.22 [−.45, .06]	−.12 [−.40, .15]	.13 [−.14, .38]	−.11 [−.33, .12]	−.08 [−.34, .19]	.11 [−.23, .37]	−.08 [−.32, .17]	−.01 [−.27, .25]	−.04 [−.29, .21]	−.06 [−.31, .21]	−.08 [−.37, .23]

Note: Values in square brackets represent 95% confidence intervals for each correlation based on 1,000 bootstrapped replicates. All correlations were calculated with df= 54 except correlations with the variable NV-IQ which were calculated with df=45 due to missing data. SES: Socioeconomic status, NV-IQ: Nonverbal intelligence, NUM Lang: Number of languages used, L2: Second language, AoA: Age of acquisition, PRO_S: Subjective proficiency, L1: Native language, IMM: Immersion, DOM: Dominance, PRO_O: Objective proficiency, OA: Overall, AVG: Average, CON: Congruent, RoC: Rate of change, INC: Incongruent, NEU: Neutral. Bold values indicate statistically significant results: ^ = $p \le .10$ * = $p \le .05$, ** = $p \le .01$, *** = $p \le .001$.

[a]Identifies variables that were non-normally distributed, with results generated using Spearman correlations.

correlation with average congruent trial RT was not significant (p = .143).

3.1.2. Regression results

Informed by the results of the correlation analysis, separate generalized linear models were constructed for each conventional and performance trajectory measure that was either significantly or marginally significantly correlated with a measure of language experience. As summarized in Table 5, L2 AoA, L2 immersion, and L1 proficiency were all significantly correlated with Simon task performance while number of languages used and L2/L1 dominance ratio were marginally significantly correlated. Given the strong significant correlation between L2 AoA and L2 immersion (Table 4), the observation that all performance measures correlated with L2 AoA were also correlated with L2 immersion, and the more informative nature of the L2 immersion measure which gives equal weight to both AoA and years of use, the decision was made to exclude L2 AoA from models in favor of the L2 immersion measure. Finally, due to the observed significant correlation between SES and Simon task performance (Table 5), and the importance of accounting for differences in SES when investigating the impact of bilingualism on cognitive control (Morton & Harper, 2007; Naeem et al., 2018; Xie & Pisano, 2019), SES was included as a control variable on the first step of all models. Accordingly, models containing SES and language experience variables will be compared against null models containing only SES, not intercept only models. Nonverbal intelligence was not included in any model due to the absence of any significant or marginally significant correlations with conventional or performance trajectory measures (Table 5).

Conventional performance measures. Full analysis results are presented in Table 6. The model for average overall RT including SES alone was not significant, R^2 = .01, F(1, 54) = 0.279, p = .600, explaining 1% of the variance in performance. The addition of L2 immersion and number of languages used on the second step resulted in a significantly improved model, ΔF(2, 52) = 3.995, p = .024, explaining an additional 13% of variance in performance, although the overall model was marginally significant, R^2 = .14, F (3, 52) = 2.767, p = .051. The model for average congruent trial RT with SES alone was not significant, R^2 = .00, F(1, 54) = 0.184, p = .670, and explained less than 1% of the variance in performance. The addition of L2 immersion on the second step resulted in a significantly improved model, ΔF(1, 53) = 5.914, p = .018, explaining an additional 10% of variance in performance, but the overall model was marginally significant, R^2 = .10, F(2, 53) = 3.058, p = .055. The model for average incongruent trial RT with SES alone was not significant, R^2 = .00, F(1, 54) = 0.132, p = .718, and the addition of L2 immersion and number of languages used on the second step marginally improved the model, ΔF(2, 52) = 3.057, p = .056, with the final model accounting for 11% of variance in performance, but the overall model was not significant, R^2 = .11, F (3, 52) = 2.086, p = .113. Finally, the model for average neutral trial RT with SES alone was not significant, R^2 = .01, F(1, 54) = 0.611, p = .438, and explained 1% of the variance in performance. The addition of L2 immersion and number of languages used on the second step resulted in a significantly improved model, ΔF (2, 52) = 4.029, p = .024, explaining an additional 13% of variance in performance, R^2 = .14, F(3, 52) = 2.913, p = .043. To summarize, although only the overall model for average neutral trial RT was significant, the presence of significant coefficients supports that higher levels of L2 immersion were generally associated with slower average performance across all trial types (marginally significant coefficient for incongruent trials) while higher numbers of languages used were associated with faster average performance on all but congruent trials.

Table 6. Regression analysis results for conventional performance measures

Average overall RT (p = .051)				
Effect	Estimate	SE	95% CI	p
Intercept	478.660	9.112	460.45, 495.73	**<.001**
SES	−9.524	8.701	−27.19, 6.85	.232
NUM lang	−15.660	7.791	−32.03, −2.31	**.022**
L2 IMM	22.221	10.560	4.04, 45.45	**.020**
Average congruent RT (p = .055)				
Effect	Estimate	SE	95% CI	p
Intercept	458.162	9.396	439.86, 477.21	**<.001**
SES	−8.499	8.852	−27.09, 8.03	.284
L2 IMM	23.672	11.074	2.37, 46.08	**.030**
Average incongruent RT (p = .113)				
Effect	Estimate	SE	95% CI	p
Intercept	506.695	9.845	487.28, 525.66	**<.001**
SES	−8.202	9.671	−26.42, 11.8	.405
NUM lang	−16.498	8.960	−37.02, −1.06	**.037**
L2 IMM	21.088	11.004	−0.31, 42.64	.054
Average neutral RT (***p* = .043**)				
Effect	Estimate	SE	95% CI	p
Intercept	474.119	9.043	457.3, 492.86	**<.001**
SES	−12.186	8.974	−29.18, 6.39	.176
NUM lang	−18.106	7.893	−35.09, −4.53	**.008**
L2 IMM	20.986	10.510	2.03, 44.61	**.031**

Note: P-values reported next to each model represent results of overall model ANOVA. Coefficient estimates are based on the median of bootstrap distributions after 5,000 replicates. All models met the assumptions for linear regression. SE: Standard error, CI: Confidence interval, SES: Socioeconomics status, NUM Lang: Number of languages used, L2 IMM: Second language immersion. Bold values indicate statistically significant results.

Performance trajectory: Initial RT. Full analysis results are presented in Table 7. The model for initial overall RT including SES alone was not significant, R^2 = .00, F(1, 54) = 0.025, p = .874, explaining less than 1% of the variance in performance. The addition of L2 immersion on the second step resulted in a significantly improved model, ΔF(1, 53) = 8.644, p = .005, explaining an additional 14% of variance in performance, R^2 = .14, F(2, 53) = 4.336, p = .018. The model for initial congruent RT including SES alone was not significant, R^2 = .01, F(1, 54) = 0.291, p = .592, explaining 1% of the variance in performance. The addition of L2 immersion on the second step resulted in a significantly improved model, ΔF (1, 53) = 11.221, p = .001, explaining an additional 17% of variance in performance, R^2 = .18, F(2, 53) = 5.783, p = .005. The model for initial incongruent RT including SES alone was not significant, R^2 = .00, F(1, 54) = 0.151, p = .699, explaining less than 1% of the variance in performance. The addition of L2 immersion on the second step resulted in a significantly improved model, ΔF (1, 53) = 11.164, p = .002, explaining an additional 17% of variance

https://doi.org/10.1017/S1366728925100576 Published online by Cambridge University Press

Table 7. Regression analysis results for performance trajectory measures

Initial overall RT (***p* = .018**)

Effect	Estimate	SE	95% CI	p
Intercept	504.708	9.967	485.00, 524.36	**<.001**
SES	−7.179	10.008	−29.49, 10.26	.410
L2 IMM	29.810	12.284	8.44, 56.80	**.006**

Initial congruent RT (***p* = .005**)

Effect	Estimate	SE	95% CI	*p*
Intercept	462.408	8.783	444.92, 479.20	**<.001**
SES	−10.914	9.157	−29.73, 6.16	.212
L2 IMM	30.292	10.914	12.34, 55.15	**.001**

Initial incongruent RT (***p* = .006**)

Effect	Estimate	SE	95% CI	*p*
Intercept	560.406	9.700	540.65, 578.75	**<.001**
SES	−10.353	10.222	−30.58, 9.60	.290
L2 IMM	32.949	11.748	13.09, 59.64	**<.001**

Initial neutral RT (***p* = .007**)

Effect	Estimate	SE	95% CI	*p*
Intercept	505.139	8.853	487.06, 521.81	**<.001**
SES	−11.843	9.185	−30.60, 5.68	.187
L2 IMM	29.402	11.302	10.72, 56.09	**.001**

Final congruent RT (*p* = .129)

Effect	Estimate	SE	95% CI	p
Intercept	458.712	4.416	449.83, 467.30	**<.001**
SES	−2.292	4.359	−10.60, 6.45	.611
L1 PRO	9.180	4.900	−0.40, 18.91	.060

Final incongruent RT (*p* = .141)

Effect	Estimate	SE	95% CI	*p*
Intercept	499.244	4.605	489.99, 507.79	**<.001**
SES	−0.803	4.193	−9.01, 7.58	.882
Age	6.726	3.908	−0.10, 15.48	.054
L1 PRO	8.832	5.076	−1.60, 18.51	.095

Final neutral RT (*p* = .140)

Effect	Estimate	SE	95% CI	*p*
Intercept	469.511	4.547	460.23, 477.99	**<.001**
SES	−2.794	4.359	−10.96, 6.41	.592
L1 PRO	9.118	5.001	−1.14, 18.56	.082

Overall rate of change (***p* < .001**)

Effect	Estimate	SE	95% CI	p
Intercept	1.024	0.045	0.94, 1.11	**<.001**
SES	−0.142	0.042	−0.23, −0.06	**.002**
L2/L1 DOM	−0.108	0.053	−0.22, −0.01	**.026**

(*Continued*)

Table 7. (*Continued*)

Note: P-values reported next to each model represent results of overall model ANOVA. Coefficient estimates are based on the median of bootstrap distributions after 5,000 replicates. All models met the assumptions for linear regression. SE: Standard error, CI: Confidence interval, SES: Socioeconomics status, NUM Lang: Number of languages used, L2 IMM: Second language immersion. Bold values indicate statistically significant results.

in performance, R^2 = .18, $F(2, 53)$ = 5.672, p = .006. Finally, the model for initial neutral RT including SES alone was not significant, R^2 = .01, $F(1, 54)$ = 0.396, p = .532, explaining 1% of the variance in performance. The addition of L2 immersion on the second step resulted in a significantly improved model, $\Delta F(1, 53)$ = 10.375, p = .002, explaining an additional 16% of variance in performance, R^2 = .17, $F(2, 53)$ = 5.420, p = .007. While controlling for SES, higher levels of L2 immersion were associated with slower performance on all measures of initial performance.

Performance trajectory: Final RT. Full analysis results are presented in Table 7. The model for final congruent trial RT including SES alone was not significant, R^2 = .00, $F(1, 54)$ = 0.102, p = .750, explaining less than 1% of the variance in performance. The addition of L1 proficiency on the second step resulted in a significantly improved model, $\Delta F(1, 53)$ = 4.155, p = .047, explaining an additional 7% of variance in performance, but the overall model was not significant, R^2 = .07, $F(2, 53)$ = 2.132, p = .129. The model for final incongruent trial RT including SES alone was not significant, R^2 = .00, $F(1, 54)$ = 0.045, p = .833, explaining less than 1% of the variance in performance. The addition of L1 proficiency and age on the second step resulted in a marginally significantly improved model, $\Delta F(2, 52)$ = 2.822, p = .069, explaining an additional 10% of variance in performance, but the overall model was not significant, R^2 = .10, $F(3, 52)$ = 1.897, p = .141. Finally, the model for final neutral trial RT including SES alone was not significant, R^2 = .00, $F(1, 54)$ = 0.143, p = .707, explaining less than 1% of the variance in performance. The addition of L1 proficiency on the second step resulted in a marginally improved model, $\Delta F(1, 53)$ = 3.931, p = .053, explaining an additional 7% of variance in performance, R^2 = .07, $F(2, 53)$ = 2.041, p = .140. Together, the lack of significant overall model tests and coefficients for predictors within each model suggests that these predictors do not have a meaningful relationship with final task performance.

Performance trajectory: Rate of change. Full analysis results are presented in Table 7. Pre-modeling data checking identified the presence of outliers in our dataset, which were successfully addressed through log transformation. The model for overall rate of change including SES alone was significant, R^2 = .15, $F(1, 54)$ = 9.615, p = .003, explaining 15% of the variance in performance. The addition of L2/L1 dominance ratio on the second step resulted in a significantly improved model, $\Delta F(1, 53)$ = 5.470, p = .023, explaining an additional 8% of variance in performance, R^2 = .23, $F(2, 53)$ = 7.940, p < .001. While controlling for SES, higher usage of English relative to Mandarin was associated with faster overall learning.

https://doi.org/10.1017/S1366728925100576 Published online by Cambridge University Press

4. Discussion

In a significant departure from all previous work, the present study investigated the relationship between language experience and Simon task performance modeled as a continuous function of time. The novelty of this study lies in the use of behavioral measures that do not assume stable task performance, therefore treating the Simon task as a learning task. We identified distinct patterns of results across both conventional performance and performance trajectory measures while controlling for SES, with each supporting a different relationship between bilingualism and cognitive control. While differences in some aspects of language experience, most reliably L2 immersion, were generally predictive of average and initial task performance, no significant associations with final task performance were identified. Additionally, L2/L1 dominance ratio was the only variable predictive of overall rate of change on the Simon task. Taken together, our findings provide additional support for the idea that different aspects of bilingual experience are predictive of different dimensions of cognitive control (Yurtsever et al., 2024), extending these findings into the domain of performance trajectory measures (Cochrane & Green, 2021).

4.1. Trajectory measures capture unique aspects of performance

Correlations between conventional performance measures and their respective measures of initial and final performance suggest that these separate indices reflect similar aspects of performance, but the absence of significant correlations between measures of initial and final performance support that these measures are independent. We interpret this pattern of results as support for initial and final performance measures reflecting distinct and independent aspects of performance at different stages, with conventional performance measures (i.e., average) reflecting an integrative measure of these independent contributions. Significant positive correlations between rate of change and final but not initial performance further supports the independence of these measures, suggesting that heterogeneity in learning is more strongly exhibited in final performance. Significant positive correlations for rate of change and conventional performance measures were only observed for incongruent and neutral trials, suggesting that overall performance on congruent trials is less sensitive to differences in learning. This interpretation is further supported by the absence of a significant reduction in RT between initial and final performance measures, suggesting that congruent trial performance may begin at near ceiling levels, potentially due to their faciliatory influence (Simon & Rudell, 1967).

4.2. Distinct relationships of L2 immersion and multilingualism with task performance

Although models for overall, congruent, and incongruent trial RT did not significantly explain variance in task performance, L2 immersion and number of languages used exhibited meaningful relationships with all measures after accounting for differences in SES. Specifically, higher levels of L2 immersion, reflecting both earlier AoA and more years of L2 use (Li et al., 2020), were generally associated with slower average performance across all trial types while higher numbers of languages used were associated with faster average performance on all but congruent trials. Similarly, the same association between L2 immersion and performance was observed for initial performance measures, although all models were significant and no meaningful relationship was identified for number of languages used. This pattern of results highlights that the relationship between bilingual experience and cognitive control can differ significantly depending on the specific dimension of language experience under investigation (Yurtsever et al., 2024).

Although negative associations between cognitive control and bilingualism in general (reviewed in Lehtonen et al., 2018), and L2 immersion specifically (Privitera, Momenian, & Weekes, 2023), have been reported, previous work has generally identified either positive (e.g., Luk et al., 2011; Yow & Li, 2015) or null associations (e.g., Pelham & Abrams, 2014) with the highly related measure of AoA. One explanation for our observed findings relates to the characteristics of our sample: Mandarin-English bilinguals who were Mandarin dominant, had lower English proficiency, and lived in a Mandarin-immersive environment. Positive bilingual effects on cognitive control have been reported in late bilingual young adults (Bak et al., 2014; Luk et al., 2011), but studies of this kind generally recruit samples who are highly proficient in their L2 and living in L2 dominant environments (e.g., international students studying abroad), conditions under which significant language control demands are exerted (Green & Abutalebi, 2013). Taking the characteristics of our sample into consideration, along with previous reports of a similar relationship in a linguistically comparable but slightly younger sample (Privitera, Momenian, & Weekes, 2023), we interpret the negative association between L2 immersion and cognitive control as reflecting the atypical environment our study was conducted in which did not necessitate proficient, habitual use of English. This interpretation aligns with prior empirical findings that emphasize the role of linguistic context in accounting for conflicting results regarding the relationship between bilingual experience and cognitive control (e.g., Woumans et al., 2015), as well as broader theoretical frameworks emphasizing the influence of interactional context (Green & Abutalebi, 2013). Further work is needed to more clearly identify factors that modulate the relationship between L2 immersion and cognitive control.

Relative to managing two languages, the use of three or more can be considered a more intensive form of language experience, which would be expected to exert a more pronounced influence on cognitive control (Bialystok, 2024). Support for this prediction remains limited, in part because studies often group individuals who use more than one language into a single 'bilingual' category, obscuring distinctions between bilingual and multilingual experiences (e.g., Bialystok et al., 2014). Previous empirical studies have reported mixed results, with participants using three or more languages demonstrating superior performance on only some dimensions of cognitive control compared to bilinguals (e.g., Durand López, 2021; Madrazo & Bernardo, 2018; Poarch & Van Hell, 2012; reviewed in Schroeder & Marian, 2017). In one previous study conducted using a linguistically similar sample, Privitera, Momenian, and Weekes (2023) observed a main effect of number of languages used on the Attention Network Test (ANT; Fan et al., 2002), with higher numbers of languages associated with better cognitive control, manifesting as faster RTs on all trial types. While our pattern of results would support a similar conclusion, the absence of significant overall models for all but average neutral trial performance force us to interpret this finding with caution. The strength in our conclusions is further tempered due to the use of a measure of multilingualism that is limited in scope, only asking participants to indicate what languages they used beyond Mandarin and English. Although lacking granularity, the measure of multilingualism used in the present study is similar to those used in previous investigations of the impact of speaking more than two languages on cognitive control (e.g., Durand López, 2021;

Guðmundsdóttir & Lesk, 2019; Poarch & Van Hell, 2012). Future investigations should consider including more detailed assessments of multilingualism to better understand the nature of this result.

4.3. No bilingual effects associated with final performance measures

In stark contrast with conventional and initial performance measures, no language experience variables exhibited a meaningful relationship with final performance measures after accounting for differences in SES. Differences between the pattern of results observed for initial and final performance measures suggests that bilingual effects on cognitive control may emerge more readily toward the beginning of the task or, conversely, when using tasks with fewer experimental trials. Aligning with this interpretation, comparisons between bilingual and monolingual samples support a negative association between the number of experimental trials and the magnitude of bilingual effects on cognitive control in adults (reviewed in Hilchey et al., 2015 ; Hilchey & Klein, 2011), although conflicting results have also been reported (Abutalebi et al., 2012). As posited by Costa et al. (2008), experience with a task is negatively associated with its cognitive demands, with well-practiced tasks being unlikely to identify bilingual effects. Empirical evidence supporting this claim was later identified (Costa et al., 2009), serving as a key component of current theoretical accounts of bilingual effects on cognitive control (Bialystok, 2024; Bialystok & Craik, 2022). The near-ceiling accuracy rates observed on the Simon task suggest it was not particularly challenging, which may have obscured any effects of bilingual language experience on final performance measures, especially considering the number of trials on the task (i.e., 150). Our finding, consistent with previous evidence syntheses (Hilchey & Klein, 2011; Hilchey et al., 2015), underscores the importance of identifying methodological factors that may prevent the identification of authentic bilingual effects on cognitive control, and further highlights the utility of alternative measures that capture different aspects of performance.

4.4. Higher L2 relative to L1 usage associated with faster learning

A novel contribution of the present study is the identification of a significant bilingual effect on rate of change, a performance trajectory measure capturing differences in learning. Higher usage of English relative to Mandarin was associated with faster overall learning on the Simon task. Investigations of the relationship between bilingualism and learning suggest a positive association (Adesope et al., 2010), although studies on specific kinds of learning, such as statistical learning, have reported mixed findings, potentially due to unaccounted heterogeneity in bilingual experience (Bulgarelli et al., 2018). Improved learning associated with higher levels of bilingual experience likely reflects the broader impact of bilingualism on cognitive control given its crucial role in learning and memory (Diamond, 2013; Duff et al., 2005).

Higher use of English relative to Mandarin has been previously associated with both positive and negative bilingual effects on cognitive control. Mandarin-English bilingual adolescents in China demonstrated a negative bilingual effect on a Flanker task, with higher L2/L1 dominance ratio associated with slower global RTs, but no influence on Simon task performance (Privitera et al., 2022). An identical finding was also observed on no-cue trials from the ANT, but emerged alongside a positive bilingual effect on orienting network function, with higher levels of English use relative to Mandarin associated with faster responses on spatial cue trials (Privitera, Momenian, & Weekes, 2023). Considering that the present study only identified a positive bilingual effect of L2/L1 dominance ratio on overall rate of change and not any conventional, initial, or final performance measure, and that no explicit orienting cues were presented on the Simon task, our finding may reflect the unmeasured influence of improved orienting network function on learning in general (e.g., Leclercq & Seitz, 2012; Pederson & Guion-Anderson, 2010). Support for this interpretation can be found in the same study by Privitera, Momenian, and Weekes (2023) in which a positive bilingual effect of L2/L1 dominance ratio on orienting, together with a positive correlation between L2/L1 dominance ratio and L2 proficiency, was interpreted as increased sensitivity to environmental cues supporting the identification of opportunities to use and, subsequently, improve (i.e., learn) English. The same correlation between L2/L1 dominance ratio and L2 proficiency was identified in the present study, supporting a similar interpretation. Alternatively, the cognitive demands placed on bilinguals who use their L2 more relative to their L1 in an L1-dominant environment may support the emergence of bilingual effects that are more readily observed on measures of learning. Considering that our participants were recruited from Mainland China, a Mandarin-dominant environment, those with higher L2/L1 dominance ratios may be more adept at learning when, where, and with whom they can use English. However, this interpretation should be considered speculative in the absence of strong evidence.

To summarize, the present study identified distinct associations between separable dimensions of bilingual language experience and both conventional and performance trajectory measures from a Simon task. While earlier accounts predicted that bilingual effects would manifest exclusively on incongruent trials or interference scores, reflecting improved inhibitory control (Bialystok, 2001), more recent accounts suggest these effects are likely more general (Bialystok, 2024; Bialystok & Craik, 2022; Hilchey et al., 2015; Hilchey & Klein, 2011). Our pattern of results aligns with these more recent accounts, although this included both positive and negative bilingual effects. Additionally, historic methodological reliance on categorical language status labels (Luk & Bialystok, 2013) and ecologically flawed monolingual-bilingual comparisons (Rothman et al., 2023) complicated the interpretation of our findings regarding the relationship between a given dimension of bilingual experience and cognitive control. Negative associations between dimensions of bilingual language experience and cognitive control may consistently be present, but their identification can be obscured by these common methodological approaches and by the masking effects of positive and null relationships with other dimensions. Only recently have investigations of bilingual effects on cognitive control considered separable dimensions of language experience, with growing evidence supporting that these relationships can differ across dimensions of both language experience and cognitive control (e.g., Privitera, Momenian, & Weekes, 2023; Xie et al., 2024; Xie & Pisano, 2019; Yow & Li, 2015; Yurtsever et al., 2024). For these reasons, interpretations described above may change considerably as future investigations better characterize the impact of separable dimensions of bilingual experience and cognitive control.

4.5. Limitations

The present study represents the first investigation of how bilingual language experience impacts time-sensitive measures of

https://doi.org/10.1017/S1366728925100576 Published online by Cambridge University Press

performance trajectory derived from Simon task performance. Accordingly, these results should be interpreted with caution as they require further validation through replication across diverse bilingual samples and a range of cognitive control tasks. Additionally, we did not observe robust bilingual effects on cognitive control, and the effect sizes associated with significant findings were generally modest. This may have resulted due to our reliance on a single behavioral task (Ware et al., 2020) or the use of a large number of experimental trials (Hilchey & Klein, 2011). These methodological decisions were made based on the desire to reduce the time burden on participants while ensuring a sufficient number of trials for generation of performance trajectory measures (Cochrane & Green, 2021). While we employed bootstrapping in our analyses to improve the accuracy of estimates and help address sample size issues, these limitations nonetheless underscore the need for further research with larger samples and more sensitive measures. Finally, as described in our discussion, the unique linguistic profile of our sample may reduce generalizability of our findings to other bilingual populations.

5. Conclusions

There is a need to reconsider how bilingualism influences cognitive control and which analytical approaches most effectively reveal these effects. Findings from the present study suggest that the nearly universal reliance on overall task performance measures may be partially responsible for mixed results across previous reports. Furthermore, their use perpetuates the assumption that performance on tasks measuring cognitive control is stable, an assumption our data do not support. We suggest that complementing conventional performance measures with those that reflect different aspects of task performance, including learning, should be considered best practice. Finally, our observed pattern of results underscores the importance of investigating the contribution of separable dimensions of language experience in order to understand the complex relationship between bilingualism and cognitive control.

Acknowledgements. Thank you to Dr Xia Lihua at Huazhong University of Science and Technology for supporting the ethical approval and data collection for this study and Dr Aaron Cochrane at Brown University for his extraordinary patience in answering my seemingly endless stream of questions about data analysis.

Data availability statement. Full data and analyses can be accessed on Open Science Framework (https://doi.org/10.17605/OSF.IO/H5XVG).

Competing interests. The authors have no competing interests to declare.

References

Abutalebi, J., **Della Rosa, P. A.**, **Green, D. W.**, **Hernandez, M.**, **Scifo, P.**, **Keim, R.**, **Cappa, S. F.**, & **Costa, A.** (2012). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cerebral Cortex*, **22**(9), 2076–2086.

Abutalebi, J., & **Green, D.** (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, **20**(3), 242–275.

Abutalebi, J., **Guidi, L.**, **Borsa, V.**, **Canini, M.**, **Della Rosa, P. A.**, **Parris, B. A.**, & **Weekes, B. S.** (2015). Bilingualism provides a neural reserve for aging populations. *Neuropsychologia*, **69**, 201–210.

Adesope, O. O., **Lavin, T.**, **Thompson, T.**, & **Ungerleider, C.** (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research*, **80**(2), 207–245.

Anthony, J. J. R., & **Blumenfeld, H. K.** (2019). Language dominance predicts cognate effects and inhibitory control in young adult bilinguals. *Bilingualism: Language and Cognition*, **22**(5), 1068–1084.

Antoniou, M. (2019). The advantages of bilingualism debate. *Annual Review of Linguistics*, **5**, 395–415.

Anwyl-Irvine, A. L., **Massonnié, J.**, **Flitton, A.**, **Kirkham, N.**, & **Evershed, J. K.** (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, **52**(1), 388–407.

Arthur, W., **Tubre, T. C.**, **Paul, D. S.**, & **Sanchez-Ku, M. L.** (1999). College-sample psychometric and normative data on a short form of the Raven advanced progressive matrices test. *Journal of Psychoeducational Assessment*, **17**(4), 354–361.

Bak, T. H., **Vega-Mendoza, M.**, & **Sorace, A.** (2014). Never too late? An advantage on tests of auditory attention extends to late bilinguals. *Frontiers in Psychology*, **5**, 485.

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

Bialystok, E. (1992). Selective attention in cognitive processing: The bilingual edge. In R. J. Harris (Ed.), *Advances in Psychology* (Vol. 83, pp. 501–513). North-Holland.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge University Press.

Bialystok, E. (2024). Bilingualism modifies cognition through adaptation, not transfer. *Trends in Cognitive Sciences*, **28**(11), 987–997.

Bialystok, E., & **Craik, F. I. M.** (2022). How does bilingualism modify cognitive function? Attention to the mechanism. *Psychonomic Bulletin & Review*, **29**(4), 1246–1269.

Bialystok, E., **Craik, F. I. M.**, **Klein, R.**, & **Viswanathan, M.** (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, **19**(2), 290.

Bialystok, E., **Poarch, G.**, **Luo, L.**, & **Craik, F. I. M.** (2014). Effects of bilingualism and aging on executive function and working memory. *Psychology and Aging*, **29**(3), 696.

Branzi, F. M., **Della Rosa, P. A.**, **Canini, M.**, **Costa, A.**, & **Abutalebi, J.** (2016). Language control in bilinguals: Monitoring and response selection. *Cerebral Cortex*, **26**(6), 2367–2380.

Bulgarelli, F., **Lebkuecher, A. L.**, & **Weiss, D. J.** (2018). Statistical learning and bilingualism. *Language, Speech, and Hearing Services in Schools*, **49**(3S), 740–753.

Bürkner, P.-C. (2017). Brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, **80**, 1–28.

Cochrane, A. (2020). TEfits: Nonlinear regression for time-evolving indices. *Journal of Open Source Software*, **5**(52), 2535.

Cochrane, A., & **Green, C. S.** (2021). Trajectories of performance change indicate multiple dissociable links between working memory and fluid intelligence. *npj Science of Learning*, **6**(1), 33.

Costa, A., **Hernández, M.**, **Costa-Faidella, J.**, & **Sebastián-Gallés, N.** (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, **113**(2), 135–149.

Costa, A., **Hernández, M.**, & **Sebastián-Gallés, N.** (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, **106**(1), 59–86.

Costa, A., **Miozzo, M.**, & **Caramazza, A.** (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language*, **41**(3), 365–397.

Craney, T. A., & **Surles, J. G.** (2002). Model-dependent variance inflation factor cutoff values. *Quality Engineering*, **14**(3), 391–403.

D'Souza, D., & **D'Souza, H.** (2021). Bilingual adaptations in early development. *Trends in Cognitive Sciences*, **25**(9), 727–729.

De Bruin, A. (2019). Not all bilinguals are the same: A call for more detailed assessments and descriptions of bilingual experiences. *Behavioral Sciences*, **9**(3), 33.

Del Maschio, N., **Del Mauro, G.**, **Bellini, C.**, **Abutalebi, J.**, & **Sulpizio, S.** (2022). Foreign to whom? Constraining the moral foreign language effect on bilinguals' language experience. *Language and Cognition*, **14**(4), 511–533.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, **64**, 135–168.

Donnelly, S., **Brooks, P. J.**, & **Homer, B. D.** (2019). Is there a bilingual advantage on interference-control tasks? A multiverse meta-analysis of global reaction time and interference cost. *Psychonomic Bulletin & Review*, **26**(4), 1122–1147.

D'Souza, D., **Brady, D.**, **Haensel, J. X.**, & **D'Souza, H.** (2020). Is mere exposure enough? The effects of bilingual environments on infant cognitive development. *Royal Society Open Science*, **7**(2), 180191.

Duff, K., **Schoenberg, M. R.**, **Scott, J. G.**, & **Adams, R. L.** (2005). The relationship between executive functioning and verbal and visual learning and memory. *Archives of Clinical Neuropsychology*, **20**(1), 111–122.

Durand López, E. M. (2021). A bilingual advantage in memory capacity: Assessing the roles of proficiency, number of languages acquired and age of acquisition. *International Journal of Bilingualism*, **25**(3), 606–621.

Eriksen, B. A., & **Eriksen, C. W.** (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, **16**(1), 143–149.

Fan, J., **McCandliss, B. D.**, **Sommer, T.**, **Raz, A.**, & **Posner, M. I.** (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, **14**(3), 340–347.

Friedman, N. P., & **Miyake, A.** (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, **86**, 186–204.

Giovannoli, J., **Martella, D.**, **Federico, F.**, **Pirchio, S.**, & **Casagrande, M.** (2020). The impact of bilingualism on executive functions in children and adolescents: A systematic review based on the PRISMA method. *Frontiers in Psychology*, **11**, 2398.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, **1**(2), 67–81.

Green, D. W., & **Abutalebi, J.** (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, **25**(5), 515–530.

Grundy, J. G. (2020). The effects of bilingualism on executive functions: An updated quantitative analysis. *Journal of Cultural Cognitive Science*, **4**(2), 177–199.

Grundy, J. G., **Chung-Fat-Yim, A.**, **Friesen, D. C.**, **Mak, L.**, & **Bialystok, E.** (2017). Sequential congruency effects reveal differences in disengagement of attention for monolingual and bilingual young adults. *Cognition*, **163**, 42–55.

Guðmundsdóttir, M. D., & **Lesk, V. E.** (2019). Does the bilingual advantage extend to trilingualism? *International Journal of Multilingualism*, **16**(4), 549–562.

Gullifer, J. W., **Kousaie, S.**, **Gilbert, A. C.**, **Grant, A.**, **Giroud, N.**, **Coulter, K.**, **Klein, D.**, **Baum, S.**, **Phillips, N.**, & **Titone, D.** (2021). Bilingual language experience as a multidimensional spectrum: Associations with objective and subjective language proficiency. *Applied Psycholinguistics*, **42**, 1–34.

Hamel, R., & **Schmittmann, V. D.** (2006). The 20-minute version as a predictor of the Raven advanced progressive matrices test. *Educational and Psychological Measurement*, **66**(6), 1039–1046.

Hilchey, M. D., & **Klein, R. M.** (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review*, **18**(4), 625–658.

Hilchey, M. D., **Saint-Aubin, J.**, & **Klein, R. M.** (2015). Does bilingual exercise enhance cognitive fitness in traditional non-linguistic executive processing tasks. *The Cambridge Handbook of Bilingual Processing*, 586–613.

JASP Team. (2024). *JASP (Version 0.19.1)* [Computer software]. https://jasp-stats.org/.

Kattner, F., **Cochrane, A.**, & **Green, C. S.** (2017). Trial-dependent psychometric functions accounting for perceptual learning in 2-AFC discrimination tasks. *Journal of Vision*, **17**(11), 3–3.

Kroll, J. F., **Dussias, P. E.**, **Bogulski, C. A.**, & **Kroff, J. R. V.** (2012). Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In *Psychology of learning and motivation* (Vol. **56**, pp. 229–262). Elsevier.

Leclercq, V., & **Seitz, A. R.** (2012). The impact of orienting attention in fast task-irrelevant perceptual learning. *Attention, Perception, & Psychophysics*, **74**, 648–660.

Lehtonen, M., **Soveri, A.**, **Laine, A.**, **Järvenpää, J.**, **De Bruin, A.**, & **Antfolk, J.** (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin*, **144**(4), 394.

Li, P., **Zhang, F.**, **Yu, A.**, & **Zhao, X.** (2020). Language history questionnaire (LHQ3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition*, **23**(5), 938–944.

Luk, G., & **Bialystok, E.** (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, **25**(5), 605–621.

Luk, G., **De Sa, E.**, & **Bialystok, E.** (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control? *Bilingualism: Language and Cognition*, **14**(4), 588–595.

Madrazo, A. R., & **Bernardo, A. B. I.** (2018). Measuring two types of inhibitory control in bilinguals and trilinguals: Is there a trilingual advantage? *Psychological Studies*, **63**, 52–60.

Marian, V., & **Spivey, M.** (2003). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition*, **6**(2), 97–115.

Miyake, A., **Friedman, N. P.**, **Emerson, M. J.**, **Witzki, A. H.**, **Howerter, A.**, & **Wager, T. D.** (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, **41**(1), 49–100.

Morton, J. B., & **Harper, S. N.** (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental Science*, **10**(6), 719–726.

Naeem, K., **Filippi, R.**, **Periche-Tomas, E.**, **Papageorgiou, A.**, & **Bright, P.** (2018). The importance of socioeconomic status as a modulator of the bilingual advantage in cognitive ability. *Frontiers in Psychology*, **9**, 1818.

Novitskiy, N., **Shtyrov, Y.**, & **Myachykov, A.** (2019). Conflict resolution ability in late bilinguals improves with increased second-language proficiency: Ant evidence. *Frontiers in Psychology*, **10**, 2825.

Ong, G., **Sewell, D. K.**, **Weekes, B.**, **McKague, M.**, & **Abutalebi, J.** (2017). A diffusion model approach to analysing the bilingual advantage for the flanker task: The role of attentional control processes. *Journal of Neurolinguistics*, **43**, 28–38.

Paap, K. (2019). The bilingual advantage debate: Quantity and quality of the evidence. *The Handbook of the Neuroscience of Multilingualism*, 701–735.

Peal, E., & **Lambert, W. E.** (1962). The relation of bilingualism to intelligence. *Psychological Monographs: General and Applied*, **76**(27), 1.

Pederson, E., & **Guion-Anderson, S.** (2010). Orienting attention during phonetic training facilitates learning. *The Journal of the Acoustical Society of America*, **127**(2), EL54–EL59.

Pelham, S. D., & **Abrams, L.** (2014). Cognitive advantages and disadvantages in early and late bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **40**(2), 313.

Poarch, G. J., & **Van Hell, J. G.** (2012). Executive functions and inhibitory control in multilingual children: Evidence from second-language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology*, **113**(4), 535–551.

Privitera, A. J. (2024a). Is dialect proficiency associated with improved executive function? *Applied Psycholinguistics*, **45**(5), 811–833. https://doi.org/10.1017/S0142716424000286.

Privitera, A. J. (2024b). Is there a foreign language effect on academic integrity? *Higher Education*. https://doi.org/10.1007/s10734-023-01134-5

Privitera, A. J., **Li, S.**, **Zhou, Y.**, & **Wang, M.** (2023). Modulatory role of foreign language experience on the moral foreign language effect. *Bilingualism: Language and Cognition*, **26**(5), 1038–1050.

Privitera, A. J., **Momenian, M.**, & **Weekes, B. S.** (2022). Task-specific bilingual effects in Mandarin-English speaking high school students in China. *Current Research in Behavioral Sciences*, **3**, 100066.

Privitera, A. J., **Momenian, M.**, & **Weekes, B. S.** (2023). Graded bilingual effects on attentional network function in Chinese high school students. *Bilingualism: Language and Cognition*, **26**(3), 527–537. https://doi.org/10.1017/S1366728922000803.

Privitera, A. J., & **Weekes, B. S.** (2023). Scoping review of research practices in the investigation of bilingual effects on inhibition and attention in young people. *International Journal of Bilingualism*, **27**(5), 749–775.

R Core Team. (2023). *R: A language and environment for statistical computing*. In Raven, J. C. (1962). Advanced progressive matrices: Sets I and II. *H.K.* London: Lewis.

Rothman, J., **Bayram, F.**, **DeLuca, V.**, **Di Pisa, G.**, **Dunabeitia, J. A.**, **Gharibi, K.**, **Hao, J.**, **Kolb, N.**, **Kubota, M.**, & **Kupisch, T.** (2023). Monolingual comparative normativity in bilingualism research is out of "control": Arguments and alternatives. *Applied Psycholinguistics*, **44**(3), 316–329.

https://doi.org/10.1017/S1366728925100576 Published online by Cambridge University Press

Schroeder, S. R., & **Marian, V.** (2017). Cognitive consequences of trilingualism. *International Journal of Bilingualism*, **21**(6), 754–773.

Simon, J. R., & **Rudell, A. P.** (1967). Auditory SR compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, **51**(3), 300.

Speelman, C. P., & **McGann, M.** (2013). How mean is the mean? *Frontiers in Psychology*, **4**, 451.

Stoica, P., & **Selen, Y.** (2004). Model-order selection: A review of information criterion rules. *IEEE Signal Processing Magazine*, **21**(4), 36–47.

Tao, L., **Marzecová, A.**, **Taft, M.**, **Asanowicz, D.**, & **Wodniecka, Z.** (2011). The efficiency of attentional networks in early and late bilinguals: The role of age of acquisition. *Frontiers in Psychology*, **2**, 123.

Van den Noort, M., **Struys, E.**, **Bosch, P.**, **Jaswetz, L.**, **Perriard, B.**, **Yeo, S.**, **Barisch, P.**, **Vermeire, K.**, **Lee, S.-H.**, & **Lim, S.** (2019). Does the bilingual advantage in cognitive control exist and if so, what are its modulating factors? A systematic review. *Behavioral Sciences*, **9**(3), 27.

Ware, A. T., **Kirkovski, M.**, & **Lum, J. A. G.** (2020). Meta-analysis reveals a bilingual advantage that is dependent on task and age. *Frontiers in Psychology*, **11**, 1458.

Wermelinger, S., **Gampe, A.**, & **Daum, M. M.** (2017). Bilingual toddlers have advanced abilities to repair communication failure. *Journal of Experimental Child Psychology*, **155**, 84–94.

Woumans, E., **Ceuleers, E.**, **Van der Linden, L.**, **Szmalec, A.**, & **Duyck, W.** (2015). Verbal and nonverbal cognitive control in bilinguals and interpreters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **41**(5), 1579.

Xie, Z., & **Pisano, T. S.** (2019). Second language (L2) proficiency, socioeconomic status (SES), and intelligence (IQ) are significant predictors of cognitive control differences among young adult unbalanced Chinese–English bilinguals. *Bilingualism: Language and Cognition*, **22**(4), 866–882.

Xie, Z., **Privitera, A. J.**, & **Huang, Y.** (2024). Bilingualism from a holistic perspective: Investigation of bilingual effects on cognitive control. *Current Psychology*, **43**, 32437–32450. https://doi.org/10.1007/s12144-024-06800-0.

Yow, W. Q., & **Li, X.** (2015). Balanced bilingualism and early age of second language acquisition as the underlying mechanisms of a bilingual executive control advantage: Why variations in bilingual experiences matter. *Frontiers in Psychology*, **6**, 164.

Yurtsever, A., **Sotebeer, K. N.**, & **Grundy, J. G.** (2024). *Different bilingual experiences predict different executive functions: Evidence from mouse-tracking* (pp. 1–12). Bilingualism: Language and Cognition.

Zhou, B., & **Krott, A.** (2016). Data trimming procedure can eliminate bilingual cognitive advantage. *Psychonomic Bulletin & Review*, **23**(4), 1221–1230. https://doi.org/10.3758/s13423-015-0981-6.