Brain and Cognition 187 (2025) 106308

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

Cognitive efficiency and expertise-dependent automaticity in the working memory performance of bilinguals

Federico Gallo [a,*], Liliia Terekhina [b], Jubin Abutalebi [a,b,c], Yury Shtyrov [c,d], Andriy Myachykov [c,e]

[a] PoLaR Lab, Center for Language, Brain and Learning, UiT-The Arctic University of Norway, Tromsø, Norway
[b] Centre for Neurolinguistics and Psycholinguistics, Vita-Salute San Raffaele University, Milan, Italy
[c] Cognitive Health and Intelligence Centre, Institute for Cognitive Neuroscience, HSE University, Moscow, Russian Federation
[d] Center of Functionally Integrative Neuroscience, Aarhus University, Aarhus, Denmark
[e] University of Macau, Centre for Cognitive and Brain Sciences, N21, Avenida da Universidade, Taipa, Macau

ARTICLE INFO

Keywords:
Bilingual experience
Working memory
Cognitive efficiency
Expertise-dependent automaticity

ABSTRACT

The concepts of cognitive efficiency (CE) and expertise-dependent automaticity are central to the understanding of cognitive adaptations related to bilingual experiences. This study examined their behavioral manifestations in bilingual young adults by manipulating the cognitive load of a working memory task; the possibility to manipulate the difficulty of a cognitive task is necessary to observe behavioral outcomes associated with CE and automaticity. To this end, participants completed an n-back task ranging in difficulty from 0-back to 3-back, with the 3-back condition being commonly recognized as being a highly cognitively demanding one. We aimed to determine whether degree of bilingual experience could predict performance outcomes – accuracy, reaction times, and the speed/accuracy tradeoff – reflecting bilingualism's putative dynamic impact on CE and automaticity in working memory. The results showed a positive relationship between degree of bilingual experience and working memory performance, particularly when the task-induced cognitive load increased. More experienced bilinguals demonstrated a smaller decline in performance when task difficulty intensified, a behavioral manifestation compatible with increased CE. Additionally, a relationship with expertise-dependent automaticity emerged, with the speed/accuracy tradeoff trajectory unfolding differentially across varying task difficulties at different degrees of bilingual experience.

1. Introduction

Bi/multilingual[1] engagement has been suggested to induce (under specific circumstances at least) functional and structural adaptations in the executive neural substrate and respective cognitive abilities (Bialystok, 2017; DeLuca et al., 2020; Kroll et al., 2015; Pliatsikas, 2020; Tao et al., 2021). One could argue that the constant necessity to manage and control multiple languages in order to resolve potential cross-linguistic conflicts and to avoid unwanted intrusions from non-target language(s) in order to carry out successful communication is a cognitive trait unique to bilingual speakers (e.g., Kroll et al., 2008). To accomplish this, bilinguals rely on *language control*, a specific cognitive ability employing executive mechanisms such as working memory, inhibition, attentional control and task switching. The neural substrate of language control has been widely investigated, and the available evidence points to a network that extensively overlaps with that subserving the domain-general executive functions (Abutalebi & Green, 2007, 2016; Green & Abutalebi, 2013). Thus, the routine engagement of executive resources for bilingual language control is hypothesized to strain – and subsequently train and strengthen – the cognitive functioning of bilingual individuals (Bialystok, 2017; Kroll et al., 2015).

A recent proposal (Bialystok, 2024; Bialystok & Craik, 2022)

* Corresponding author at: Psycholinguistics of Language Representation (PoLaR) lab, Center for Language, Brain and Learning (C-LaBL), UiT, The Arctic University of Norway, Hansine Hansens veg 18, 9019 Tromsø, Norway.
E-mail address: federico.gallo@uit.no (F. Gallo).

[1] To streamline the discussion, except when strictly referring to the current experimental sample, the term "bilingual(−ism)" will be used throughout this paper as an umbrella term referring to all forms of bi- and multilingualism. This term allows us to distinguish between individuals who know only one language and those who know more than one, without any assumption that all forms of bi- or multilingualism are identical.

https://doi.org/10.1016/j.bandc.2025.106308
Received 20 February 2025; Received in revised form 6 May 2025; Accepted 7 May 2025
Available online 12 May 2025

broadens this traditional view focused specifically on executive functions aiming to unify diverse findings across the lifespan within a comprehensive theoretical framework. Central to this proposal is the idea that adaptations in selective attention serve as the primary mechanism through which bilingualism influences cognitive processes. Adaptions in selective attention – an underlying resource affecting cognitive performance broadly – would eventually result in downstream effects for a wide range of more specific cognitive functions, depending on the specific circumstances of an individual's bilingual experience. This proposal argues that the cognitive consequences of bilingual experience come into existence via *increased efficiency in managing attentional resources*. An argument for bilingualism-induced increases in cognitive efficiency, irrespective of the specifically targeted cognitive function (i.e., selective attention or other executive functions) and/or preferred label system, is a linking point between this and previous attempts to detail the mechanisms underlying bilingualism's cognitive consequences (e.g., DeLuca et al., 2020; Grundy et al., 2017; Pliatsikas, 2020; Stocco et al., 2014). This argument is corroborated by experimental findings showing that bilinguals need to deploy less attentional resources than their monolingual peers to achieve comparable task performance (Abutalebi et al., 2012; Gold et al., 2013).

Thus, *cognitive efficiency* (CE) emerges as a key pathway through which bilingualism modifies cognitive functioning. CE is a concept that has been widely researched in cognitive science outside the bilingualism field (e.g., Pyc & Rawson, 2007; Rypma et al., 2006; Stilley et al., 2010; Vernon, 1983). Although its conceptualizations and operationalizations vary broadly in the literature, CE can been defined as the ratio between the quality of cognitive performance, or conceptual clarity of knowledge, and costs, such as cognitive effort, required to attain said performance or knowledge (see Hoffman, 2012 for review). It is important to note that CE is linked to but not identical to the concept of *neural efficiency* – the dynamic changes in the levels of neural activation that underlie the attribution of CE (Haier et al., 1992). As the current study focuses exclusively on behavioral data, we use the term CE as solely reflecting behavioral benefits and/or penalties resulting from bilingual experience. As such, higher CE is expected to manifest as relatively smaller drops in performance despite increased task difficulty, since individuals with higher efficiency would be able to retain higher amounts of cognitive resources to be deployed in the face of higher cognitive demands (Neubauer & Fink, 2009; Reuter-Lorenz & Cappell, 2008; Rypma et al., 2006). This, importantly, also implies that it is impossible to investigate the behavioral indices of CE without manipulating task difficulty.

At the same time, task difficulty manipulation has also been proposed as a means to address the long-standing debate regarding the lack of consistency across experimental studies of bilingualism-induced cognitive modifications (see e.g., Antoniou, 2019; Baum & Titone, 2014; Bialystok, 2017). Indeed, some research groups have repeatedly questioned the magnitude (and, in some cases, the actual existence) of the cognitive consequences of bilingualism, based on the irregular emergence of supporting results across the literature (e.g., De Bruin et al., 2015; Duñabeitia & Carreiras, 2015; Paap, 2022). Several arguments have in turn been put forward to explain these mixed results. One such argument is that limited task sensitivity could prevent the observation of the cognitive consequences of bilingual experience (Kroll & Bialystok, 2013; Leivada et al., 2021; Valian, 2015). This is particularly relevant for young adults—the predominant population studied in cognitive science research (see e.g., Henrich et al., 2010)—who generally possess optimal cognitive resources and engage in numerous cognitively stimulating activities (Hartshorne & Germine, 2015) alongside bilingualism, potentially obscuring the detection of its cognitive benefits. This observation highlights the critical need for the field to adopt more fine-grained tasks that permit systematic manipulation of difficulty levels. This is particularly important given that behavioral manifestations of CE typically require manipulation of task difficulty to become observable, and CE is regarded as a central mechanism through which the cognitive effects of bilingualism are realized (Bialystok & Craik, 2022; DeLuca et al., 2020; Pliatsikas, 2020, *inter alia*). For instance, a task designed to systematically vary cognitive load could be employed to tax the cognitive resources of young individuals, thereby facilitating the emergence of individual differences more distinctly.

One sub-domain of executive functions for which such an opportunity is readily available is *working memory*, whose efficiency has been repeatedly shown to be affected by bilingualism (e.g., Anderson et al., 2021; Comishen & Bialystok, 2021; Kudo & Swanson, 2014; Morales et al., 2013; Morrison et al., 2019, but see also Lehtonen et al., 2018). Indeed, one of the most widely adopted measures of working memory capacity, the *n-back task*, readily allows for manipulation of task difficulty, thereby enabling substantial variation in cognitive demands. In this task, participants are to compare the current stimulus with one presented *n* trials before, and to indicate whether the two stimuli match. Thus, the task requires sustained involvement of working memory which must be updated in a continuous (trial-by-trial) fashion, and – crucially – the cognitive demand can be easily escalated by increasing the magnitude of *n*. A 3-back task will therefore be more cognitively taxing than a 2-back one, which will be more cognitively taxing than 1-back, and so on. This is typically reflected in lower accuracy and slower reaction times (RTs) for higher *n* (Mattay et al., 2006), allowing for objective evaluation of the task-specific cognitive load, on the one hand, and of the participants' efficiency – on the other.

Furthermore, differences in performance on more complex tasks (such as the 2- or 3-back) may also originate from individual differences in *automaticity* (Logan, 1988; Moors & De Houwer, 2006) as a function of acquired skill/expertise relative to a certain cognitive function. Automaticity is classically defined by reduced reliance on controlled processing and more efficient task execution (Logan, 1988), with increased proficiency leading to faster responses without sacrificing accuracy. Differences in automaticity are reflected, for instance, in differential speed/accuracy tradeoffs (SATs) between skilled and novice performers (Beilock et al., 2008). Research has shown that experienced individuals perform better when attention is limited during execution – in other words, when performance is more *automated* – compared to situations that encourage step-by-step control, which tend to impair their performance and benefit that of novices (e.g., Beilock et al., 2002, 2008; Perkins-Ceccato et al., 2003). Because novice performance requires more attentional resources than skilled performance, conditions that divert attention away from the primary task are more detrimental for novices than for experts. Thus, novice participants typically demonstrate the "traditional" SAT pattern (i.e., slower performance = better accuracy) while skilled performers show its reversal (i.e., faster performance = better accuracy).

Building on these premises, we aimed to examine the behavioral manifestations of CE in bilingual young adults by systematically manipulating the difficulty of a working memory task. This approach was intended to tax participants' cognitive resources to varying degrees, thereby enhancing the likelihood of detecting the cognitive consequences associated with bilingualism. Our bilingual young adult participants underwent an n-back task in its iterations from 0- to 3-back, with the latter broadly considered to be a highly demanding task. We subsequently examined whether, and in what ways, individual differences in the degree of bilingual experience modulate trajectories of accuracy, RTs and SATs across increasing levels of task difficulty—aiming to capture the effects of CE and expertise-dependent automaticity on executive functioning. Based on existing evidence, we expected higher degrees of bilingual experience to be associated with (1) increased CE as measured at the behavioral level, manifesting as better performance maintenance—and increasingly stronger positive contributions to performance—under escalating task difficulty; (2) expertise-dependent automaticity, whereby varying degrees of bilingual experience would give rise to distinct trajectories in SATs as task difficulty increases.

2. Materials and methods

2.1. Participants

Forty-five bilingual participants from the student population of HSE University in Moscow, Russian Federation (13 males; L1: Russian, L2: English; mean age = 20.5, SD ± 3) took part in the study. All participants were right-handed, as assessed via the Edinburgh Handedness Inventory scale (Oldfield, 1971), and reported no psychiatric or neurological impairments. The socio-demographic characteristics of participants, including age, educational attainment, and socio-economic status (SES), were evaluated using the MacArthur Scale of Subjective Social Status (MacArthur Foundation, 2007). To estimate SES, annual household income categories were adapted to local standards based on data from the European Social Survey 2020 (ESS Round 10: European Social Survey Round 10 Data, 2020). General intelligence was assessed using a subset of Raven's Standard Progressive Matrices for adults (Court & Raven, 1992). Individual profiles of bilingual experience were assessed via the Russian version of the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007) on several dimensions including L2 exposure across different modalities (i.e., writing, reading, speaking, listening) and L2 age of acquisition (AoA). We also tested participants on the Cambridge Test for Adult Learners (http://www.cambridgeenglish.org/test-your-english/general-english/) to obtain an objective measure of their general proficiency in L2 English. Table 1 presents the demographics and language background characteristics of the study sample. The study received approval from the local research ethics committee; all participants gave their written informed consent and were made aware that they could withdraw from the experiment at any time.

2.2. Working memory assessment

To assess working memory performance, we used an n-back task (Dobbs & Rule, 1989) of increasing difficulty. The task was administered on a 75-cm diagonal computer screen using the PsychoPy software (Peirce et al., 2019). Ten consonant letters in native L1 Cyrillic script (Б, Г, Ж, К, Л, М, П, Ф, Ч, Ш) were used as stimuli, with vowels excluded to reduce the likelihood of participants developing chunking strategies for memorization across trials (Grimes et al., 2008; Pereira Soares et al., 2024).

The n-back task included four memory load conditions. Each experimental block tested one condition. In the 0-back block, participants were required to press the rightmost button on a response box (RB-740, Cedrus Corp.) if the letter currently on the screen matched the first letter of the block. Otherwise, they needed to press the leftmost button. For instance, if a block started with the letter M, every time M appeared on the screen subsequently participants had to press the right button. For any other letter, they had to press the leftmost button. In the 1-back block, participants had to press the right button if the current letter matched the one immediately preceding it (e.g., an M followed by another M). In the 2-back block, participants pressed the right button if the current letter matched the one presented two trials earlier (e.g., an M followed by a K and then another M, as in the sequence xxMKMxx). In the 3-back block, participants pressed the right button if the current letter matched the one presented three trials earlier (e.g., a M followed by two different letters and then another M, as in the sequence xxMKΦMxx).

Each block included a practice session followed by the experimental session and lasted approximately 10 min in total. The whole experimental session lasted ~ 40 min. Both the practice and the experimental sessions within each block contained 25 % match trials (4 out of 16 trials in the practice session and 60 out of 240 trials in the experimental session). After receiving experimental instructions, participants proceeded with the first block (0-back). In each block, a brief practice session of 16 trials with feedback was presented, followed by the experimental session including 240 trials presented in a randomized order. After each block, participants were allowed to take a break or proceed directly to the next block. The order of blocks was constant across participants (0-back, 1-back, 2-back, 3-back). Trials lasted until the participant responded, with a 2000-millisecond response timeout. A 1700 ms inter-trial interval (ITI) with a blank screen was presented between trials.

Table 1
Descriptive statistics of socio-demographic, cognitive, and language background measures. Mean, standard deviation (SD), and range for each measure are reported.

	Mean (SD)	Range (min – max)
Age (years)	20.467 (3.02)	18 – 34
Education (years)	13.6 (2.1)	10 – 19
SES (score)	3.78 (1.97)	1 – 6
Raven's Matrices (score)	14.53 (0.66)	12 – 15
L2 AoA (years)	7.8 (2.37)	3 – 13
L2 exposure (daily percentage)	19.44 (12.94)	2 – 51
L2 proficiency (Cambridge test score)	16.78 (4.33)	8 – 25

2.3. Statistical analyses

2.3.1. Data preprocessing

Both accuracies and RTs were analyzed. The analyses of accuracy and SAT (see below for more details) included all target trials. For the RT analyses, data pre-processing included removal of incorrect and false-start (RT < 100 ms) trials, as well as outlier trials displaying RTs beyond 3 SDs from the individual mean RT values (Baayen & Milin, 2010). This procedure resulted in discarding 8.3 % of the total data for the 0-back condition, 13.6 % of the total data for the 1-back condition, 22.9 % of the total data for the 2-back condition and 39.6 % of the total data for the 3-back condition. All statistical analyses were conducted using Stata 17 (StataCorp, 2021).

2.3.2. Accuracy

To assess whether the degree of bilingual experience modulates the relationship between variations in task difficulty and accuracy of working memory performance, we proceeded as follows. Data was coded in long format, with each participant appearing four times in the data matrix: the total accuracy score (0–60) for each of the four n-back blocks (0-, 1-, 2- and 3-back) constituted one row. To account for repeated measures, we used a linear-mixed effects approach including random intercepts for participants. Then, we included an interaction term between bilingual experience factor(s) and task difficulty (n-back block) in our models. This allowed the models to distinguish between the accuracy score of each block for each individual and estimate whether variations in bilingual experience factors modulated the expected progressive decrease in accuracy with increasing task difficulty.

Thus, we tested three linear mixed regression models with three different bilingual experience factors as predictors, respectively: (a) L2 proficiency as measured by the Cambridge test score, (b) average L2 exposure across four modalities (i.e., reading, speaking, listening, writing) and (c) L2 AoA. Each model included: (1) overall n-back accuracy across the four blocks as the dependent variable, (2) main effects of task difficulty (0- vs. 1- vs. 2- vs. 3-back) and bilingual experience factor, as well as their interaction, as predictors, and (3) covariates for age, sex, general intelligence, SES, and maximal educational attainment as factors known to potentially affect working memory performance. The models also included random intercepts for participants to account for repeated measures (the four different n-back blocks). The model specification was as follows:

mixed block_level_accuracy c.L2_proficiency##i.nback_block i.sex c.general_intelligence c.education_years c.age c.SES || participant:

(The above example is for L2 proficiency; replace L2_proficiency with other bilingual experience factors for their respective models).

We first tested models inserting task difficulty as a continuous factor to facilitate assessing the general direction and strength of the effect (see e.g., Long & Freese, 2006; Pasta, 2009; Robitzsch, 2020; Williams, 2020). Once significance was confirmed, we tested the models again with task difficulty (n-back block) as a categorical variable. Estimates and *p*-values for both continuous and categorical models are reported in the respective Results sections below.

2.3.3. Reaction times

Statistical analysis of RTs was similar to that of accuracy. The data was coded in long format with RT of each single trial (across all four n-back blocks) constituting one row in the data matrix. To account for repeated measures, we used a linear-mixed effects approach including random intercepts for participants and random slopes for single trials. We tested three single-trial linear mixed regression models with the same three bilingual experience factors (L2 proficiency, exposure and AoA) used as predictors. Each model included: (1) single-trial n-back RTs across the four blocks as the dependent variable, (2) main effects of task difficulty (0- vs. 1- vs. 2- vs. 3-back) and bilingual experience factor, as well as their interaction, as predictors, and (3) covariates for age, sex, general intelligence, SES and maximal educational attainment. The model also included random intercepts for participants and random slopes for trials. The model specification was as in the following example for L2 proficiency:

mixed single_trial_RT c.L2_proficiency##i.nback_block i.sex c.general_intelligence c.education_years c.age c.SES || participant: trial_number.

We first tested models inserting task difficulty as a continuous factor to assess the general direction and strength of the effect. Once significance was confirmed, we tested the models again with task difficulty as a categorical variable. Estimates and *p*-values for both continuous and categorical models are reported in Results.

2.3.4. Speed/accuracy tradeoff

To assess whether skill development associated with degree of bilingual experience affects the automaticity of working memory performance, we investigated speed/accuracy tradeoffs (SATs) and how they varied with increasing levels of task difficulty. To do this, we calculated SATs as the predictive effect of single-trial accuracy (correct vs. incorrect response) on single-trial RT. We then estimated the interaction between single-trial accuracy, task difficulty and L2 proficiency to test whether individual differences in L2 proficiency modulate the influence of task difficulty on the SAT. Using RTs – and not accuracy – as the dependent variable allowed us to deploy a linear mixed effect model (as RT is a continuous dependent variable) instead of a logit model (single-trial accuracy being a binary variable), which improved the estimation power of our statistical model.

Since only L2 proficiency was significant in the accuracy analyses above (see Results for details), we only tested one single-trial linear mixed regression model with L2 proficiency as the sole exponent of bilingual experience. The model included: (1) single-trial n-back RTs across the four blocks as the dependent variable, (2) main effects of single-trial accuracy, task difficulty (0- vs. 1- vs. 2- vs. 3-back) and L2 proficiency, as well as their interaction, as predictors, and (3) covariates for age, sex, general intelligence, SES and maximal educational attainment. The model also included random intercepts for participants and random slopes for trials. The model specification was as follows:

mixed single_trial_RT i.single_trial_accuracy##c.L2_proficiency##i.nback_block i.sex c.general_intelligence c.education_years c.age i.SES || participant: trial_number.

Similar to the other analyses, we first tested the model using task difficulty as a continuous factor to assess the general direction and strength of the effect, and then, after confirming significance, we tested the model again with task difficulty as a categorical variable. Estimates and *p*-values for both continuous and categorical models are reported in Results.

3. Results

3.1. Accuracy

The model revealed a significant interaction between L2 proficiency and task difficulty, when measured both as a continuous factor ($\beta = 0.23$; $p = 0.017$) and as a categorical factor, indicating that varying L2 proficiency differentially affects n-back accuracy at different levels of difficulty. Relative to 0-back performance, L2 proficiency showed a stronger positive association with accuracy for both 2-back ($\beta = 0.676$; $p = 0.019$) and 3-back ($\beta = 0.603$; $p = 0.037$). The interaction plot(s) confirmed that (i) the beneficial effect of L2 proficiency on n-back accuracy was reliably stronger for higher (2- and 3-back) as compared to lower (0- and 1-back) task difficulty levels (see Fig. 1a) and (ii) the overall decrease of accuracy with increasing task difficulty was less pronounced for higher levels of L2 proficiency (see Fig. 1b). No significant interactions emerged for L2 exposure and L2 AoA.

3.2. Reaction times

Models revealed a significant interaction between L2 proficiency and L2 AoA and task difficulty, measured both as a continuous and as a categorical factor, indicating that the degree of bilingual experience differentially affects n-back RTs at different levels of difficulty. Namely, for L2 proficiency, the interaction with continuous task difficulty was significant ($\beta = -1.659$; $p = 0.001$) and the interaction with categorical task difficulty (with 0-back as referent) was significant for 3-back ($\beta = -6.708$; $p < 0.001$). For L2 AoA, the interaction with continuous task difficulty was significant ($\beta = 4.312$; $p < 0.001$) and the interaction with categorical task difficulty (with 0-back as referent) was significant for the 2-back ($\beta = 7.402$; $p = 0.004$) and 3-back ($\beta = 13.252$; $p < 0.001$) conditions. For L2 proficiency, the interaction plots revealed that (i) a beneficial effect of degree of bilingual experience (i.e., higher L2 proficiency) on n-back RTs emerged only at the highest level of task difficulty (3-back, see Fig. 2a) and (ii) the trajectory of RT increase with increasing task difficulty was less pronounced for higher levels of L2 proficiency when moving from 2- to 3-back (see Fig. 2c). Similarly, for L2 AoA the interaction plots revealed that (i) a beneficial effect of degree of bilingual experience (i.e., earlier L2 AoA) on n-back RTs emerged only at higher levels of task difficulty (2- and 3-back, see Fig. 2b) and (ii) the trajectory of RT increase with increasing task difficulty was less pronounced for earlier L2 AoA when moving from 1- to 2- back and from 2- to 3-back (see Fig. 2d). For L2 exposure, the interaction was significant for continuous, but not for categorical, task difficulty, preventing its further analyses.

3.3. Speed/accuracy tradeoff

The model revealed a significant interaction between L2 proficiency, task difficulty and single-trial accuracy, measured both as a continuous ($\beta = -7.6$; $p < 0.001$) and as a categorical factor (0-back vs. 3-back $\beta = -16.369$; $p = 0.005$), indicating that varying L2 proficiency differentially affected the SAT at low and high levels of difficulty. The interaction plots revealed that the expected speed/accuracy relationship (i.e., faster = less accurate) inverted at higher levels of difficulty, potentially indicating skill-related effects in working memory processes (see the Discussion section below). Interestingly, the level of difficulty, at which this inversion manifested, diverged between different proficiency levels. Namely, low-proficient bilinguals exhibited this inversion already in the 1-back condition, maintained it for the 2-back, but returned to the expected SAT in the high cognitive load condition (i.e., 3-back). High-proficient bilinguals, instead, seemed to delay the change of strategy to higher levels of difficulty (i.e., 2-back) and maintained it throughout the entire experiment, including the 3-back block (see Fig. 3a-d).

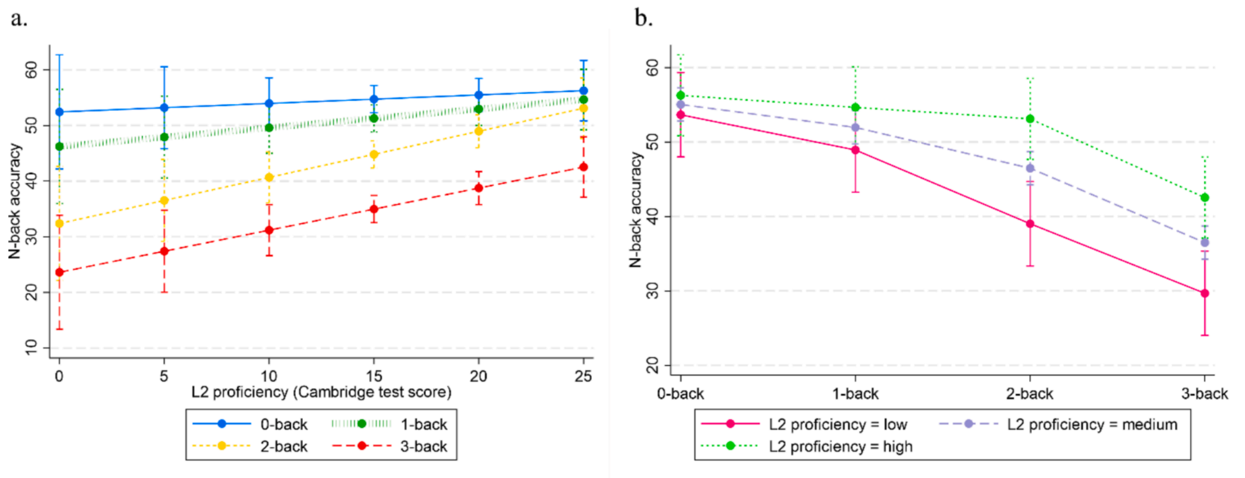

Fig. 1. Accuracy analyses. Interaction plot for the L2 proficiency*task difficulty interaction predicting n-back accuracy. a. The beneficial effect of L2 proficiency on n-back accuracy was stronger for higher task difficulty (2- and 3- back) as compared to lower task difficulty (0- and 1-back). b. The overall decrease of accuracy with increasing task difficulty was less pronounced for higher levels of L2 proficiency. For plotting purposes, the three levels of L2 proficiency, i.e., low, medium and high, represent 1σ below the mean value, mean value, and 1σ above the mean value of the sample, respectively.

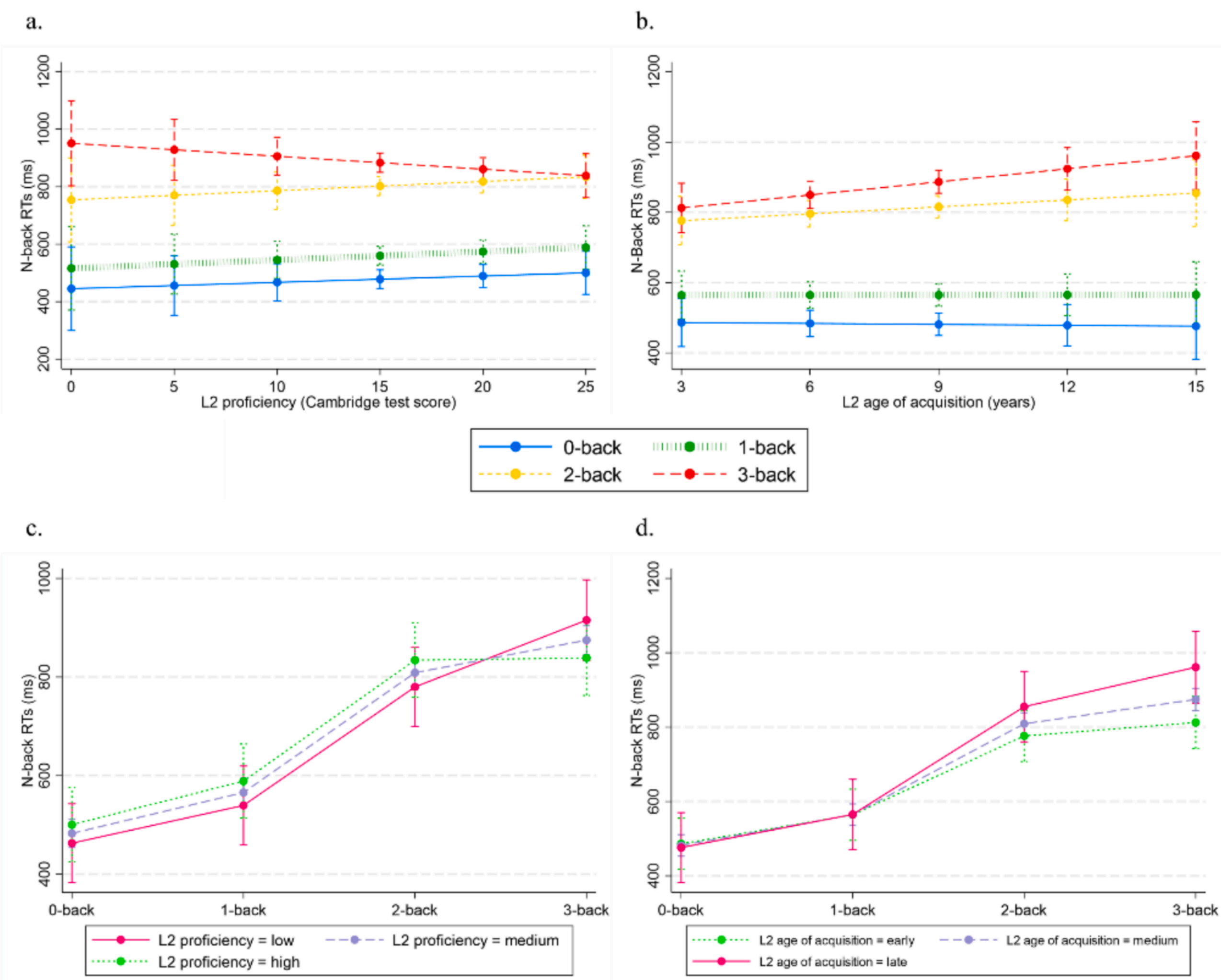

Fig. 2. RT analyses. Interaction plot for bilingual experience factor*task difficulty interactions predicting n-back RTs. a. A beneficial effect of higher L2 proficiency on n-back RTs emerged only at higher levels of task difficulty. b. A beneficial effect of earlier L2 AoA on n-back RTs emerged only at higher levels of task difficulty. c. The trajectory of RT increase with increasing task difficulty was less pronounced for higher levels of L2 proficiency. d. The trajectory of RT increase with increasing task difficulty was less pronounced for earlier levels of L2 AoA. For plotting purposes, the three levels of L2 proficiency and L2 age of acquisition, i.e., low, medium and high, represent 1σ below the mean value, mean value, and 1σ above the mean value of the sample, respectively.

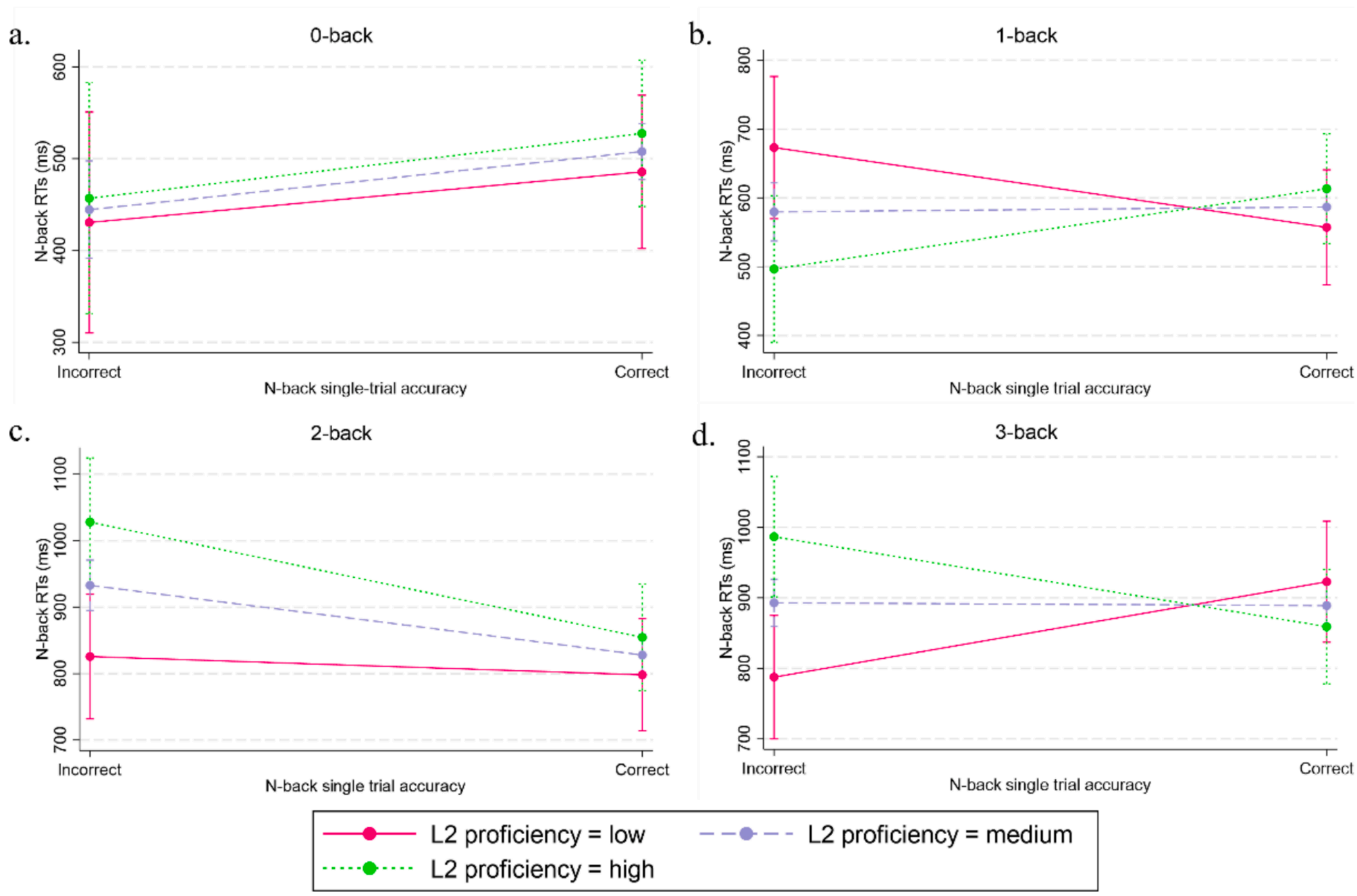

Fig. 3. Speed/accuracy tradeoff analyses. Interaction plot for the three-way interaction between L2 proficiency*task difficulty*single-trial accuracy predicting n-back RTs. Varying L2 proficiency differentially affects the speed/accuracy tradeoff trajectory with increasing task difficulty. For plotting purposes, the three levels of L2 proficiency, i.e., low, medium and high, represent 1σ below the mean value, mean value, and 1σ above the mean value of the sample, respectively.

4. Discussion

The concepts of cognitive efficiency and of expertise-dependent automaticity are central to models detailing bilingualism-induced cognitive adaptations in the executive system (e.g., Bialystok, 2024; Bialystok & Craik, 2022; (DeLuca et al., 2020); Grundy et al., 2017; Pliatsikas, 2020; Stocco et al., 2014). In our current study, we focused on the behavioral manifestations of these concepts. Such behavioral manifestations, especially those related to CE, are intrinsically linked to the possibility of manipulating the cognitive load of tasks used to assess relevant cognitive functions. This manipulation enables individual differences in efficiency to surface in the form of reduced performance decline as task difficulty increases (Neubauer & Fink, 2009; Reuter-Lorenz & Cappell, 2008; Rypma et al., 2006). At the same time, the adoption of more fine-grained and adjustable tasks has been suggested as a potential tool to inform the debate about the inconsistent appearance of bilingualism-induced cognitive effects, particularly in behavioral studies (Kroll & Bialystok, 2013; Leivada et al., 2021; Valian, 2015). Following this approach, here we examined the behavioral counterpart of CE in bilingual young adults by varying the cognitive load imposed by a working memory task to strain their cognitive resources and thereby address the putative cognitive impacts of bilingualism in a more graded fashion. Participants completed versions of the n-back task ranging from 0- to 3-back, with the 3-back condition widely recognized as presenting a highly demanding cognitive challenge. We investigated whether individual differences in the degree of bilingual experience could predict performance measures—such as accuracy, RT, and SAT—as task difficulty and cognitive load increased. Our RT and accuracy analyses showed an interactive contribution of degree of bilingual experience and task difficulty to the participants' n-back performance, indexing bilingualism-related consequences for CE. Moreover, our SAT analyses revealed that the relationship between SAT and task difficulty is calibrated to the degree of bilingual experience, hinting at bilingualism-induced consequences for expertise-dependent automaticity as well. Below we discuss these two sets of results and their implications in more detail.

4.1. RT and accuracy analyses

Our RT and accuracy analyses revealed that the positive contribution of bilingual experience to working memory performance became stronger with increasing levels of task-induced cognitive load, and that performance decreased to a lesser extent for more expert bilinguals when task difficulty increased. These results carry dual theoretical importance. First, they contribute to the debate on the inconsistent appearance of bilingualism's effects on cognition (see e.g., Antoniou, 2019; Baum & Titone, 2014; Bialystok, 2017). Namely, both patterns point to task sensitivity as a potential reason for failures to observe bilingual effects on cognition in some studies, particularly in behavioral investigations on young adults, who are typically at peak cognitive condition and live highly cognitively stimulating lives (Kroll & Bialystok, 2013; Leivada et al., 2021; Valian, 2015). To put it simply, if an experimental manipulation does not allow for driving the cognitive load to highly challenging levels, the sample's performance can simply be at (near-)ceiling, limiting the variability in obtained measures and rendering the design statistically insensitive to the targeted effects. In other words, participants' cognitive resources, especially in young adult populations, need to be strained enough to effectively probe the volume of each individual's "cognitive warehouse" and allow linking it to their bilingual parameters. Our results support this view: the contribution of bilingual experience to working memory performance emerged more strongly (in the case of accuracy, see Fig. 1a) or even exclusively (in the case of RTs, see Fig. 2a and 2b) at higher levels of difficulty. For example, had we tested the impact of L2 proficiency on RTs in a 0-back, 1-back or 2-back tasks, we would have failed to observe any effects. Only by increasing the working memory load to 3-back were we able to drive the system to the levels when bilingualism-related benefits could be revealed. This argument echoes and extends the one put forward by Pereira Soares et al. (2024) in a recent study also deploying an n-back task. In their investigation, which only included 0- to 2-back conditions, Pereira Soares and colleagues observed relationships between degree of bilingual experience and n-back performance at the brain level (measured via on-task EEG) but not at the behavioral level. The authors suggest that the failure to observe behavioral effects may be ascribed to the absence of a sufficiently difficult condition, i.e., a 3-back. Not only are our results are in line with this interpretation but they also extend and confirm those findings by demonstrating the effect emergence at the 3-back level.

The importance of manipulating task difficulty for detecting cognitive effects of bilingualism in young adults is also supported by the differential drop in performance with increased difficulty. For example, when investigating the modulatory effect of degree of bilingual experience on RT variations with increasing task difficulty, differential trajectories mainly (or only in the case of L2 Proficiency) emerged between 2-back and 3-back conditions (see Fig. 2c/2d). Once again, we would have simply failed to observe such differential trajectories, had we not included a 3-back condition in our study. Furthermore, these results are in line with a previous investigation by Comishen and Bialystok (2021) who also used an n-back task of increasing difficulty with young bilingual and monolingual groups. The authors found that while accuracy was similar across groups in low-difficulty conditions, bilinguals outperformed monolinguals at high-difficulty conditions. Here, we effectively not only replicate these behavioral findings but also verify and extend them using a continuous operationalization of bilingualism in a single group, rather than a dichotomous variable (see e.g. Leivada et al., 2021; Luk & Bialystok, 2013 for discussion on this topic).

Furthermore, our results suggest that bilingual experience may foster the accrual of CE, at least as manifested in its behavioral indices analyzed here. Indeed, CE is expected to manifest as smaller performance decreases at the behavioral level in the face of increasing task difficulty. While our findings may in principle parallel the associated concept of "neural efficiency" (i.e., lower degrees of activation being required to achieve a given level of performance, Neubauer & Fink, 2009; Reuter-Lorenz & Cappell, 2008; Rypma et al., 2006), we cannot make this link directly in the absence of neuroimaging data from our participants. However, it is at least logical to assume that, by requiring less resources to perform at low difficulty levels, individuals with higher neural efficiency may be able to maintain extra resources to deploy at more complex stages of a task, resulting in mitigated deterioration of performance standards in highly demanding conditions. Moreover, enhanced efficiency of the executive control network has been proposed as a key outcome of neuroplastic adaptations resulting from bilingual experience (Abutalebi & Green, 2007; DeLuca et al., 2020; Green & Abutalebi, 2013; Pliatsikas, 2020; Stocco et al., 2014).

Therefore, our present results, although limited to a behavioral investigation, generally agree with predictions at the neural level and their purported behavioral manifestations. They show that increasing degrees of bilingual experience mitigate performance deterioration at higher levels of task difficulty– the standard behavioral manifestation of increased CE. Further corroborating this view is the finding that the contribution of bilingual experience to working memory performance emerges more intensely at high difficulty levels, when fewer cognitive resources are typically available. In other words, when cognitive resources dwindle, expert bilinguals separate themselves from the less expert ones even more clearly in terms of performance. This line of reasoning is corroborated by the pattern of findings emerging in the abovementioned study by Comishen and Bialystok (2021). Beyond behavioral results similar to those emerging in the present investigation, the authors also reported EEG signatures (larger P2/P3 amplitudes across conditions, and earlier P3 peaks during 3-back) of less effortful processing by bilinguals, as compared to monolinguals.

An intriguing aspect of our findings is that, at least with respect to accuracy measures, not all dimensions of bilingual experience appeared to significantly influence performance. Specifically, while L2 proficiency emerged as a significant predictor, neither L2 AoA nor L2 exposure contributed meaningfully to our models. One plausible explanation is that, within the sociolinguistic context sampled, these variables may reflect partially distinct facets of bilingual experience. Our participants were L1-dominant individuals living in an L1-dominant environment, with variable, individually determined opportunities to engage with their L2. In such a context, the relationship between the *quantity* of bilingual experience (as indexed by L2 exposure) and its *quality* (as indexed by L2 proficiency) may be modulated by individual differences in (1) the meaningfulness of L2 engagement and (2) the pace of individual learning. Under these conditions, L2 proficiency may better capture the qualitative aspects of bilingual experience and, consequently, account for a larger portion of variance in cognitive performance than L2 exposure.

L2 AoA, in this framework, occupies a somewhat intermediary position. On the one hand, it can reflect acquisition during periods of heightened neurocognitive plasticity—suggesting a qualitative component—yet, beyond a certain threshold, it primarily indexes the duration of bilingual experience, a more quantitative measure. This dual nature may explain why AoA was associated with response speed (i.e., reaction times in the n-back task) but not with accuracy.

These interpretations should be understood in light of the specific sociolinguistic context of our sample. In environments with more sustained and intensive L2 exposure (e.g., immersion settings), bilingualism-related cognitive adaptations may surpass the threshold at which L2 proficiency alone explains variance in performance (see DeLuca et al., 2019; Pereira Soares et al., 2022). Further empirical research directly comparing such distinct sociolinguistic contexts is warranted to test this hypothesis.

4.2. Speed/accuracy tradeoff analyses

Our second main finding relates to the bilingualism-related modulation of the SAT trajectory unfolding as task demands increase. Our SAT analyses show that the trajectory of strategies adopted to cope with increasing working memory demands had different dynamics for different levels of bilingual experience. Nonetheless, irrespective of the degree of bilingual experience, the expected speed/accuracy relationship (i.e., more speed = less accuracy) appeared to invert at some level of task difficulty. On the one hand, this result may be interpreted as indexing a "fine tuning" of bilinguals' working memory during task progression (e.g., Abutalebi et al., 2012; DeLuca, Rothman, Bialystok, & Pliatsikas, 2020). On the other hand, it may be framed in terms of expertise-dependent automaticity (Beilock et al., 2008; Mann et al., 2007; Moors & De Houwer, 2006) and the so-called *flow state* (Csikszentmihalyi, 2000; Csikszentmihalyi et al., 2014; Engeser & Rheinberg, 2008; Harris et al., 2017; Jackson et al., 2001).

In our study, the speed/accuracy relationship was inverted at some point for all degrees of bilingual experience, but what is more interesting is the progression of this trajectory, which appeared to be related to the degree of bilingual experience. For lower L2 proficiency levels, the inversion emerged already in the 1-back condition, persisted for the 2-back, but returned to expected SAT (more speed = less accuracy) in the highest cognitive load condition (3-back). In high-proficiency bilinguals, instead, this inversion appeared at a higher level of difficulty (2-back) and persisted through the 3-back block (see Fig. 3a-d).

Such an inversion of the traditional SAT has been documented for individuals who are highly skilled in a given task (e.g., Beilock et al., 2002, 2008; Perkins-Ceccato et al., 2003). For example, Beilock and colleagues (2008) showed that skilled golf players performed better when being instructed to perform as fast *and* as accurately as possible, compared to when being instructed to take as much time as needed to be accurate. The opposite pattern emerged for novices, who performed better when they were slower. This pattern is traditionally explained in terms of differing levels of *automaticity* when performing a task for skilled vs. novice performers. Because performance is more automated for skilled performers, it declines in situations that require step-by-step control. Novice performance, which requires more attentional resources, typically benefits from such step-by-step instructions (e.g., Castaneda & Gray, 2007; Ford et al., 2005; Gray, 2004; Jackson et al., 2006).

Indeed, bilinguals' executive system has also been posited to become more automated as a function of extensive training (DeLuca et al., 2020; Grundy et al., 2017; Pliatsikas, 2020). Among other studies, corroboration for these claims comes from the abovementioned EEG investigation by Comishen and Bialystok (2021), who deployed a similar design to the one adopted in the present study. Indeed, they found that bilinguals, compared to monolinguals, displayed earlier P3 peak latency during 3-back performance, indicating earlier recruitment of attentional resources at high difficulty stages.

Altogether, these observations may explain the present finding of reversed SAT in our participants. It may also be useful to refer to the concept of the *flow state* (see e.g., Csikszentmihalyi et al., 2014), tightly linked with that of automaticity of cognitive performance (Harris et al., 2017). Indeed, the flow state is frequently defined as a state of enhanced automaticity or an absence of controlled attention during cognitive performance (Harris et al., 2017), in which increased task demands are faced with no increase in effort (Csikszentmihalyi, 2000) and peak performance can thus be achieved (Jackson et al., 2001). For the flow state to be triggered, the individual has to perceive that the challenge imposed by a task and their skill are both at high levels (Engeser & Rheinberg, 2008). If the personal skill related to a task is perceived as high, but the task is perceived as easy, the flow theory predicts a state of relaxation and disengagement from automated performance. Similarly, if the level of task complexity is perceived as high but the personal skill is not, the flow state cannot be triggered or maintained. Our results fit this framework well. For instance, it could be argued that on the 0-back level, no participant would enter the flow state, as the task was perceived as too easy. Bilinguals with lower L2 proficiency, who have lower degrees of bilingual experience and thus a working memory system trained to a lower extent, could possibly enter the flow state already at 1-back. At 3-back, the task became too difficult given the cognitive resources available to these individuals, and thus they disengaged from the flow state, returning to show a typical SAT. Highly expert bilinguals, instead, are hypothesized to have a more trained working memory system; as a result, their entrance in the flow state is delayed by one difficulty level, showing a reversal of the SAT at 2- and 3-, but not at 1-back. That said, the present experiment was not geared towards testing the flow state as such, and this putative interpretation remains to be elaborated in further studies.

5. Conclusions

Taken together, our findings suggest that the degree of bilingual experience is associated with behavioral manifestations compatible with increased cognitive efficiency, in the form of improved maintenance of working memory performance level under high task demands, with stronger relative contributions of accumulated bilingual experience when task difficulty increases. Moreover, they indicate a relationship between degree of bilingual experience and expertise-dependent automaticity, as indexed by different speed-accuracy tradeoff trajectories with increasing task difficulty. While our findings are limited to the behavioral manifestations of cognitive efficiency and automaticity, future neuroimaging studies will need to confirm and expand the theoretical interpretations of results emerging from this and similar investigations.

CRediT authorship contribution statement

Federico Gallo: Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Liliia Terekhina:** Writing – review & editing, Investigation, Data curation. **Jubin Abutalebi:** Writing – review & editing, Supervision, Conceptualization. **Yury Shtyrov:** Writing – review & editing, Supervision, Conceptualization. **Andriy Myachykov:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is an output of a research project implemented as part of the Basic Research Program at the National Research University Higher School of Economics. FG has received funding from the European Union's Horizon Europe research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 101106069. JA acknowledges funding by the Trond Mohn Foundation, grant No. TMS2023UiT01.

Data availability

Data will be made available on request.

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. https://doi.org/10.1093/cercor/bhr287

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics, 20*(3), 242–275. https://doi.org/10.1016/j.jneuroling.2006.10.003

Abutalebi, J., & Green, D. W. (2016). Neuroimaging of language control in bilinguals: Neural adaptation and reserve. *Bilingualism, 19*(4), 689–698. https://doi.org/10.1017/S1366728916000225

Anderson, J. A. E., Grundy, J. G., Grady, C. L., Craik, F. I. M., & Bialystok, E. (2021). Bilingualism contributes to reserve and working memory efficiency: Evidence from structural and functional neuroimaging. *Neuropsychologia, 163*, Article 108071. https://doi.org/10.1016/j.neuropsychologia.2021.108071

Antoniou, M. (2019). The advantages of bilingualism debate. *Annual Review of Linguistics, 5*(1), 395–415.

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3*(2), 12–28.

Baum, S., & Titone, D. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Applied Psycholinguistics, 35*(5), 857–894. https://doi.org/10.1017/S0142716414000174

Beilock, S. L., Bertenthal, B. I., Hoerger, M., & Carr, T. H. (2008). When does haste make waste? Speed-accuracy tradeoff, skill level, and the tools of the trade. *Journal of Experimental Psychology: Applied, 14*(4), 340.

Beilock, S. L., Carr, T. H., MacMahon, C., & Starkes, J. L. (2002). When paying attention becomes counterproductive: Impact of divided versus skill-focused attention on novice and experienced performance of sensorimotor skills. *Journal of Experimental Psychology: Applied, 8*(1), 6.

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin, 143*(3), 233–262. https://doi.org/10.1037/bul0000099

Bialystok, E. (2024). Bilingualism modifies cognition through adaptation, not transfer. *Trends in Cognitive Sciences*.

Bialystok, E., & Craik, F. I. M. (2022). How does bilingualism modify cognitive function? Attention to the mechanism. *Psychonomic Bulletin & Review, 29*(4), 1246–1269. https://doi.org/10.3758/s13423-022-02057-5

Castaneda, B., & Gray, R. (2007). Effects of focus of attention on baseball batting performance in players of differing skill levels. *Journal of Sport and Exercise Psychology, 29*(1), 60–77.

Comishen, K. J., & Bialystok, E. (2021). Increases in attentional demands are associated with language group differences in working memory performance. *Brain and Cognition, 147*, Article 105658.

Court, J. H., & Raven, J. (1992). *Raven manual: Section 3. The standard progressive matrices*. Oxford: Oxford Psychologists Press.

Csikszentmihalyi, M. (2000). Beyond boredom and anxiety. In *Beyond boredom and anxiety*. (pp. xxx, 231–xxx, 231). Jossey-Bass.

Csikszentmihalyi, M., Abuhamdeh, S., & Nakamura, J. (2014). *Flow BT - Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi* (M. Csikszentmihalyi (ed.); pp. 227–238). Springer Netherlands. Doi: 10.1007/978-94-017-9088-8_15.

De Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive advantage in bilingualism: An example of publication bias? *Psychological Science, 26*(1), 99–107.

DeLuca, V., Rothman, J., Bialystok, E., & Pliatsikas, C. (2020). Duration and extent of bilingual experience modulate neurocognitive outcomes. *NeuroImage, 204*, Article 116222. https://doi.org/10.1016/j.neuroimage.2019.116222

DeLuca, V., Rothman, J., & Pliatsikas, C. (2019). Linguistic immersion and structural effects on the bilingual brain: A longitudinal study. *Bilingualism: Language and Cognition, 22*(5), 1160–1175. https://doi.org/10.1017/S1366728918000883

DeLuca, V., Segaert, K., Mazaheri, A., & Krott, A. (2020). Understanding bilingual brain function and structure changes? U bet! A unified bilingual experience trajectory model. *Journal of Neurolinguistics, 56*, Article 100930.

Dobbs, A. R., & Rule, B. G. (1989). Adult age differences in working memory. *Psychology and Aging, 4*(4), 500.

Duñabeitia, J. A., & Carreiras, M. (2015). The bilingual advantage: Acta est fabula. *Cortex, 73*, 371–372.

Engeser, S., & Rheinberg, F. (2008). Flow, performance and moderators of challenge-skill balance. *Motivation and Emotion, 32*, 158–172.

Ford, P., Hodges, N. J., & Williams, A. M. (2005). Online attentional-focus manipulations in a soccer-dribbling task: Implications for the proceduralization of motor skills. *Journal of Motor Behavior, 37*(5), 386–394.

Gold, B. T., Kim, C., Johnson, N. F., Kryscio, R. J., & Smith, C. D. (2013). Lifelong bilingualism maintains neural efficiency for cognitive control in aging. *Journal of Neuroscience, 33*(2), 387–396. https://doi.org/10.1523/JNEUROSCI.3837-12.2013

Gray, R. (2004). Attending to the execution of a complex sensorimotor skill: Expertise differences, choking, and slumps. *Journal of Experimental Psychology: Applied, 10*(1), 42.

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology, 25*(5), 515–530. https://doi.org/10.1080/20445911.2013.796377

Grimes, D., Tan, D. S., Hudson, S. E., Shenoy, P., & Rao, R. P. N. (2008). Feasibility and pragmatics of classifying working memory load with an electroencephalograph. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 835–844).

Grundy, J. G., Anderson, J. A. E., & Bialystok, E. (2017). Neural correlates of cognitive processing in monolinguals and bilinguals. *Annals of the New York Academy of Sciences, 1396*(1), 183–201. https://doi.org/10.1111/nyas.13333

Haier, R. J., Siegel, B., Tang, C., Abel, L., & Buchsbaum, M. S. (1992). Intelligence and changes in regional cerebral glucose metabolic rate following learning. *Intelligence, 16*(3–4), 415–426.

Harris, D. J., Vine, S. J., & Wilson, M. R. (2017). Chapter 12 - Neurocognitive mechanisms of the flow state. In M. R. Wilson, V. Walsh, & B. B. T.-P. in B. R. Parkin (Eds.), *Sport and the Brain: The Science of Preparing, Enduring and Winning, Part B* (Vol. 234, pp. 221–243). Elsevier. Doi: Doi: 10.1016/bs.pbr.2017.06.012.

Hartshorne, J. K., & Germine, L. T. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychological Science, 26*(4), 433–443.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2–3), 61–83.

Hoffman, B. (2012). Cognitive efficiency: A conceptual and methodological comparison. *Learning and Instruction, 22*(2), 133–144.

Jackson, R. C., Ashford, K. J., & Norsworthy, G. (2006). Attentional focus, dispositional reinvestment, and skilled motor performance under pressure. *Journal of Sport and Exercise Psychology, 28*(1), 49–68.

Jackson, S. A., Thomas, P. R., Marsh, H. W., & Smethurst, C. J. (2001). Relationships between flow, self-concept, psychological skills, and performance. *Journal of Applied Sport Psychology, 13*(2), 129–153.

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology, 25*(5), 497–514.

Kroll, J. F., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica, 128*(3), 416–430. https://doi.org/10.1016/j.actpsy.2008.02.001

Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2015). Bilingualism, Mind, and Brain. *Annual Review of Linguistics, 1*, 377–394. https://doi.org/10.1146/annurev-linguist-030514-124937

Kudo, M., & Swanson, H. L. (2014). Are there advantages for additive bilinguals in working memory tasks? *Learning and Individual Differences, 35*, 96–102.

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. In *Psychological Bulletin* (Vol. 144, Issue 4, pp. 394–425). American Psychological Association. Doi: 10.1037/bul0000142.

Leivada, E., Westergaard, M., Duñabeitia, J. A., & Rothman, J. (2021). On the phantom-like appearance of bilingualism effects on neurocognition: (How) should we proceed? *Bilingualism: Language and Cognition, 24*(1), 197–210. https://doi.org/10.1017/S1366728920000358

Logan, G. D. (1988). Automaticity, resources, and memory: Theoretical controversies and practical implications. *Human Factors, 30*(5), 583–598.

Long, J. S., & Freese, J. (2006). *Regression models for categorical dependent variables using Stata* (Vol. 7). Stata press.

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology (Hove, England), 25*(5), 605–621. https://doi.org/10.1080/20445911.2013.795574

MacArthur Foundation. (2007). *The MacArthur Research Network on Socioeconomic Status and Health*. https://macses.ucsf.edu/research/psychosocial/subjective.php#measurement.

Mann, D. T. Y., Williams, A. M., Ward, P., & Janelle, C. M. (2007). Perceptual-cognitive expertise in sport: A meta-analysis. *Journal of Sport and Exercise Psychology, 29*(4), 457–478.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research, 50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Mattay, V. S., Fera, F., Tessitore, A., Hariri, A. R., Berman, K. F., Das, S., Meyer-Lindenberg, A., Goldberg, T. E., Callicott, J. H., & Weinberger, D. R. (2006). Neurophysiological correlates of age-related changes in working memory capacity. *Neuroscience Letters, 392*(1–2), 32–37.

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin, 132*(2), 297.

Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology, 114*(2), 187–202.

Morrison, C., Kamal, F., & Taler, V. (2019). The influence of bilingualism on working memory event-related potentials. *Bilingualism: Language and Cognition, 22*(1), 191–199.

Neubauer, A. C., & Fink, A. (2009). Intelligence and neural efficiency. *Neuroscience & Biobehavioral Reviews, 33*(7), 1004–1023. https://doi.org/10.1016/j.neubiorev.2009.04.001

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Paap, K. (2022). *The bilingual advantage in executive functioning hypothesis: How the debate provides insight into psychology's replication crisis*. Routledge.

Pasta, D. J. (2009). Learning when to be discrete: Continuous vs. categorical predictors. *SAS Global Forum, 248*, 1–10.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*, 195–203.

Pereira Soares, S. M., Prystauka, Y., DeLuca, V., Botija, C. P. P., & Rothman, J. (2024). Brain Correlates of Attentional Load Processing Reflect Degree of Bilingual Engagement: Evidence from EEG. *NeuroImage, 120786*.

Pereira Soares, S. M., Prystauka, Y., DeLuca, V., & Rothman, J. (2022). Type of bilingualism conditions individual differences in the oscillatory dynamics of inhibitory control. *Frontiers in Human Neuroscience, 16*, Article 910910.

Perkins-Ceccato, N., Passmore, S. R., & Lee, T. D. (2003). Effects of focus of attention depend on golfers' skill. *Journal of Sports Sciences, 21*(8), 593–600.

Pliatsikas, C. (2020). Understanding structural plasticity in the bilingual brain: The Dynamic Restructuring Model. *Bilingualism: Language and Cognition, 23*(2), 459–471. https://doi.org/10.1017/S1366728919000130

Pyc, M. A., & Rawson, K. A. (2007). Examining the efficiency of schedules of distributed retrieval practice. *Memory & Cognition, 35*(8), 1917–1927.

Reuter-Lorenz, P. A., & Cappell, K. A. (2008). Neurocognitive Aging and the Compensation Hypothesis. *Current Directions in Psychological Science, 17*(3), 177–182. https://doi.org/10.1111/j.1467-8721.2008.00570.x

Robitzsch, A. (2020). Why ordinal variables can (almost) always be treated as continuous variables: Clarifying assumptions of robust continuous and ordinal factor analysis estimation methods. *Frontiers in Education, 5*, Article 589965.

Rypma, B., Berger, J. S., Prabhakaran, V., Martin Bly, B., Kimberg, D. Y., Biswal, B. B., & D'Esposito, M. (2006). Neural correlates of cognitive efficiency. *NeuroImage, 33*(3), 969–979. https://doi.org/10.1016/j.neuroimage.2006.05.065

StataCorp.. (2021). *Stata Statistical Software: Release 17*. College Station, TX: StataCorp LLC.

Stilley, C. S., Bender, C. M., Dunbar-Jacob, J., Sereika, S., & Ryan, C. M. (2010). The impact of cognitive function on medication management: Three studies. *Health Psychology, 29*(1), 50.

Stocco, A., Yamasaki, B., Natalenko, R., & Prat, C. S. (2014). Bilingual brain training: A neurobiological framework of how bilingual experience improves executive function. *International Journal of Bilingualism, 18*(1), 67–92. https://doi.org/10.1177/1367006912456617

Tao, L., Wang, G., Zhu, M., & Cai, Q. (2021). Bilingualism and domain-general cognitive functions from a neural perspective: A systematic review. *Neuroscience & Biobehavioral Reviews, 125*, 264–295.

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition, 18*(1), 3–24. https://doi.org/10.1017/S1366728914000522

Vernon, P. A. (1983). Speed of information processing and general intelligence. *Intelligence, 7*(1), 53–70.

Williams, R. A. (2020). Ordinal Independent Variables. In P. Atkinson, S. Delamont, A. Cernat, J. W. Sakshaug, & R. A. Williams (Eds.), *SAGE Research Methods Foundations*. SAGE Publications Ltd.. https://doi.org/10.4135/9781526421036938055