**Appendix S1. Additional analyses and methodological information**

1. **Additional information on participants’ language background**

Forty-five multilinguals reported that at least one of their parents had a non-Greek language as their native or dominant language or that their parents had balanced proficiency in Greek (Cypriot Greek and/or Standard Modern Greek) and a non-Greek language. In this respect, these multilinguals had some experience with a non-Greek language from the family. Moreover, they all self-reported that they started being exposed to/using the non-Greek language from or before the age of five. Five multilinguals grew up in a native Cypriot Greek-speaking family. Of these (1) two had just completed an English-instructed school in Cyprus (for seven years), (2) two had attended an English-instructed school in Cyprus (for seven years) and were undergraduate students in the United Kingdom; and (3) one was born in an English-speaking country, had an English-speaking spouse, had studied English and French in a country outside Cyprus and was teaching English after graduation[[1]](#footnote-1). Two of the latter multilinguals indicated that they started being exposed to/using English at a relatively late age, at six and seven years.

Bidialectals were native Greek Cypriots who spoke Cypriot Greek (CG) and Standard Modern Greek (SMG), while “monolinguals” were Hellenic Greeks who came to Cyprus for work or studies. At the group level, bidialectals and “monolinguals” indicated some experience with non-Greek languages. This was expected because most, if not all, European (including Cypriot and Greek) public schools introduce one or more foreign (or, generally, additional) language(s) during primary education (e.g., Eurostat, 2022). Moreover, globally, including Cyprus and Greece, English is widely used in various contexts; particularly, the internet, social media, and television. Bidialectals and “monolinguals” were studying or had studied in Greek-instructed university programmes, apart from four cases; specifically, two bidialectals and two “monolinguals” reported that they were studying/had studied subjects related to foreign languages at undergraduate level, including Turkish Studies (one participant), English Language and Literature (two participants), and Modern Languages and European Studies (one participant). These participants were retained in the bidialectal and “monolingual” groups because they did not have rich, sustained and continuous use of a non-Greek language throughout their life and, particularly, before adulthood (e.g., from a parent who was a native/dominant speaker of a non-Greek language or through school instruction in a non-Greek language), apart from the few hours of additional-language learning that is typical in most European schools. In any case, the experience that bidialectals and “monolinguals” had with non-Greek languages (if any) was captured by their scores from the Language Background and Socioeconomic Status Questionnaire for (1) age of onset of exposure to/use of, (2) overall use of, and (3) proficiency in English and other non-Greek languages. These scores did not significantly differ between bidialectals and “monolinguals” but both groups had significantly less experience with non-Greek languages than multilinguals. Finally, “monolinguals”, as a group, had some limited CG experience given that they lived in Cyprus at the time of testing. However, this, again, was significantly lower than both multilinguals and bidialectals.

Table S1 presents descriptive statistics and group comparisons on participants’ language characteristics based on the Language Background and Socioeconomic Status Questionnaire.

|  |
| --- |
| Table S1.*Participants’ language characteristics based on the Language Background and Socioeconomic Status Questionnaire.* |
|  | “Monolinguals” | Bidialectals | Multilinguals | Statistic |
|  | *N*1 | *N*1 | *N*1 |  |
| Language proficiency |  |  |  |  |
| CG  | No proficiencyLimited proficiencyModerate proficiencyGood proficiencyVery good proficiency | 21012142 | 0001064 | 2231031 | *H*(2) = 80.82, *p* < .001Mult. > Monol. Mult. < Bid. Monol. < Bid. |
| SMG | No proficiencyLimited proficiencyModerate proficiencyGood proficiencyVery good proficiency | 000344 | 0001460 | 003839 | *H*(2) = 5.136, *p* = .077Mult. = Monol. Mult. = Bid.Monol. = Bid. |
| English | No proficiencyLimited proficiencyModerate proficiencyGood proficiencyVery good proficiency | 068249 | 05143421 | 0071527 | *H*(2) = 14.16, *p* = .001Mult. > Monol. Mult. > Bid. Monol. = Bid.  |
| Other | No proficiencyLimited proficiencyModerate proficiencyGood proficiencyVery good proficiency | 341031 | 711830 | 036921 | *H*(2) = 40.76, *p* < .001Mult. > Monol. Mult. > Bid. Monol. = Bid. |
|  | Mean (*SD*) | Mean (*SD*) | Mean (*SD*) |  |
| Language use |  |  |  |  |
| CG | 0.62 (0.74) | 3.62 (0.39) | 2.51 (1.18) | *F*(2, 157) = 182.4, *p* < .001Mult. > Monol. Mult. < Bid. Monol. < Bid. |
| SMG | 3.83 (0.19) | 2.58 (0.76) | 2.58 (0.96) | *F*(2, 169) = 51.82, *p* < .001Mult. < Monol. Mult. = Bid. Monol. > Bid. |
| English | 1.59 (0.68) | 1.67 (0.66) | 2.22 (0.89) | *F*(2, 168) = 11.12, *p* < .001Mult. > Monol. Mult. > Bid. Monol. = Bid. |
| Other | 1.08 (1.07) | 1.27 (1.04) | 2.54 (1) | *F*(2, 92) = 19.81, *p* < .001Mult. > Monol. Mult. > Bid. Monol. = Bid. |
|  |  |  |  |  |
| Onset of language exposure/use (in months after birth) |
| CG | 224.86 (91.1) | 1.35 (6.2) | 47.5 (69.2) | *F*(2, 134) = 143.2, *p* < .001Mult. < Monol. Mult. > Bid. Monol. > Bid. |
| SMG | 0 (0) | 17.8 (27.6) | 45.5 (56.9) | *F*(2, 152) = 19.17, *p* < .001Mult. > Monol. Mult. > Bid. Monol. < Bid. |
| English | 89.5 (22.5) | 79.03 (37.44) | 58.98 (60.1) | *F*(2, 137) = 4.92, p = .009Mult. < Monol. Mult. = Bid. Monol. = Bid. |
| Other | 159.60 (41.49) | 185.36 (81.39) | 28.6 (89) | *F*(2, 55) = 21.24, *p* < .001Mult. < Monol. Mult. < Bid. Monol. = Bid. |
| *Note.* *N*: Number, *SD*: Standard Deviation, CG: Cypriot Greek, SMG: Standard Modern Greek, Other: any other non-Greek language, Mult.: Multilinguals, Monol.: “Monolinguals”, Bid.: Bidialectals, >: group on the left has a significantly higher score than the group on the right (Bonferroni correction applied), <: group on the left has a significantly lower score than the group on the right (Bonferroni correction applied), =: no significant difference between the group on the left and the group on the right (Bonferroni correction applied). 1 Total sum of participants for each language/dialect may not coincide with the total number of participants tested for each group (50 multilinguals, 79 bidialectals, 52 “monolinguals”) because not all participants provided a response to the relevant items of the questionnaire.  |

1. **Schematic illustrations of the Flanker, Stroop, Colour-Shape, Number-Letter, and N-Back executive function tasks**

Figure S1. The Flanker task.



*Note.* FS= Fixation Start duration, R= Random number, RT= Reaction Time, FE= Fixation End duration.

Figure S2. The Stroop task.



*Note.* RT= Reaction Time, R= Random number.

Figure S3. The Colour-Shape task.

**

*Note.* RT= Reaction Time, R= Random number.

Figure S4. The Number-Letter task.



*Note.* RT= Reaction Time, R= Random number.

Figure S5. The N-Back task.



1. **Bivariate correlations between reaction times and accuracy in the critical conditions (incongruent, switch) of the inhibition and switching tasks**

|  |
| --- |
| Table S2.*Bivariate correlations (Pearson’s r) between reaction times and accuracy in the critical conditions (incongruent, switch) of the inhibition and switching tasks.*  |
|  | CS RT | NL RT | Stroop RT | Flanker RT | CS Acc. | NL Acc. | Stroop Acc. |
| NL RT | .557\*\* |  |  |  |  |  |  |
| Stroop RT | .414\*\* | .564\*\* |  |  |  |  |  |
| Flanker RT | .218\*\* | .398\*\* | .372\*\* |  |  |  |  |
| CS Acc. | -.294\*\* | -.311\*\* | -.316\*\* | -.321\*\* |  |  |  |
| NL Acc. | -.258\*\* | -.471\*\* | -.344\*\* | -.191\* | .328\*\* |  |  |
| Stroop Acc. | -.048 | -.073 | -.148\* | -.069 | .341\*\* | .182\* |  |
| Flanker Acc. | .022 | -.114 | -.055 | -.129 | .084 | .351\*\* | .170\* |
| \*\*. Correlation is significant at the 0.01 level (two-tailed). \*. Correlation is significant at the 0.05 level (two-tailed).*Note.* RT: mean reaction times for correct trials, CS RT: RT for switch trials in the Colour-Shape task, NL RT: RT for switch trials in the Number-Letter task, Stroop RT: RT for incongruent trials in the Stroop task, Flanker RT: RT for incongruent trials in the Flanker task, CS Acc.: proportion of correct responses for switch trials in the Colour-Shape task, NL Acc.: proportion of correct responses for switch trials in the Number-Letter task, Stroop Acc.: proportion of correct responses for incongruent trials in the Stroop task, Flanker Acc.: proportion of correct responses for incongruent trials in the Flanker task. |

|  |
| --- |
| 1. **Descriptive statistics by task and condition for the executive function tasks**

Table S3.*Descriptive statistics by task and condition for the executive function tasks (collapsed across the three groups).* |
| Working Memory |
| N-Back | Target | Lure | Non-target | N2 (Target) | N3 (Target) | N4 (Target) | RT Statistic | Accuracy Statistic |
|  | Acc. (*SD*) | Acc. (*SD*) | Acc. (*SD*) | Acc. (*SD*) | Acc. (SD) | Acc. (*SD*) |  |  |
|  | .54 (.16) | .72 (.12) | .87 (.10) | .64 (.17) | .50 (.18) | .49 (.20) | n. a. | *F*(1.3, 224.995) = 334.32, *p* < .001Target < LureTarget < Non-targetLure < Non-target*F*(2, 342) = 91.71, *p* < .001N2 > N3N2 > N4N3 = N4 |
| Rotation Span | Partial score (*SD*)1  | Distractor task Acc. (*SD*) |  |  |  |  |  |  |
|  | 25.7 (7.2) | .92 (.09) |  |  |  |  | n. a. | n. a. |
| Corsi blocks | Forward | Backward |  |  |  |  |  |  |
|  | Partial score (*SD*)2 | Partial score (*SD*)2 |  |  |  |  |  |  |
|  | 44.3 (12.9) | 46.6 (11.7) |  |  |  |  | n. a. | *F*(1, 178) = 5.2, *p* = .024Forward < Backward |
| Switching |
|  | Switch | Repeat | Pure |  |  | RT Statistic | Accuracy Statistic |
| RT (*SD*) | Acc. (*SD*) | RT (*SD*) | Acc. (*SD*) | RT (*SD*) | Acc. (*SD*) |
| Number-Letter task | 1231 (403) | .92 (.09) | 1028 (340) | .94 (.08) | 744 (172) | .96 (.08) |  |  | *F*(1.4, 241.3) = 374.04, *p* < .001Switch > RepeatSwitch > PureRepeat > Pure | *F*(1.2, 213.5) = 25.08, *p* < .001Switch < RepeatSwitch < PureRepeat < Pure  |
| Colour-Shape task | 1766 (443) | .87 (.16) | 1247 (344) | .91 (.16) | 937 (333) | .92 (.19) |  |  | *F*(1.7, 290.96) = 549.2, *p* < .001Switch > RepeatSwitch > PureRepeat > Pure | *F*(1.4, 242.2) = 15.4, *p* < .001Switch < RepeatSwitch < PureRepeat = Pure |
| Inhibition |
|  | Incongruent | Congruent | Neutral |  |  |  |
|  | RT (*SD*) | Acc. (*SD*) | RT (*SD*) | Acc. (*SD*) | RT (*SD*) | Acc. (*SD*) |  |  |  | RT Statistic | Accuracy Statistic |
| Stroop task | 816 (184) | .95 (.08) | 693 (134) | .98 (.05) | 699 (126) | .97 (.06) |  |  |  | *F*(1.5, 264.2) = 277.2, *p* < .001Incongruent > CongruentIncongruent > NeutralCongruent = Neutral | *F*(1.2, 210.3) = 16.7, *p* < .001Incongruent < CongruentIncongruent < NeutralCongruent > Neutral |
| Flanker task | 606.2 (97) | .97 (.03) | 547 (82) | .99 (.01) | 533 (72) | .99 (.02) |  |  |  | *F*(1.4, 246.9) = 423, *p* < .001Incongruent > CongruentIncongruent > NeutralCongruent > Neutral | *F*(1.3, 218.6) = 44, *p* < .001Incongruent < CongruentIncongruent < NeutralCongruent > Neutral |
| 1Maximum score was 42.2Maximum score was 88.*Note.* Acc.: proportion correct, RT: mean reaction times for correct trials, *SD*: Standard Deviation, >: score on the left is significantly higher than score on the right (Bonferroni correction applied), <: score on the left is significantly lower than score on the right (Bonferroni correction applied), =: no significant difference between the score on the left and the score on the right (Bonferroni correction applied).  |

1. **Reliability and convergent validity of target executive function measures and of the difference scores from inhibition and switching tasks**

In this section, we discuss the reliability and convergent validity of all executive function (EF) measures, with a particular focus on the reaction-time (RT) difference scores from the inhibition and switching tasks. We also explore if and the conditions under which these might show better and/or adequate reliability and convergent validity. Moreover, based on these reliability and convergent validity results, in the following sections of the Supplementary Materials Online (SMO), we report additional (group and other) analyses on EF composite scores that used difference scores. To foreshadow the results, these analyses provide further, complementary evidence that the multilingual/bidialectal effect lies in EF and not in simple processing speed. This is because those difference scores that do not suffer from the limitations outlined in the *Introduction* (main manuscript) better isolate and reflect EF processes.

For difference scores, we examined the reliability and convergent validity of two measures from each inhibition and switching task: RT incongruent-RT neutral and RT incongruent-RT congruent for the Flanker and Stroop tasks (Flanker interference effect 1, Stroop interference effect 1, Flanker interference effect 2, and Stroop interference effect 2, respectively); and RT switch-RT pure and RT switch-RT repeat for the Number-Letter and Colour-Shape tasks (Number-Letter switch cost 1, Colour-Shape switch cost 1, Number-Letter switch cost 2 and Colour-Shape switch cost 2, respectively).

Table S4 reports the reliabilities of all target EF measures[[2]](#footnote-2) and Table S5 shows the reliabilities of the RT difference scores from inhibition and switching tasks. As reported in the main manuscript, all target measures showed reliabilities close to or above Cronbach’s *α* = .8. Also, for each EF construct, highest reliabilities were found for target measures from the longest task of that construct (Flanker, Number-Letter, N-Back). Moreover, for inhibition and switching tasks, the target RT measures always had higher reliabilities than the difference scores from the same task.

For the difference scores, again, measures from longer tests (Flanker, Number-Letter) showed higher reliabilities. Moreover, difference scores that used the least EF-involving condition as baseline (neutral RT from inhibition and pure RT from switching tasks[[3]](#footnote-3)) also had higher reliabilities compared to difference scores that involved a relatively more EF-involving baseline (congruent RT from inhibition and repeat RT from switching tasks). Consequently, difference scores that satisfied both of these conditions (Flanker interference effect 1 and Number-Letter switch cost 1) showed the highest, close to the acceptable Cronbach’s *α* = .8 level, reliabilities.

Table S6 shows the correlations between all EF measures (including difference scores). As reported in the main manuscript, correlations between target EF measures from separate tasks of the same process were in the small-to-moderate range and all statistically significant. Furthermore, generally, for all RT measures—whether target RT or difference scores—the reliability results were reflected in the correlations. First, for inhibition and switching, correlations between target RT measures (most reliable scores from inhibition and switching tasks) from separate tasks of the same process were always higher than correlations involving one or two difference scores. Second, of the two difference scores from each inhibition and switching task, the most reliable difference score always had the largest correlations with the other three measures (RT target and two difference scores) from the second task of the same process. For example, the Number-Letter switch cost 1 measure had larger (and significant) correlations with the Colour-Shape RT target, Colour-Shape switch cost 1 and 2 measures compared to the correlations between the Number-Letter switch cost 2 score and the three Colour-Shape measures.

Moreover, contrary to expectation, the difference scores from the Colour-Shape and Number-Letter switching tasks significantly correlated with each other, with correlations in the small-to-moderate range. The highest correlation (*r* = .41, *p*(two-tailed) <. 05) was between the Colour-Shape switch cost 1 and Number-Letter switch cost 1 measures (each being the most reliable difference score from the corresponding switching task). Finally, for switching tasks, the largest correlation involving a difference score was between the Number-Letter switch cost 1 (most reliable switching difference score) and the Colour-Shape RT target measure (*r* = .5, *p*(two-tailed) < .05).

In contrast, there were no significant correlations between the difference scores from the inhibition tasks. However, again, the largest correlation (and in the expected direction) was between the two most reliable difference scores from each task (Flanker interference effect 1 and Stroop interference effect 1; *r* =. 07, *p*(two-tailed) > .05). We speculate that correlations between the inhibition difference scores were non-significant (and always very small) because, in the inhibition tasks, all trials, including neutral trials, were intermixed in the same block. In line with this, as already reported, the two switching difference scores, calculated by subtracting RTs from the least EF-demanding trials *in a separate block* (Colour-Shape switch cost 1 and Number-Letter switch cost 1), showed the highest correlation among all eight difference scores. These results are also consistent with theoretical arguments that all trial types, even those that are often considered easier, *baseline* trials, involve EF when intermixed in the same EF-demanding block (see e.g., Section 1.1. in main manuscript and in Antoniou, 2023); and that calculating difference scores by subtracting these easier trials removes EF variance. This, in turn, may be contributing to the low convergent validity of difference scores calculated from same-block trial types. This argument is further reinforced when considering the correlations between mean RT for the critical, most EF-demanding trials and the other two trial types in each inhibition and switching task: The two correlations were very high and almost indistinguishable in the inhibition tasks where all trials were intermixed in the same block (Stroop incongruent and congruent: *r* = .90, Stroop incongruent and neutral: *r* = .89, Flanker incongruent and congruent: *r* = .93, Flanker incongruent and neutral: *r* = .91), suggesting that these trials types tap, to a large extent, into very similar processes. However, correlations between switch trials and separate-block pure trials in the two switching tasks were substantially lower (Colour-Shape switch and pure: *r* = .51, Number-Letter switch and pure: r = .74) than correlations between the intermixed switch and repeat trials (Colour-Shape switch and repeat: *r* = .74, Number-Letter switch and repeat: *r*= .92).

Finally, for inhibition tasks, correlations were larger and significant when considering a difference score from one inhibition task and a target RT measure from the other inhibition task. The highest correlation was between the Flanker interference effect 1 (most reliable inhibition difference score) and the Stroop RT target measure (*r* = .23, *p*(two-tailed) < .05).

In the main manuscript, we reported analyses that, for inhibition and switching, used the target RT measures. However, because some inhibition and switching difference scores showed adequate reliabilities and some degree of convergent validity, in the following sections of the SMO (Sections S6 and S8), we report additional analyses (Confirmatory Factor Analyses and group comparisons) with these difference scores.

Overall, the findings in this section suggest that highest reliability is evident for difference scores that (1) stem from longer tests (e.g., ≥ 96 trials for each relevant condition, based on our Flanker results) and (2) are calculated by subtracting a least EF-involving baseline condition. In turn, high reliability determines convergent validity: A score with higher reliability exhibits higher correlations with other measures from separate tasks of the same process. Also, convergent validity seems to be further, independently affected by another factor; specifically, it is higher for difference scores calculated by subtracting a least EF-involving baseline condition *from a separate block*.

The link between reliability and convergent validity is a manifestation of the broader positive relation between reliability and statistical power, which has been also reported for between-subject, group comparisons (e.g., Hajcak et al., 2017; discussion in Antoniou, 2023). Assuming that an effect exists, low reliability leads to an underestimation of the true effect. In contrast, higher reliability leads to an observed effect (e.g., correlation or group difference) that is closer to the true effect and thus, ceteris paribus, to greater likelihood of a statistical test to detect this effect (assuming it exists). Overall, these results suggest that reliability and the way difference scores are calculated are important factors to consider in the EF and multilingualism literatures (Draheim et al., 2019). First, they determine convergent validity, which helps evaluate whether scores measure the construct of interest. Second, reliability affects the size of the observed relative to the true effect; and, thus, statistical power.

|  |
| --- |
| Table S4. *Reliabilities (Cronbach’s α) for the target measures.* |
| Task | Target measure | *α* |
| Mill Hill Definitions1 | Total correct  | .91 |
| Mill Hill Multiple Choice1 | Total correct | .93 |
| WASI1 | Total correct | .77 |
| N-Back1 | Proportion correct | .88 |
| Rotation Span1 | Partial score | .74 |
| Corsi Forward1 | Partial score | .82 |
| Corsi Backward1 | Partial score | .80 |
| Number-Letter2 | Switch RT  | .95 |
| Colour-Shape2 | Switch RT  | .83 |
| Flanker2 | Incongruent RT  | .95 |
| Stroop2 | Incongruent RT  | .91 |
| *Note*. RT: mean reaction times for correct trials, WASI: Matrix Reasoning sub-test from the Wechsler Abbreviated Scale of Intelligence.1Reliability calculated based on all critical items.2Reliability calculated based on critical block performance. |
|  |

|  |
| --- |
| Table S5.*Reliabilities (Cronbach’s α) of RT difference scores from the inhibition and switching tasks.* |
| Task | Target measure | *α* |
| Number-Letter1 | Switch RT-Repeat RT | .72 |
| Number-Letter2 | Switch RT-Pure RT | .77 |
| Colour-Shape1 | Switch RT-Repeat RT | .57 |
| Colour-Shape3 | Switch RT-Pure RT | .71 |
| Flanker1 | Incongruent RT- Congruent RT | .65 |
| Flanker1 | Incongruent RT- Neutral RT | .78 |
| Stroop1 | Incongruent RT- Congruent RT | .60 |
| Stroop1  | Incongruent RT- Neutral RT | .66 |
| *Note*. RT: mean reaction times for correct trials.1Reliability calculated based on two or more scores, one for each critical block.2Reliability calculated based on two scores: RT for 1st and 2nd mixed block minus RT for 1st pure block; and RT for 3rd and 4th mixed block minus RT for 2nd pure block.3Reliability calculated based on two scores: RT for 1st mixed block minus RT for 1st pure block; and RT for 2nd mixed block minus RT for 2nd pure block. |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table S6.*Bivariate (Pearson’s r) correlations between all executive function RT measures.*  |  |  |  |  |  |  |  |  |
|  | Stroop RT | Flanker RT | NL RT | CS RT | N-Back | Corsi For. | Corsi Back. | RS | Flanker IE1 | Flanker IE2 | Stroop IE1 | Stroop IE2 | NL SC1 | NL SC2 | CS SC1 |
| Flanker RT | .37\*\* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NL RT | .56\*\* | .40\*\* |  |  |  |  |  |  |  |  |  |  |  |  |  |
| CS RT | .41\*\* | .22\*\* | .56\*\* |  |  |  |  |  |  |  |  |  |  |  |  |
| N-Back | -.23\*\* | -.20\* | -.29\*\* | -.19\* |  |  |  |  |  |  |  |  |  |  |  |
| Corsi For. | -.30\*\* | -.25\*\* | -.23\*\* | -.18\* | .24\*\* |  |  |  |  |  |  |  |  |  |  |
| Corsi Back. | -.31\*\* | -.25\*\* | -.28\*\* | -.21\*\* | .32\*\* | .38\*\* |  |  |  |  |  |  |  |  |  |
| RS | -.25\*\* | -.16\* | -.32\*\* | -.32\*\* | .26\*\* | .34\*\* | .46\*\* |  |  |  |  |  |  |  |  |
| Flanker IE1 | .23\*\* | .74\*\* | .22\*\* | .09 | -.23\*\* | -.22\*\* | -.23\*\* | -.08 |  |  |  |  |  |  |  |
| Flanker IE2 | .17\* | .57\*\* | .19\* | .10 | -.21\*\* | -.14 | -.17\* | -.01 | .85\*\* |  |  |  |  |  |  |
| Stroop IE1 | .78\*\* | .15 | .37\*\* | .26\*\* | -.11 | -.19\* | -.18\* | -.19\* | .07 | -.01 |  |  |  |  |  |
| Stroop IE2 | .74\*\* | .09 | .28\*\* | .22\*\* | -.06 | -.16\* | -.12 | -.09 | .03 | -.02 | .84\*\* |  |  |  |  |
| NL SC1 | .43\*\* | .21\*\* | .92\*\* | .51\*\* | -.25\*\* | -.16\* | -.24\*\* | -.35\*\* | .13 | .13 | .32\*\* | .24\* |  |  |  |
| NL SC2 | .18\* | .08 | .56\*\* | .25\*\* | -.04 | -.00 | .01 | -.14 | .04 | .02 | .16\* | .13 | .67\*\* |  |  |
| CS SC1 | .14 | -.03 | .37\*\* | .75\*\* | -.11 | .05 | -.05 | -.19\* | -.11 | -.08 | .11 | .13 | .41\*\* | .26\*\* |  |
| CS SC2 | .02 | -.07 | .23\*\* | .63\*\* | -.04 | .07 | -.02 | -.20\*\* | -.12 | -.06 | -.02 | .04 | .28\*\* | .23\*\* | .74\*\* |
| \*\*. Correlation is significant at the 0.01 level (two-tailed). \*. Correlation is significant at the 0.05 level (two-tailed).*Note.* RT: mean reaction times for correct trials, Flanker RT: RT for incongruent trials in the Flanker task, Stroop RT: RT for incongruent trials in the Stroop task, CS RT: RT for switch trials in the Colour-Shape task, NL RT: RT for switch trials in the Number-Letter task, Corsi For.: partial score in the forward Corsi blocks task, Corsi Back.: partial score in the backward Corsi blocks task, RS: partial score in the Rotation span task, IE1: interference effect 1, IE2: interference effect 2, SC1: switch cost 1, SC2: switch cost 2. |

1. **Additional confirmatory factor analyses with difference scores from the switching and inhibition tasks**

Some inhibition and switching RT difference scores in our study showed adequate reliabilities and some degree of convergent validity (see Section S5, Table S5 and Table S6 in the SMO). For this reason, we conducted additional Confirmatory Factor Analyses (CFA) with these difference scores; specifically, a second CFA included the Colour-Shape switch cost 1 and Number-Letter switch cost 1 difference scores as indicators of the switching component; and the Flanker interference effect 1 difference score and Stroop mean RT target measure as indicators of inhibition (see Kane et al., 2016; McVay & Kane, 2012, for a similar approach). However, this model provided an improper and inadmissible solution, regardless of fit indices. This was evident in a standardised factor loading above 1 and a negative standardised error term for the Number-Letter switch cost 1 measure (Heywood case; e.g., Brown, 2016: 162; Karr et al., 2018: 1153). Thus, a third CFA was performed with the same indicators as above for all factors but with the Colour-Shape mean RT target measure (instead of the Colour-Shape switch cost 1 difference score) as a switching indicator. This final CFA converged properly, with acceptable fit indices: *χ*2(17, *n* = 165) = 24.1, *p* > .05, *χ*2/df = 1.42, CFI = 0.97, NNFI = 0.95, RMSEA value = 0.05, 90% confidence interval 0.00-0.09. Moreover, there were significant moderate-to-large correlations between the three factors: .78 between switching and inhibition, .63 between inhibition and working memory; and .52 between switching and working memory.

1. **Tables** **with summary results of the initial linear mixed-effects comparison model for the analyses in the main manuscript indicating a significant Group effect**

|  |
| --- |
| Table S7. *Summary of the initial linear mixed-effects comparison model for the analysis with the whole participant sample indicating a significant Group effect on executive function (with only Standard Modern Greek Vocabulary included as a covariate).* |
| Effect | Coefficient | *SE* | *df* | t-value | Pr(>|*z*|) | Random Slope(s) |
| Intercept | -0.25 | 0.10 | 229.8 | -2.48 | = .014 | no |
| Group 1 | 0.31 | 0.12  | 175 | 2.61  | = .01 | no |
| Group 2 | 0.39  | 0.14  | 175 | 2.75  | = .007 | no |
| Component 2 | 0.01  | 0.06  | 356 | 0.16  | > .05 | no |
| Component 3  | 0.04  | 0.06  | 356 | 0.07  | > .05 | no |
| Vocabulary | 0.15  | 0.06  | 175 | 2.60 | = .01 | no |
| *Note 1.* Group 1: Bidialectals. Group 2: Multilinguals. Component 2: Switching composite score. Component 3: Working Memory composite score, Vocabulary: Standard Modern Greek vocabulary composite score.*Note 2*. Degrees of freedom and *p* values were calculated using Kenward-Roger's method from the *pbkrtest* package (Halekoh & Højsgaard, 2014) in R (R Core Team, 2020). |

|  |
| --- |
| Table S8. *Summary of the initial linear mixed-effects comparison model for the analysis with the matched participant samples indicating a significant Group effect on executive function (with Standard Modern Greek Vocabulary and all other background variables included as covariates).* |
| Effect | Coefficient | *SE* | *df* | t-value | Pr(>|*z*|) | Random Slope(s) |
| Intercept | 0.06 | 0.62 | 89.05 | 0.09 | > .05 | no |
| Group 1 | -0.31 | 0.12  | 88 | -2.52  | = .014 | no |
| Component 2 | 0.04  | 0.08  | 192 | 0.44  | > .05 | no |
| Component 3  | 0.04  | 0.08  | 192 | 0.45  | > .05 | no |
| Vocabulary | 0.41  | 0.13  | 88 | 3.05 | = .003 | no |
| Gender 1 | -0.27 | 0.16 | 88 | -1.70 | > .05 | no |
| SES | 0.03 | 0.1 | 88 | 0.25 | > .05 | no |
| Age | -0.14 | 0.07 | 88 | -1.93 | > .05 | no |
| Education 1 | 0.06 | 0.61 | 88 | 0.10 | > .05 | no |
| Education 2 | 0.1 | 0.65 | 88 | 0.16 | > .05 | no |
| Education 3 | 0.28 | 0.67 | 88 | 0.42 | > .05 | no |
| *Note 1.* Group 1: “Monolinguals”. Component 2: Switching composite score. Component 3: Working Memory composite score, Vocabulary: Standard Modern Greek vocabulary composite score, Gender 1= Female, SES: Socioeconomic Status, Education 1: participants who started studying for or have a first degree, Education 2: participants who started studying for or have a master’s degree, Education 3: participants who started studying for or have a doctoral degree.*Note 2*. Degrees of freedom and *p* values were calculated using Kenward-Roger's method from the *pbkrtest* package (Halekoh & Højsgaard, 2014) in R (R Core Team, 2020). |

1. **Additional group analyses with the inhibition and switching composite scores that were based on difference scores**

As reported in Section S6 (SMO), the third CFA that, for switching, included the Number-Letter switch cost 1 difference score and Colour-Shape mean RT target measures; and, for inhibition, included the Flanker interference effect 1 difference score and Stroop mean RT target measures converged properly and with acceptable fit indices. For this reason, further group analyses were conducted on the inhibition and switching composite scores that included these difference scores, as dependent measures. The working memory composite score was the same as in the main analysis (Section 3.1., main manuscript).

Specifically, the inhibition composite score was computed by averaging the *z*-transformed and reverse-scored Flanker interference effect 1 difference score and Stroop target RT measures; and the switching composite was calculated based on the *z*-transformed and reverse-scored Number-Letter switch cost 1 and Colour-Shape target RT measures. Importantly, these composite scores for inhibition and switching reflect the variance common to (1) mean RT in the critical EF condition (Stroop incongruent and Colour-Shape switch), which, as already discussed, subsumes basic processing speed and the relevant EF process; and (2) the difference scores, which isolate the target EF process. Thus, these composite scores do not reflect simple processing speed (McVay & Kane, 2012: 308). In turn, a potential group effect on overall EF (as discussed in the main manuscript) or in any of these composite scores or their combination can be attributed to EF but not to simple processing speed.

Table S9 presents descriptive statistics for the composite EF scores in this analysis. The first analysis (with Group, Component, and the Group by Component interaction) showed no significant results (Group by Component interaction: *χ2*(4, *n* = 181) = 2.77, *p* > .05, Group effect: *χ2*(2, *n* = 181) = 3.23, *p* > .05). To examine whether the Group effect depends on SMG vocabulary proficiency, two further analyses were conducted. First, we included SMG Vocabulary in the comparison models. This analysis showed a non-significant Group by Component interaction (*χ2*(4, *n* = 179) = 2.78, *p* > .05) but a significant Group effect (*χ2*(2, *n* = 179) = 7.1, *p* = .029), in the direction of better multilingual and bidialectal EF performance (estimate = 0.32, *SE* = 0.13, *z* value = 2.39, *p* =.031 and estimate = 0.26, *SE* = 0.11, *z* value = 2.36, *p* = .034, respectively)[[4]](#footnote-4).

The second analysis was performed on the matched samples. Table S10 presents descriptive statistics for the composite EF scores in this analysis. Model comparisons (with SMG Vocabulary, Gender, SES, Age, and Education covaried) showed a non-significant Group by Component interaction (*χ2*(2, *n* = 97) = 0.21, *p* > .05) but a significant Group effect (*χ2*(1, *n* = 97) = 6.69, *p* = .01) in the direction of better “multilingual” EF performance. The “multilingual” effect in this analysis remained significant even when not covarying SMG Vocabulary in the comparison models (*χ2*(1, *n* = 97) = 5.46, *p* = .019).

Thus, overall, the group results, indicating better multilingual/bidialectal performance, are largely similar, in terms of statistical significance, when using composite scores that, for inhibition and switching, were based on difference scores instead of the target RT measures. In this respect, the results in this section further confirm that the multilingual/bidialectal cognitive effect lies in EF rather than in simple processing speed.

|  |
| --- |
| Table S9. *Descriptive statistics for the composite executive function scores (based on difference scores, where appropriate) for the whole sample.*  |
|  | “Monolinguals” | Bidialectals | Multilinguals | Mean difference: Unstandardised [95% CIs], Standardised [95% CIs] |
|  | Mean [95% CIs] | Mean [95% CIs] | Mean [95% CIs] |  |
| Inhibition1 | -.14 [-.37, .09] | 0.03 [-.12, .19] | .08 [-.16, .33] | Mult-Monol.: .23 [-.11, .57], .26 [-.13, .66]Mult.-Bid.: .05 [-.23, .33], .07 [-.29, .43]Bid.-Monol.: .17 [-.09, .44], .23 [-.12, .59] |
| Switching1 | -.12 [-.32, .08] | .11 [-.08, .29] | -.04 [-.33, .24] | Mult-Monol.: .07 [-.28, .42], .08 [-.31, .48]Mult.-Bid.: -.15 [-.47, .18], -.16 [-.52, .19]Bid.-Monol.: .22 [-.06, .50], .28 [-.07, .64] |
| Working Memory | -.13 [-.33, .07] | .01 [-.15, .17] | .10 [-.08, .29] | Mult-Monol.: .24 [-.04, .51], .34 [-.05, .74]Mult.-Bid.: .09 [-.16, .34], .13 [-.22, .49]Bid.-Monol.: .14 [-.11, .40], .20 [-.15, .56] |
| Executive Function1 | -.14 [-.30, .03] | .04 [-.09, .17] | .06 [-.12, .24] | Mult-Monol.: .19 [-.05, .44], .31 [-.08, .71]Mult.-Bid.: .02 [-.20, .24], .03 [-.32, .39]Bid.-Monol.: .17 [-.03, .38], .30 [-.05, .66] |
| *Note.* CIs: Confidence Intervals, Mult.: Multilinguals, Bid.: Bidialectals, Monol.: “Monolinguals”. 1Composite scores were calculated from the reverse-scored (multiplied with -1) and *z*-transformed individual measures (from the inhibition and switching tasks) so that a higher value indicates better performance. The executive function composite score was calculated by averaging all individual *z*-transformed (and reverse-scored, as appropriate) executive function measures. |

|  |
| --- |
| Table S10. *Descriptive statistics for the composite executive function scores (based on difference scores, where appropriate) for the matched samples.*  |
|  | “Monolinguals”(*n* = 52) | “Multilinguals”(*n* = 52) | Mean difference: Unstandardised [95% CIs], Standardised [95% CIs] |
|  | Mean [95% CIs] | Mean [95% CIs] |  |
| Inhibition1 | -.14 [-.37, .09] | .15 [-.06, .36] | Mult.-Monol.: .29 [.02, .6], .36 [.03, .76] |
| Switching1 | -.12 [-.32, .08] | .09 [-.14, .32] | Mult.-Monol.: .21 [-.1, .51], .26 [-.12, .66] |
| Working Memory | -.13 [-.33, .07] | .11 [-.09, .30] | Mult.-Monol.: .24 [-.04, .52], .33 [-.05, .73] |
| Executive Function1 | -.14 [-.30, .03] | .11 [-.06, .27] | Mult.-Monol.: .24 [.01, .47], .40 [.02, .81] |
| *Note.* CIs: Confidence Intervals, Mult.: “Multilinguals” (multilinguals and bidialectals), Monol.: “Monolinguals”. 1Composite scores were calculated from the reverse-scored (multiplied with -1) and *z*-transformed individual measures (from the inhibition and switching tasks) so that a higher value indicates better performance. The executive function composite score was calculated by averaging all individual *z*-transformed (and reverse-scored, as appropriate) executive function measures. |

1. **Regression analysis and interpretation of its results when the covariate and the Group factor are related**

From an interpretation perspective, the linear mixed-effect model (regression) analysis with Group as a between-subjects factor and SMG Vocabulary as a covariate examines the Group effect when SMG Vocabulary is held constant at a fixed value for all participants. In other words, the analysis examines the Group effect when, because of covarying SMG Vocabulary, neither Group nor EF are related to SMG Vocabulary (e.g., Cohen et al., 2002; see also Miller & Chapman, 2001, for a discussion in the context of Analysis of Covariance). Thus, at a conceptual (interpretation) level, this analysis assumes that (1) the part of SMG vocabulary that correlates with EF is not an inherent, essential, central part of EF and (2) the part of SMG vocabulary that correlates with Group, that is, the group differences in SMG vocabulary are not an inherent, essential, central part of the Group factor. In other words, for the Group factor, the analysis assumes that there exist real-life multilinguals, bidialectals, and “monolinguals” who have approximately[[5]](#footnote-5) similar SMG vocabulary proficiency (Zinbarg et al., 2010: 308, for a discussion in the context of Analysis of Covariance). In turn, results can be generalised only to these sub-populations; that is, to multilinguals, bidialectals, and “monolinguals” who have approximately equal SMG vocabulary proficiency.

From a statistical perspective, there is no statistical requirement in regression that the predictors should be completely independent, even though very high correlations (multicollinearity) between the predictors (roughly, *r* > .8) can bias the regression model (e.g., Field, 2013: 324-326, 486). Looking at Table S11, there are no variables in our case which correlate strongly (highest correlation is *r* = .49; correlation between the Group factor and SMG Vocabulary is *r* = -.44, *p*(two-tailed) < .001). We also note that due to the very small correlation between SMG Vocabulary and EF processes (Table S11) any bias (towards an artifactual, spurious effect of Group) in the statistical analysis would be very small (see Zinbarg et al., 2010: 314-315).

To directly examine whether the results from a regression analysis with Group and SMG Vocabulary as predictors (correlated negatively and moderately) might be spurious, we calculated the Type I error rate for such an analysis using the *Ivy* (*Incremental Validity* Error Rate Calculator) shiny app (<https://jakewestfall.shinyapps.io/ivy_app/>; Westfall and Yarkoni, 2016). We performed two analyses, one for the null hypothesis that there is no Group effect (i.e., simple correlation between Group and overall EF performance is *0*) and one for the null hypothesis that there is no Group effect after covarying SMG Vocabulary (i.e., partial correlation, after covarying SMG Vocabulary, between Group and overall EF performance is *0*). We specified a simple correlation of *r* = .09 between SMG Vocabulary and overall EF performance (Table S11), a simple correlation of *r* = -.443 between Group and SMG Vocabulary, reliabilities of .9 for the Group and .9 for the SMG Vocabulary predictors (Table S4) and a sample size of 179 participants (one bidialectal and one “monolingual” had missing data for SMG Vocabulary). Both analyses showed that the probability of rejecting the regression coefficient for the Group predictor, when there is no Group effect, was within nominal level (*p* < .05). We note that changing the reliability for the Group factor (for which there is no reliability estimate in our sample) does not alter these results. Finally, the same results hold when considering the overall EF score calculated by including the Flanker interference effect 1 and Number-Letter switch cost 1 difference scores—based on the second CFA that properly and acceptably converged—instead of the target RT measures for the critical incongruent and switch trials in the Flanker and Number-Letter tasks, respectively. Correlation between this overall EF score and SMG Vocabulary was *r* = .07.

|  |
| --- |
| Table S11.*Bivariate correlations (Pearson’s r) between the composite executive function scores and background variables.*  |
|  | EF1 | WM | Inhibition1  | Switching1 | Vocabulary | Age | Gender | SES | Education |
| WM | .85\*\* |  |  |  |  |  |  |  |  |
| Inhibition1 | .76\*\* | .43\*\* |  |  |  |  |  |  |  |
| Switching1 | .76\*\* | .41\*\* | .53\*\* |  |  |  |  |  |  |
| Vocabulary | .09 | .02 | .11 | .13 |  |  |  |  |  |
| Age | -.14 | -.13 | -.09 | -.11 | .13 |  |  |  |  |
| Gender | -.17\* | -.25\*\* | -.05 | -.04 | .08 | -.21\*\* |  |  |  |
| SES | .14 | .15\* | .07 | .11 | -.07 | -.20\*\* | .00 |  |  |
| Education | -.06 | -.05 | .03 | -.13 | .07 | .49\*\* | -.04 | -.11 |  |
| WASI | .46\*\* | .38\*\* | .36\*\* | .34\*\* | .07 | .05 | -.13 | -.02 | .11 |
| \*\*. Correlation is significant at the 0.01 level (two-tailed). \*. Correlation is significant at the 0.05 level (two-tailed).*Note.* EF: overall executive function composite score, WM: Working memory composite score, Inhibition: Inhibition composite score, Switching: Switching composite score, Vocabulary: Standard Modern Greek vocabulary composite score, Gender: *0* = male, *1* = female, SES: Socioeconomic status composite score, Education: *0* = no degree, *1* = participants who started studying for or have a first degree, *2* = participants who started studying for or have a master’s degree, *3* = participants who started studying for or have a doctoral degree, WASI: score in the Matrix Reasoning sub-test from the Wechsler Abbreviated Scale of Intelligence.1Composite scores were calculated from the reverse-scored (multiplied with -1) individual target measures from the inhibition and switching tasks so that a higher value indicates better performance. |  |

1. **Correlations between the composite executive function scores and various multilingual and bidialectal experiences measured continuously**

|  |
| --- |
| Table S12.*Bivariate correlations between the composite executive function scores and continuous measures of English- and other-language experience across the whole sample.*  |
|  | EF1 | EF11 | WM  | Inhibition1 | Inhibition11 | Switching1 | Switching11 | EnProf | EnUse | EnAoOE | OProf  | OUse |
| EF11 | .98\*\* |  |  |  |  |  |  |  |  |  |  |  |
| WM | .85\*\* | .87\*\* |  |  |  |  |  |  |  |  |  |  |
| Inhibition1 | .76\*\* | .71\*\* | .43\*\* |  |  |  |  |  |  |  |  |  |
| Inhibition11 | .70\*\* | .72\*\* | .43\*\* | .90\*\* |  |  |  |  |  |  |  |  |
| Switching1 | .76\*\* | .73\*\* | .41\*\* | .53\*\* | .44\*\* |  |  |  |  |  |  |  |
| Switching11 | .70\*\* | .71\*\* | .39\*\* | .43\*\* | .37\*\* | .98\*\* |  |  |  |  |  |  |
| EnProf | .14 | .12 | .05 | .22\*\* | .22\*\* | .15 | .12 |  |  |  |  |  |
| EnUse | .00 | .00 | -.02 | .05 | .07 | -.01 | -.02 | .45\*\* |  |  |  |  |
| EnAoOE | -.16 | -.14 | -.15 | -.07 | -.03 | -.14 | -.10 | -.21\* | -.14 |  |  |  |
| OProf | -.02 | -.04 | -.08 | .01 | -.01 | .05 | .06 | .26\* | .16 | -.17 |  |  |
| OUse | -.01 | .00 | -.04 | -.02 | -.02 | .05 | .09 | .12 | .31\*\* | -.34\*\* | .70\*\* |  |
| OAoOE | .03 | .03 | -.07 | -.03 | -.03 | .18 | .18 | .00 | -.09 | -.13 | -.47\*\* | -.31\* |

\*\*. Correlation is significant at the 0.01 level (two-tailed). \*. Correlation is significant at the 0.05 level (two-tailed).

*Note 1.* EF: overall executive function composite score (based on the target RT measures for all inhibition and switching tasks), EF1: overall executive function composite score (based on the target RT measures for the Colour-Shape switching and Stroop inhibition tasks and on the difference scores Number-Letter switch cost 1 and Flanker interference effect 1), WM: Working memory composite score, Inhibition: Inhibition composite score (based on the target RT measures from the two inhibition tasks), Inhibition1: Inhibition composite score (based on the target RT measure from the Stroop inhibition task and on the difference score Flanker interference effect 1), Switching: Switching composite score (based on the target RT measures from the two switching tasks), Switching1: Switching composite score (based on the target RT measure from the Colour-Shape switching task and on the difference score Number-Letter switch cost 1), EnProf: self-reported English proficiency, EnUse: self-reported English use at both the production and exposure level, EnAoOE: age of onset of exposure to/use of English, OProf: self-reported other-language proficiency, OUse: self-reported other-language use at both the production and exposure level, ΟAoOE: age of onset of exposure to/use of other language.

*Note 2.* Number of participants (*n*) varies depending on the pair of variables examined (e.g., regarding the executive function measures, *n* = 170 for the correlations with self-reported English proficiency, *n* = 89 for the correlations with other-language proficiency, and *n* = 95 for the correlations with overall other-language use).

*Note 3.* All correlations are Pearson’s *r*, apart from the correlations with the self-reported English and other-language proficiency scores (EnProf and OProf, respectively) which are Spearman’s *ρ* correlations.

1Composite scores were calculated from the reverse-scored (multiplied with -1) individual measures from the inhibition and switching tasks so that a higher value indicates better performance.

|  |
| --- |
| Table S13.*Bivariate correlations (Pearson’s r) between the composite executive function scores and continuous measures of Standard Modern Greek experience across the sample of multilinguals and bidialectals.* |
|  | EF1 | EF11 | WM  | Inhibition1 | Inhibition11 | Switching1 | Switching11 | Vocabulary | SMG Use |
| Vocabulary | .13 | .10 | .05 | .14 | .09 | .15 | .13 |  |  |
| SMG Use | -.05 | -.07 | -.08 | -.02 | -.04 | -.01 | -.03 | .14 |  |
| SMG AoOE | -.15 | -.16 | -.11 | -.10 | -.13 | -.16 | -.14 | -.34\*\* | -.26\*\* |

\*\*. Correlation is significant at the 0.01 level (two-tailed). \*. Correlation is significant at the 0.05 level (two-tailed).

*Note 1.* EF: overall executive function composite score (based on the target RT measures for all inhibition and switching tasks), EF1: overall executive function composite score (based on the target RT measures for the Colour-Shape switching and Stroop inhibition tasks and on the difference scores Number-Letter switch cost 1 and Flanker interference effect 1), WM: Working memory composite score, Inhibition: Inhibition composite score (based on the target RT measures from the two inhibition tasks), Inhibition1: Inhibition composite score (based on the target RT measure from the Stroop inhibition task and on the difference score Flanker interference effect 1), Switching: Switching composite score (based on the target RT measures from the two switching tasks), Switching1: Switching composite score (based on the target RT measure from the Colour-Shape switching task and on the difference score Number-Letter switch cost 1), Vocabulary: Standard Modern Greek vocabulary composite score, SMG Use: self-reported Standard Modern Greek use at both the production and exposure level, SMG AoOE: age of onset of exposure to/use of Standard Modern Greek.

*Note 2.* Number of participants (*n*) varies depending on the pair of variables examined (e.g., regarding the executive function measures, *n* = 128 for the correlations with Standard Modern Greek vocabulary, *n* = 124 for the correlations with overall Standard Modern Greek use across production and exposure, and *n* = 109 for the correlations with age of onset of exposure to/use of Standard Modern Greek).

1Composite scores were calculated from the reverse-scored (multiplied with -1) individual measures from the inhibition and switching tasks so that a higher value indicates better performance.

**References**

Antoniou, K. (2023). The ups and downs of bilingualism: A review of the literature on executive control using event-related potentials. *Psychonomic Bulletin & Review, 30*, 1187-1226. <https://doi.org/10.3758/s13423-023-02245-x>

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research*. New York: The Guilford Press.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences*. Hillsdale, NJ: Erlbaum.

Cohen, J., Cohen, P., West, S.G., & Aiken, L.S. (2002). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (3rd ed.). Routledge.

Draheim, C., Mashburn, C. A., Martin, J. D., and Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin, 145*, 508. <https://doi.org/10.1037/bul0000192>

Eurostat (2022, June). *Foreign language learning statistics*. <https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Foreign_language_learning_statistics&oldid=576746#Primary_education>.

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & Cognition, 43*, 226-236. <https://doi.org/10.3758/s13421-014-0461-7>

Field, A. (2013). *Discovering statistics using SPSS* (4th ed.). London: Sage.

Hajcak, G., Meyer, A., & Kotov, R. (2017). Psychometrics and the neuroscience of individual differences: Internal consistency limits between-subjects effects. *Journal of Abnormal Psychology, 126*, 823. <https://doi.org/10.1037/abn0000274>

Halekoh, U., & Højsgaard, S. (2014). A Kenward-Roger Approximation and Parametric Bootstrap Methods for Tests in Linear Mixed Models – The R Package pbkrtest. *Journal of Statistical Software, 59*, 1–32. <https://doi.org/10.18637/jss.v059.i09>

Kane, M. J., Meier, M. E., Smeekens, B. A., Gross, G. M., Chun, C. A., Silvia, P. J., & Kwapil, T. R. (2016). Individual differences in the executive control of attention, memory, and thought, and their associations with schizotypy. *Journal of Experimental Psychology: General, 145*, 1017-1048. <https://doi.org/10.1037/xge0000184>

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin, 144*, 1147. <https://doi.org/10.1037/bul0000160>

McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of experimental psychology: general, 141*, 302. <https://doi.org/10.1037/a0025250>

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology, 110*, 40. <https://doi.org/10.1037/0021-843X.110.1.40>

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. *Psychological Bulletin, 94*, 18. <https://doi.org/10.1037/0033-2909.94.1.18>

Zinbarg, R. E., Suzuki, S., Uliaszek, A. A., & Lewis, A. R. (2010). Biased parameter estimates and inflated Type I error rates in analysis of covariance (and analysis of partial variance) arising from unreliability: alternatives and remedial strategies. *Journal of Abnormal Psychology, 119*, 307. <https://doi.org/10.1037/a0017552>

Westfall, J., & Yarkoni, T. (2016). Statistically controlling for confounding constructs is harder than you think. *PloS one, 11*, e0152719. <https://doi.org/10.1371/journal.pone.0152719>

**Appendix S2. Translated version (from Greek to English) of the Language Background and Socioeconomic Status Questionnaire**

**Language Background and Socioeconomic Status Questionnaire**

1. **General information**

* 1. Birth Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1.2. Gender: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1.3. Place of birth: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1.4. If place of birth is not current country of residence, please indicate the date of arrival in current country of residence: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1.5. How many siblings do you have (please, circle)? 1 2 3 More than 3

1.6. Are you left-handed or right-handed (please, circle)? Left-handed Right-handed

1.7. Is there another left-handed person in your family (please, circle)? Yes No

1.8. If there is another left-handed person in your family, this is your (please, circle): Father Mother Sibling

1.9. What level of university studies are you currently at or have you attained (please, circle): Not applicable

First degree Master’s Doctoral

1.10. What do you study? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. **Language history**

2.1 At what age did you start being exposed to and/or using (please, indicate age in **MONTHS**!!):

|  |  |
| --- | --- |
|  | **Age in MONTHS** |
| Standard Greek[[6]](#footnote-6) |  |
| Cypriot Greek |  |
| English |  |
| Other (specify) |  |

2.2.Languages/dialects used in your family environment.

2.2.1. Languages/dialects that you use/used with your mother.

| **Language/dialect that your MOTHER uses/used** **towards YOU** | **Language/dialect that YOU use/used****towards your MOTHER** |
| --- | --- |
|  | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always** | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always**  |
| Standard Greek |  |  |  |  |  |  |  |  |  |  |
| Cypriot Greek |  |  |  |  |  |  |  |  |  |  |
| English |  |  |  |  |  |  |  |  |  |  |
| Other (specify) |  |  |  |  |  |  |  |  |  |  |

2.2.2. Languages/dialects that you use/used with your father.

| **Language/dialect that your FATHER uses/used****towards YOU** | **Language/dialect that YOU use/used****towards your FATHER** |
| --- | --- |
|  | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always** | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always**  |
| Standard Greek |  |  |  |  |  |  |  |  |  |  |
| Cypriot Greek |  |  |  |  |  |  |  |  |  |  |
| English |  |  |  |  |  |  |  |  |  |  |
| Other (specify) |  |  |  |  |  |  |  |  |  |  |

2.2.3. Languages/dialects that your parents use/used among themselves.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually** | **4****Always** |
| Standard Greek |  |  |  |  |  |
| Cypriot Greek |  |  |  |  |  |
| English |  |  |  |  |  |
| Other (specify) |  |  |  |  |  |

2.2.4. Languages/dialects that you use/used with your siblings (for each sibling, please, complete a separate table).

| **Language/dialect that your SIBLING 1[[7]](#footnote-7)****uses/used towards YOU** | **Language/dialect that YOU use/used towards your SIBLING 1** |
| --- | --- |
|  | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always** | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always**  |
| Standard Greek |  |  |  |  |  |  |  |  |  |  |
| Cypriot Greek |  |  |  |  |  |  |  |  |  |  |
| English  |  |  |  |  |  |  |  |  |  |  |
| Other (specify) |  |  |  |  |  |  |  |  |  |  |
| **Language/dialect that your SIBLING 2[[8]](#footnote-8)****uses/used towards YOU** | **Language/dialect that YOU use/used towards your SIBLING 2** |
|  | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always** | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always**  |
| Standard Greek  |  |  |  |  |  |  |  |  |  |  |
| Cypriot Greek |  |  |  |  |  |  |  |  |  |  |
| English |  |  |  |  |  |  |  |  |  |  |
| Other (specify) |  |  |  |  |  |  |  |  |  |  |

2.3. Languages of schooling[[9]](#footnote-9).

| **Language of instruction at pre-primary education** | **Language of instruction at primary education** |
| --- | --- |
|  | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always** | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always**  |
| Standard Greek |  |  |  |  |  |  |  |  |  |  |
| English  |  |  |  |  |  |  |  |  |  |  |
| Other (specify) |  |  |  |  |  |  |  |  |  |  |
| **Language of instruction at gymnasium (junior high school)** | **Language of instruction at lyceum (senior high school)** |
|  | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always** | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always**  |
| Standard Greek  |  |  |  |  |  |  |  |  |  |  |
| English |  |  |  |  |  |  |  |  |  |  |
| Other (specify) |  |  |  |  |  |  |  |  |  |  |

2.4. Languages after secondary education.

| **Language of instruction at undergraduate level (first degree)** | **Language of instruction at postgraduate (master’s) level**  |
| --- | --- |
|  | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always** | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always**  |
| Standard Greek |  |  |  |  |  |  |  |  |  |  |
| English  |  |  |  |  |  |  |  |  |  |  |
| Other (specify) |  |  |  |  |  |  |  |  |  |  |
| **Language of instruction at postgraduate (doctoral) level** |
|  | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always** |
| Standard Greek  |  |  |  |  |  |
| English |  |  |  |  |  |
| Other (specify) |  |  |  |  |  |

2.5. Languages/dialects that you use with your friends.

| **Language/dialect that your FRIENDS use****towards YOU** | **Language/dialect that YOU use****towards your FRIENDS** |
| --- | --- |
|  | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always** | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always**  |
| Standard Greek |  |  |  |  |  |  |  |  |  |  |
| Cypriot Greek |  |  |  |  |  |  |  |  |  |  |
| English |  |  |  |  |  |  |  |  |  |  |
| Other (specify) |  |  |  |  |  |  |  |  |  |  |

2.6. Languages/dialects used for activities in your free time (not for activities related to your studies).

| **Reading books, newspapers, magazines, and/or other written texts** | **Computer and/or internet use**  |
| --- | --- |
|  | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always** | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always**  |
| Standard Greek |  |  |  |  |  |  |  |  |  |  |
| English  |  |  |  |  |  |  |  |  |  |  |
| Other (specify) |  |  |  |  |  |  |  |  |  |  |
| **For watching TV, movies, and/or DVD** |
|  | **0****Never** | **1****Rarely** | **2****Sometimes** | **3****Usually**  | **4****Always** |
| Standard Greek  |  |  |  |  |  |
| Cypriot Greek |  |  |  |  |  |
| English |  |  |  |  |  |
| Other (specify) |  |  |  |  |  |

1. Language/dialect proficiency.
	1. How proficient would you characterise your knowledge of each of your spoken languages/dialects?

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **0****No proficiency** | **1****Limited proficiency** | **2****Moderate proficiency** | **3****Good proficiency** | **4****Very good proficiency** |
|  | No understanding or speaking ability | Some understanding ability and can form short, simple sentences in speaking | Adequate understanding ability and can express myself on various topics in speaking | Adequate understanding and speaking ability in most situations | Excellent understanding ability and comfortable use in all situations |
|  |  | e.g. can answer the phone | e.g. can go to the supermarket, explain what I want and buy things; can give directions about the location of a place  | e.g. can watch movies or television shows (without subtitles) |  |
| Standard Greek |  |  |  |  |  |
| Cypriot Greek |  |  |  |  |  |
| English |  |  |  |  |  |
| Other (specify) |  |  |  |  |  |

* 1. Which one of your languages/dialects do you know better and feel more comfortable using when talking to others (please, tick **ONLY ONE** option)?

|  |  |
| --- | --- |
|  | **√** |
| Standard Greek |  |
| Cypriot Greek |  |
| English |  |
| Other (specify) |  |
| Balanced proficiency and comfort in using two languages (specify) |  |

4. Information about your family.

4.1. General information.

4.1.1. Tick the descriptor which best summarises your family's situation/status.

|  |  |
| --- | --- |
| **Do you have a car/van/truck in your family?** | **√** |
| No |  |
| Yes, just one |  |
| Yes, more than one |  |

4.1.2. Tick the descriptor which best summarises your family's situation/status.

|  |  |
| --- | --- |
| **During the past 12 months, how many times have you travelled abroad for holidays either alone or with your family?** | **√** |
| None |  |
| Once |  |
| Twice |  |
| More than twice |  |

4.1.3. Tick the descriptor which best summarises you family's situation/status.

|  |  |
| --- | --- |
| **How many computers do you have in your family?** | **√** |
| None |  |
| One |  |
| Two |  |
| More than two |  |

* + 1. Tick the descriptor which best summarises your family’s situation/status.

|  |  |
| --- | --- |
| **Do/did you have your own bedroom at your family’s house?** | **√** |
| Yes |  |
| No |  |

**4.2 Information about the mother.**

4.2.1. In which country and region (if applicable) was she born? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

4.2.2. Is she currently working (please, circle)? Yes No

If yes, what is her job? Where does she work? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

4.2.3. Which languages/dialects does she speak?

|  |  |
| --- | --- |
|  | **√** |
| Standard Greek |  |
| Cypriot Greek |  |
| English |  |
| Other (specify) |  |

* + 1. Which languages/dialects does your mother know better and feel more comfortable using when talking to others (please, tick **ONLY ONE** option).

|  |  |
| --- | --- |
|  | **√** |
| Standard Greek |  |
| Cypriot Greek |  |
| English |  |
| Other (specify) |  |
| Balanced proficiency and comfort in using two languages (specify) |  |

* + 1. Education:

|  |  | **Number of years** | **Further information** |
| --- | --- | --- | --- |
| Primary school | Yes / No |  |  |
| High school (gymnasium/lyceum) | Yes / No |  |  |
| First degree | Yes / No |  |  |
| Master’s | Yes / No |  |  |
| Doctoral (PhD) | Yes / No |  |  |

**4.3. Information about the father.**

4.3.1. In which country and region (if applicable) was he born? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

4.3.2. Is he currently working (please, circle)? Yes No

If yes, what is his job? Where does he work? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

4.3.3. Which languages/dialects does he speak?

|  |  |
| --- | --- |
|  | **√** |
| Standard Greek |  |
| Cypriot Greek |  |
| English |  |
| Other (specify) |  |

* + 1. Which languages/dialects does your father know better and feel more comfortable using when talking to others (please, tick **ONLY ONE** option).

|  |  |
| --- | --- |
|  | **√** |
| Standard Greek |  |
| Cypriot Greek |  |
| English |  |
| Other (specify) |  |
| Balanced proficiency and comfort in using two languages (specify) |  |

* + 1. Education:

|  |  | **Number of years** | **Further information** |
| --- | --- | --- | --- |
| Primary school | Yes / No |  |  |
| High school (gymnasium/lyceum) | Yes / No |  |  |
| First degree | Yes / No |  |  |
| Master’s | Yes / No |  |  |
| Doctoral (PhD) | Yes / No |  |  |

5. Difficulties

In each cell, please tick in case you or another member of your family has/had each of the following difficulties.

|  | **You** | **Siblings (any)** | **Mother** | **Father** | **Father’s family** | **Mother’s family** |
| --- | --- | --- | --- | --- | --- | --- |
| Difficulties at school  |  |  |  |  |  |  |
|  Difficulties mainly with reading and spelling |  |   |  |  |  |  |
| Repeated one or more grades in school |  |  |  |  |  |  |
| Difficulties understanding others when they speak  |  |  |  |  |  |  |
| Difficulties expressing oneself orally (pronunciation, forming sentences, etc.) |  |  |  |  |  |  |
| Other difficulties (e.g., Attention Deficit Hyperactivity Disorder, Autism, social phobia). Please, note the kind of difficulty and whether it is you or another family member that has/had this difficulty. |  |

1. The Language Background and Socioeconomic Status Questionnaire did not ask for information about the participants’ partners or the participants’ language/dialect use with their partners (see Appendix S2, Supplementary Materials Online). Moreover, it did not inquire about language of work. We did not include such items because we aimed to mainly target and, in fact, the vast majority of our participants were university students. For the multilingual participant described in (3), information regarding their teaching and their spouse was orally communicated by the participant themselves (as was occasionally done by some participants) during the testing sessions. [↑](#footnote-ref-1)
2. Remember that *target measures* are the EF scores included in the main analyses of the main manuscript; that is, mean RT for correct responses for incongruent trials in the Flanker and Stroop tasks (for inhibition), mean RT for correct responses for switch trials in the Colour-Shape and Number-Letter tasks (for switching); and the partial score from the forward and backward Corsi blocks and the Rotation span tasks (Foster et al., 2015); and proportion of correct target trials in the N-Back test (for working memory). [↑](#footnote-ref-2)
3. For example, in a mixed inhibition block that includes incongruent, congruent, and neutral trials, performance for neutral trials is not or is less influenced by EF processes (i.e., neutral trials are the least EF-involving) compared to congruent and incongruent trials. This is because, for instance, sustained control during a mixed Flanker block (e.g., sustained focus on the position of the centre arrow throughout the block) will decrease the influence of incongruent flankers for incongruent trials, decrease facilitation from congruent flankers for congruent trials but will have no or little effect on performance for neutral trials (see e.g., Antoniou, 2023). [↑](#footnote-ref-3)
4. Results were similar, in terms of statistical significance, when the other background variables (Gender, SES, Age, Education) were covaried in the analyses, with the exception of the bidialectal-“monolingual” post-hoc contrast which was marginally non-significant (estimate = 0.25, *SE* = 0.11, *z* value = 2.16, *p* = .055). [↑](#footnote-ref-4)
5. We say *approximately* to allow for the fact that unreliability of the covariate might result in incomplete removal of the covariate variance from the Group factor and the dependent measure (Zinbarg et al., 2010: 309). However, we note that, in our case, the two measures of SMG Vocabulary were highly reliable (Table S4). Also, we expect the composite SMG Vocabulary score from the two measures to be even more reliable (e.g., Cohen, 1988: 535-542; Rushton et al., 1983). Moreover, any potential under-adjustment of the covariate in our analyses would lead to an underestimate of the Group (multilingual, bidialectal) effect rather than to a Type I error because it is “monolinguals” who had a higher SMG Vocabulary. [↑](#footnote-ref-5)
6. Standard Greek is the language spoken in everyday communication by Hellenic Greeks in Greece. In Cyprus, it is taught in school and is used for writing, by the media and, generally, for formal purposes. [↑](#footnote-ref-6)
7. Sibling 1 refers to the 1st born child in the family not counting yourself. [↑](#footnote-ref-7)
8. Sibling 2 refers to the 2nd born child in the family not counting yourself. [↑](#footnote-ref-8)
9. This includes pre-primary, primary education, gymnasium (junior high school) and lyceum (senior high school). [↑](#footnote-ref-9)