

Bilingual Language Experience and Its Effect on Conflict Adaptation in Reactive Inhibitory Control Tasks

Psychological Science
2023, Vol. 34(2) 238–251
© The Author(s) 2022



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/09567976221113764
www.psychologicalscience.org/PS



Jason W. Gullifer^{1,2}, Irina Pivneva^{1,2}, Veronica Whitford^{2,3},
Naveed A. Sheikh^{1,2}, and Debra Titone^{1,2}

1Department of Psychology, McGill University; 2Centre for Research on Brain, Language and Music, McGill University; and 3Department of Psychology, University of New Brunswick

Abstract

We used machine-learning techniques to assess interactions between language and cognitive systems related to inhibitory control and conflict adaptation in reactive control tasks. We built theoretically driven candidate models of Simon and Number Stroop task data ($N = 777$ adult bilinguals ages 18–43 years living in Montréal, Canada) that differed in whether bilingual experience interacted with inhibitory control, including two forms of conflict adaptation: shorter term sequential congruency effects and longer term trial order effects. Models with continuous aspects of bilingual experience provided signal in predicting new, unmodeled data. Specifically, mixed language usage predicted trial order adaptation to conflict. This effect was restricted to Number Stroop, which overtly involves linguistic or symbolic information and relatively higher language- and response-related uncertainty. These results suggest that bilingual experience adaptively tunes aspects of the control system and offers a novel integrative modeling approach that can be used to pursue other complex individual difference questions within the psychological sciences.

Keywords

conflict adaptation, reactive control, individual differences, bilingualism

Received 3/25/21; Revision accepted 6/27/22

Regular use of multiple languages (bilingualism) leads to adaptations in language outcomes, language control, and executive control (Abutalebi & Green, 2016; Bialystok et al., 2012; Green & Abutalebi, 2013; Green & Wei, 2014; Gullifer & Titone, 2021). Bilingual experience exhibits hallmarks of ideal cognitive training that should generalize to nonlinguistic domains (Diamond & Ling, 2016): It is taxing, continual over the lifespan, and socioculturally relevant. However, empirical links between language experience and executive control are variable, leading to questions about scientific practices, transfer between cognitive systems, and recurring epistemological issues (Titone et al., 2017; Titone & Tiv, 2022).

One theoretical perspective asserts that linguistic and cognitive systems do not interact (Paap, Anders-Jefferson, et al., 2019) or that there is weak transfer between systems (Barnett & Ceci, 2002; Sala et al., 2019). Executive control differences for bilinguals versus monolinguals are attributed to a reliance on small sample sizes and

p values, publication bias, researcher degrees of freedom, and questionable research practices (de Bruin et al., 2015; Donnelly et al., 2019; Lehtonen et al., 2018; Paap, 2019; Paap et al., 2015). These practices inflate effect sizes and cause false positives (Chambers, 2013; Cumming, 2014; Easterbrook et al., 1991; Francis, 2012; Gelman & Carlin, 2014). Accordingly, studies with larger samples may reveal null effects for bilingual versus monolingual experience on executive control (Dick et al., 2019; Nichols et al., 2020; Von Bastian et al., 2016).

Another perspective asserts that interactions between systems are conditioned by sociolinguistic contexts or usage practices. For example, dual language contexts

Corresponding Authors:

Jason W. Gullifer, McGill University
Email: jason.gullifer@mail.mcgill.ca; j.gullifer@marianopolis.edu

Debra Titone, McGill University
Email: debra.titone@mcgill.ca

engage control processes (e.g., active goal maintenance and interference control) to a greater extent than single language contexts (Beatty-Martínez et al., 2020; Green & Abutalebi, 2013). The continual engagement of such processes may transfer to broader aspects of executive control. Here, mixed evidence is ascribed to grouping individuals with diverse usage patterns, abilities, and cognitive states into monolithic categories (Beatty-Martínez & Titone, 2021; Grosjean, 1997; Gullifer et al., 2021; Luk & Bialystok, 2013; Navarro-Torres et al., 2021; Salig et al., 2021; Surrain & Luk, 2019) on the basis of socially invented language distinctions (Makoni & Pennycook, 2007; Ortega, 2018; Otheguy et al., 2019). Signal is detected by including diverse aspects of experience using multidimensional continua (Beatty-Martínez & Titone, 2021; Gullifer & Titone, 2020b).

For example, Gullifer and Titone (2020b) modeled a proactive control task (where control is applied ahead of conflict) using a large sample ($N = 450+$) and machine-learning techniques that compared out-of-sample prediction error across models (integrative modeling; Hofman et al., 2021). Large samples yield more accurate effect sizes, and cross-validation offers analytical flexibility while minimizing researcher biases and reliance on p values. Continuous features of bilingual experience (namely, language entropy, which measures joint language use) predicted performance. The effect cross-validated, minimizing prediction error relative to models with trial type alone. Thus, joint language engagement provides signal in predicting proactive control (i.e., active goal maintenance) in the Montréal, Canada, context.

Conflict is also managed reactively through immediate interference suppression, a central construct to bilingual control. Thus, an open question is whether cognitive models of reactive inhibitory control are structured similarly. Reactive control is commonly measured through tasks such as Simon Arrows and Number Stroop (Fig. 1). These tasks assess conflict monitoring and adaptation but differ in two ways, resulting in higher language- and response-related uncertainty for Number Stroop. First, only Number Stroop involves competing language representations via numeric symbols. This uncertainty may be heightened for bilinguals who experience cross-language competition for routinized number labels (Meuter & Allport, 1999). Second, Number Stroop has more response options (four) than Simon Arrows (two), heightening response uncertainty. The higher language- and response-related uncertainty might better simulate control demands experienced by bilinguals (Gullifer & Titone, 2021).

Importantly, people adapt performance in the presence of conflict. Often, conflict adaptation is assessed narrowly, in the short term, by measuring changes in

Statement of Relevance

How might language influence thought? Bilingualism—or the frequent use of more than one language—has been thought to confer advantages to broader aspects of cognition, like the ability to attend to pertinent information and ignore irrelevant information. Yet evidence for this possibility has been mixed, possibly because of failure to recognize that there are various types of bilingualism. That is, participants from diverse language backgrounds are commonly lumped together into groups of “bilinguals” and “monolinguals”, when, in reality, these categorical distinctions are not clear-cut. Here, we employ a highly rigorous statistical approach to data from a large sample of diverse bilinguals, and we measure their experiences using sensitive measures. We find that these measures relate to cognition in predictable ways: we observed cognitive impacts of bilingualism that were most prominent for tasks with high language-related uncertainties, similar to those that bilingual individuals experience in their day-to-day lives.

conflict effects as a function of immediately preceding trial types, known as sequential congruency effects (Gratton et al., 1992). Typically, conflict effects are smaller after incongruent versus congruent trials. This could reflect immediate increases in control demands following conflict (Botvinick et al., 2001) or passive carryover from prior application of control (Hubbard et al., 2017; Mansouri et al., 2009; Paap, Myuz, et al., 2019). However, as people progress through tasks, they gain familiarity with task parameters and conflict, which could result in longer term adaptations: faster task performance (speedups on all trials) and smaller conflict effects as more trials are seen (trial order adaptation).

At the group level, some studies have found bilinguals to be less susceptible to short-term sequential congruency effects than monolinguals, suggesting that bilinguals effectively disengage from prior conflict (Grundy et al., 2017). However, other researchers have questioned the generality and interpretation of these effects (Goldsmith & Morton, 2018; Paap, Myuz, et al., 2019; but see Grundy & Bialystok, 2019). There is also a well-documented reaction time (RT) advantage for bilinguals versus monolinguals (Bialystok et al., 2004; Costa et al., 2009; Hilchey & Klein, 2011; Paap, Myuz, et al., 2019; but see Hilchey et al., 2015). RT advantages suggest enhanced conflict monitoring due to bilingualism. Meta-analytic effect sizes are small ($g = 0.06$) and

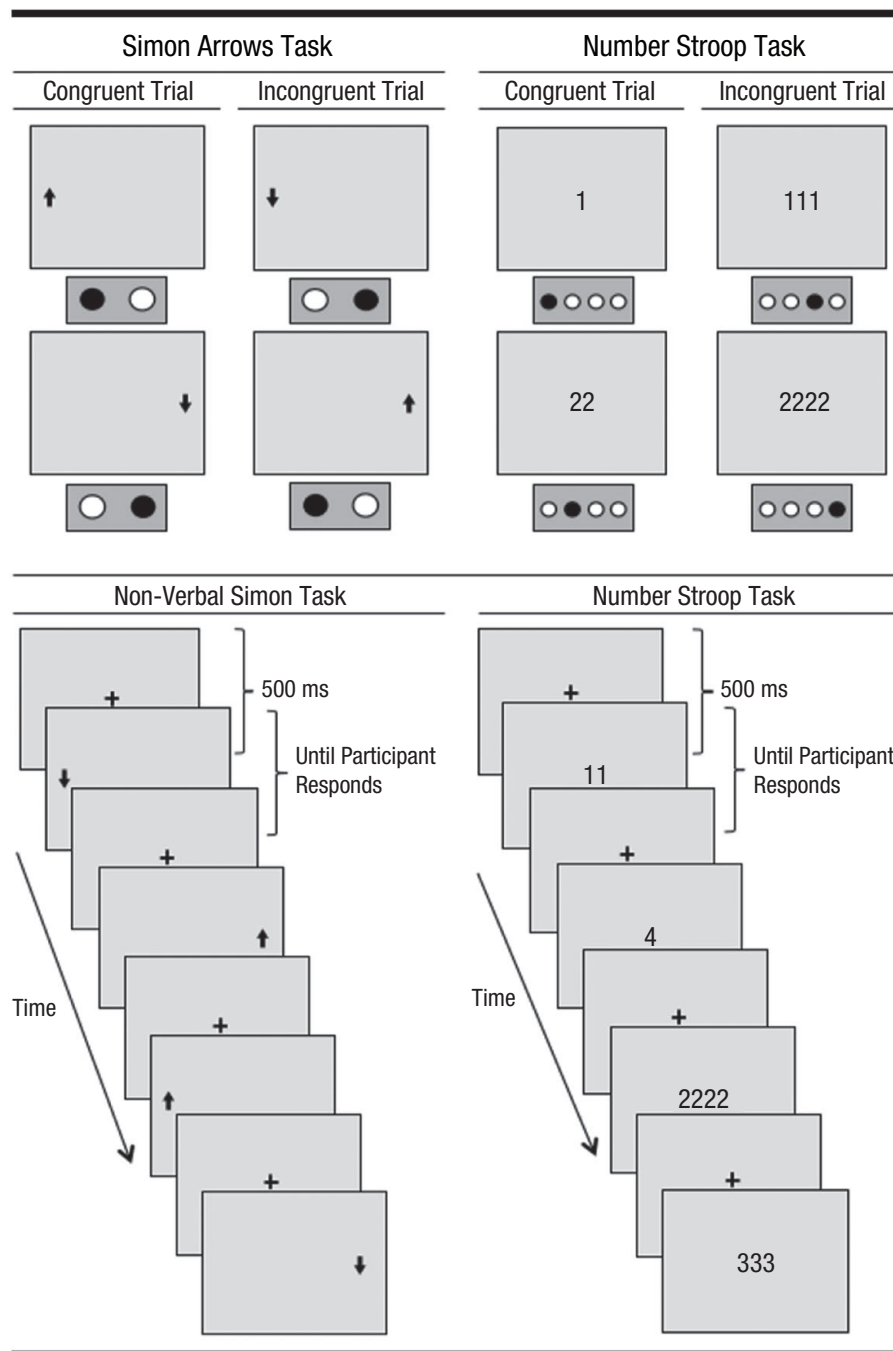


Fig. 1. Illustration of the Simon Arrows (left panel) and Number Stroop (right panel) tasks. In the Simon Arrows task, conflict arises between the position of an arrow in a display (left side vs. right side) and the location of a desired response (left response vs. right response). In the Number Stroop task, conflict arises between the number of digits that appear on a display and the quantity specified by those digits. Trials with conflict (incongruent trials) result in costs in reaction times (RTs) and accuracies relative to those without conflict (congruent trials). In tasks in which trial types are equally likely, participants cannot reliably predict upcoming trials, and control is exercised reactively.

do not survive publication bias corrections (Lehtonen et al., 2018), but meta-analytic approaches may obscure group differences (Grundy, 2020), particularly for samples with heterogeneous language practices. Within

bilingual samples (our focus here), joint language engagement matters. Frequent code switchers (who interleave multiple languages) exhibit sequential congruency effects when they encounter code-switched

Table 1. Participant Demographics

Statistic	<i>N</i>	<i>M</i>	<i>SD</i>
Chronological age (years)	771	22.51	3.63
L2 AoA (years)	777	6.39	4.22
Years bilingual	771	16.11	5.03
L1 current exposure (/100)	777	61.73	22.30
L2 current exposure (/100)	777	35.61	21.95
L3 current exposure (/100)	384	5.37	6.89
Mixed use (1–5)	739	3.79	1.43

Note: This includes self-report information about chronological age, age of acquisition of the second language (L2), and current exposure to each language (first language, or L1, to third language, or L3). AoA = age of acquisition.

sentences in a Flanker task (Adler et al., 2020), particularly in later task blocks. They also exhibit smaller conflict effects as they progress through a traditional Simon task (Kheder & Kaan, 2021). Thus, code switching may progressively increase control demands, but code switchers develop efficiency at resolving conflict. Altogether, joint language engagement is important, but it interacts in complex ways with the presence of conflict, prior trial conflict, and trial order.

Here, we assessed such interactions in Simon Arrows and Number Stroop tasks. We used a large sample ($N = 700+$) of adult, bilingual Montréalers and an integrative modeling approach (Gullifer & Titone, 2020b; Hofman et al., 2021) to compare models that differed in whether aspects of bilingualism predict conflict adaptation, broadly construed. We focused on joint language engagement (mixed use) but also tested other indicators such as age of acquisition (AoA; Kousaie et al., 2017; Tao et al., 2011) and second language (L2) exposure (Gullifer et al., 2021; Subramaniapillai et al., 2018). Given past work, we expected a modulation of sequential congruency effects or trial order adaptation effects as a function of mixed use as well as a relationship between aspects of bilingual experience and global RT.

The studies presented here were not preregistered. Readers may access the code for the analytical routines on OSF (<https://osf.io/axrtv/>). Data will be made available upon request by contacting the corresponding author.

Method

Participants

We analyzed data from 777 bilingual participants who reported language history information (including relative exposure to and use of two or more languages) and completed our target executive control tasks. This sample size is larger than typical for cognitive psychology.

Participants' language history was previously analyzed and reported by Gullifer and Titone (2020a), and a subset ($n = 459$) completed a proactive control task that was previously analyzed and reported by Gullifer and Titone (2020b). Approximately 40% of the sample reported acquiring English as their first language (L1; $n = 320$), and approximately 60% reported acquiring French as their L1 ($n = 457$). Approximately 75% of the sample identified as female ($n = 580$). Table 1 presents summary demographic and language history data. Participants were a mix of university students and members of the larger Montréal community, recruited through McGill's human participant pool system, posted flyers (digital, physical), and word of mouth. All participants provided written informed consent to have their data from experiments and questionnaires collected, stored, and analyzed. The McGill Research Ethics Board approved this research.

Materials

Assessing language experience. Approximately 60% of participants completed a questionnaire previously reported (Gullifer & Titone, 2020a, 2020b). The remaining 40% completed an earlier, shorter questionnaire version. The questionnaires allowed us to probe language experience within the Montréal context. For purposes of the analysis, we extracted or computed several background measures, outlined below.

We extracted two classic measures of L2 experience: L2 AoA (based on the onset of learning) and global exposure to the L2 (Gullifer & Titone, 2020a, 2020b). Participants also reported information about language use in different communicative contexts. Because of changes in the questionnaires over time, the response types of language usage differed. Approximately 40% reported usage in a coarse manner: whether a specific language is used (i.e., French, English, other, or combinations of two or more languages) at home, at work, and in social settings. Approximately 60% reported a more fine-grained Likert score according to the amount of usage of each language (Gullifer & Titone, 2020a, 2020b) on a scale ranging from 1 (*no usage at all*) to 7 (*usage all the time or a significant amount*).

To assess joint language engagement, we computed a mixed-use language measure. First, we converted the Likert scores to coarse usage: We treated scores greater than 1 for any language as usage of that language in that context. We then classified participants into one of five ordered categories on the basis of their joint language engagement (for decision criteria, see Table 2). Overall, mixed-use scores of 1 reflect usage of a single language across contexts, and scores of 5 reflect the engagement of two or more languages in all contexts.

Table 2. Classification Procedure for Mixed Use With Three Social Contexts

Classification decision criterion	Mixed use classification score	Interpretation
Start at 1 as a baseline. This score will reflect usage of the same language in all three contexts.	1	Usage of a single language across contexts
+1 if there exists another unique language in any context	2	Usage of two languages across contexts
+1 if any single context contains two or more languages	3	Occasional joint engagement of two languages
+1 if any single context contains three or more languages -OR- two contexts contain two or more languages	4	Frequent joint engagement of two or more languages
+1 if all contexts contain two or more languages	5	Very frequent joint engagement of two or more languages

Assessing reactive control using Simon Arrows and Number Stroop tasks.

In the Simon Arrows task, participants responded, as quickly and accurately as possible, to the direction (up or down) signified by an arrow on a display, which required a left or right response, respectively. Arrows were positioned on the left or right side of the screen, resulting in congruent trial types (response position and stimulus positions matched) and incongruent trial types (response position and stimulus positions mismatched). Conflict effects were assessed by contrasting incongruent and congruent trials and reflected participants' abilities to inhibit positional information and attend to directional information. Conflict adaptation effects were assessed locally via interactions between trial type and prior trial type. Conflict adaptation effects were assessed globally via interactions between trial type and trial order (a continuous variable reflecting current trial number, scaled and centered). The task contained 80 trials, preceded by a practice block. Half of the trials ($n = 40$) required a congruent response, and half ($n = 40$) required an incongruent response. Figure 1 (left panel) illustrates the task.

We chose the vertical version of the Simon Arrows task because pilot studies and work by Verbruggen and colleagues (2005) suggested that the vertical version results in larger conflict effects compared with the horizontal version. Paap and colleagues (2020) further suggested that the vertical Simon Arrows task evidences lower reliability compared with other tasks. Tasks with lower reliability (i.e., with high between-individuals variability) may be more reliable individual difference markers (Hedge et al., 2018).

In the Number Stroop task, participants saw a display of digits in the center of the screen (e.g., 4444 or 222). The numbers that composed the digits were semantically congruent (e.g., 4444, 22) or incongruent (e.g., 2222, 44) with the response. Participants responded, as quickly and accurately as possible, to the number of digits that

appeared in the display while ignoring the quantity specified by each digit's value. Participants used one of four fingers to press a number key (on the number row of the keyboard; 1, 2, 3, or 4) to indicate how many digits appeared on the screen. Figure 1 (right panel) illustrates the task (button presses indicated as black circles). In the illustration, participants pressed the first button when they saw one digit (e.g., "1"), the second button when they saw two digits (e.g., "22"), the third button when they saw three digits (e.g., "111"), and the fourth button when they saw four digits (e.g., "2222"). The number strings "1" and "22" are examples of congruent trials because the quantity specified by the digits (distractor information: "one" and "two," respectively) matches the number of digits on the screen (response criterion: press Buttons 1 and 2, respectively). In contrast, "111" and "2222" are examples of incongruent trials because the quantity specified by the digits (distractor information: "one" and "two," respectively) conflicts with the number of digits (response criterion: press Buttons 3 and 4, respectively).

Conflict effects were assessed by contrasting incongruent and congruent trials, which provided an indication of participants' abilities to inhibit semantic information and attend to quantity information. Conflict adaptation effects were assessed via conflict effects as a function of prior trial type and trial order. The task contained 60 trials preceded by a practice block. Half of the trials required a congruent response, and half required an incongruent response. Responses were evenly distributed across the four fingers ($n = 15$ trials each).

Results

Data were processed and analyzed in the R programming environment (Version 4; R Core Team, 2021). For trial-level Simon Arrows data, we removed RTs less than

Table 3. Model Specifications for Simon and Number Stroop Analyses

Number	Model fixed effects
1.	Trial × Current
2.	Trial × Prior × Current
3a.	Trial × Prior × Current × L2 AoA + Trial × Prior × Current × L2 Exposure
4.	Trial × Prior × Current × L2 AoA × L2 Exposure
3b.	Trial × Prior × Current × L2 AoA
3c.	Trial × Prior × Current × L2 Exposure
3d.	Trial × Prior × Current × Mixed Use

Note: Trial = trial order; Current = current trial type; Prior = prior trial type; L2 = second language; AoA = age of acquisition. Random effects always included random intercepts by subject (i.e., 1|subject).

150 ms and greater than 2,000 ms as outliers. For trial-level Number Stroop data, we removed RTs less than 250 ms and greater than 2,000 ms as outliers. Outlier detection was based on visual inspection of a density plot for each data set. Prior to analysis, trial type and prior trial type were deviation coded (congruent: −0.5, incongruent: 0.5), and L2 AoA and trial order were z scored.

Using linear mixed-effects regression (lme4; Bates et al., 2015), we fitted several theoretically driven models to RTs and accuracy for each task data set to determine the best model specification. The models differed in whether they included interactions between aspects of bilingual experience and task effects. All linear mixed-effects regression models included a random intercept for subject. We did not add a random slope for trial type by subject because our purpose was to assess fixed effects related to individual differences,

and these would have been completely captured by a random slope by subject. We computed effect sizes (d) for RTs by dividing the slope estimate by the square root of the sum of the random effect variances (following Brysbaert & Stevens, 2018; Gullifer & Titone, 2020b; Westfall et al., 2014), and we computed p values using the degrees of freedom approximated by the Satterthwaite method (Kuznetsova et al., 2017). Model specifications are listed in Table 3.

To select the best model specification, we compared four metrics: Akaike information criterion (AIC), Bayesian information criterion (BIC), likelihood-ratio tests (LRTs), and cross-validation error (described more fully by Gullifer & Titone, 2020b). Table 4 provides a summary of the best model specification selected by each metric for each task, and detailed modeling tables are provided in the Supplemental Material available online. We scrutinized the model that cross-validated.

Simon task

Aggregated data at the subject level are illustrated in Figure 2. Visual inspection of these figures indicates a sequential congruency effect (typical conflict effect after a prior congruent trial; smaller conflict effect after a prior incongruent trial) for both RTs (Fig. 2a) and accuracy (Fig. 2b).

RT analysis. For Simon Arrows RTs, most metrics (AIC, BIC, and LRT) agreed that the best model included interactions between trial order, prior trial type, and current trial type. However, cross-validation error indicated that the best model included an additional interaction with L2 exposure. Model fits are available in Tables S2 and S3 in

Table 4. Summary of Results Depicting the Best Model According to Each Metric or Test

Data set	Measure	AIC	BIC	LRT	CVAL
Simon	RT	Trial × Prior × Current	Trial × Prior × Current	Trial × Prior × Current	Trial × Prior × Current × L2 Exposure
	Accuracy	Trial × Prior × Current	Trial × Prior × Current	Trial × Prior × Current	Trial × Prior × Current
Number Stroop	RT	Trial × Prior × Current × Mixed Use	Trial × Prior × Current	Trial × Prior × Current × Mixed Use	Trial × Prior × Current × Mixed Use
	Accuracy	Trial × Prior × Current × AoA	Trial × Current	Trial × Prior × Current × AoA	Trial × Prior × Current

Note: AIC = Akaike information criterion; BIC = Bayesian information criterion; LRT = likelihood-ratio test; CVAL = cross-validation error; RT = reaction time; Trial = trial order; Prior = prior trial type; Current = current trial type; L2 = second language; AoA = age of acquisition. In the event that there was disagreement between metrics/tests, we selected the model that cross-validated (indicated by bold face). For Simon RTs, we scrutinized the model containing Trial × Prior × Current × L2 Exposure; for Simon accuracy, we scrutinized Trial × Prior × Current; for Number Stroop RTs, we scrutinized Trial × Prior × Current × Mixed Use; and for Number Stroop accuracy, we scrutinized Trial × Prior × Current. More detailed results are reported in the Supplemental Material available online.

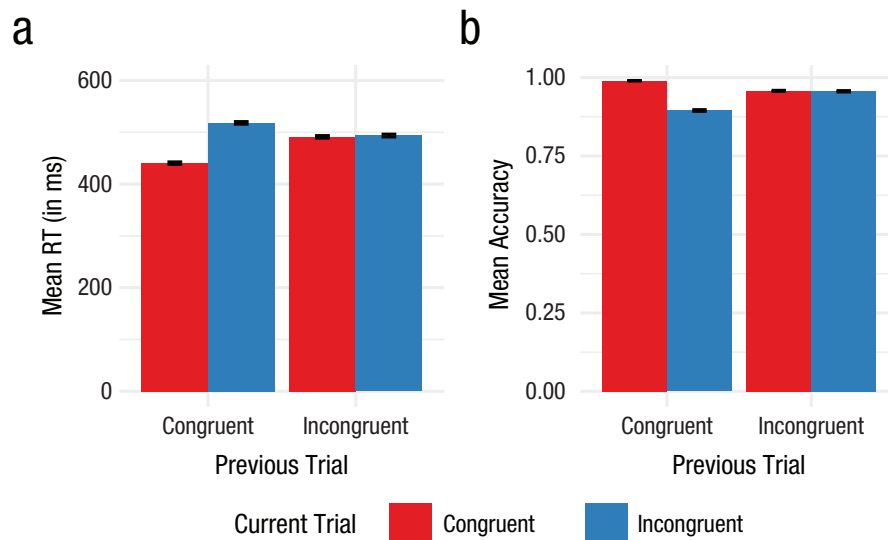


Fig. 2. Subject aggregate (a) reaction times (RTs) and (b) accuracy for the Simon Arrows task. Error bars represent standard error of the mean.

the Supplemental Material. We scrutinized the model that cross-validated and observed the following pattern of results.

There was a significant effect of current trial type ($b = 39.76$, 95% confidence interval, or CI = [37.6, 41.92], $d = 0.28$, $t = 36.05$, $p < .001$), indicating that incongruent trials were slower than congruent trials. Current trial type interacted with prior trial type ($b = -75.57$, 95% CI = [-79.91, -71.22], $d = 0.53$, $t = -34.07$, $p < .001$) and with trial order ($b = -4.2$, 95% CI = [-6.36, -2.03], $d = 0.03$, $t = -3.80$, $p < .001$). This indicated that conflict effects were smaller after an incongruent trial (i.e., the presence of a sequential congruency effect) and toward the end of the task, suggesting that participants exhibited conflict adaptation. There were no significant interactions with L2 exposure, although there was a main effect of L2 exposure ($b = -5.48$, 95% CI = [-10.29, -0.67], $d = 0.04$, $t = -2.23$, $p = .026$). This indicated that participants with higher L2 exposure were faster overall. Note that the main effect of L2 exposure remained stable in an ad hoc model in which it was treated as a simple additive effect ($b = -5.37$, 95% CI = [-10.17, -0.57], $t = -2.19$, $p = .029$). Full regression results are available in Table S4 in the Supplemental Material.

Accuracy analysis. For Simon Arrows accuracy, all metrics agreed that the best model included interactions between trial order, prior trial type, and current trial type. Model fits are available in Tables S5 and S6 in the Supplemental Material. We scrutinized this model and observed the following pattern of results.

There was a significant effect of current trial type, indicating that incongruent trials were less accurate

than congruent trials ($b = -1.23$, 95% CI = [-1.34, -1.12], $z = -21.72$, $p < .001$). Current trial type interacted with prior trial type ($b = 2.44$, 95% CI = [2.22, 2.66], $z = 21.57$, $p < .001$) but not trial order ($b = 0.11$, 95% CI = [0, 0.22], $z = 1.90$, $p = .057$), again indicative of a sequential congruency effect, suggesting that participants exhibited conflict adaptation. However, these accuracy effects did not build over the task. Full regression results are available in Table S7 in the Supplemental Material.

Summary. Overall, there was no indication through inspection of this model that aspects of bilingual experience, as measured in this experiment, modulated conflict effects or sequential congruency effects in the Simon Arrows task. However, L2 exposure modulated RTs across trials. There were no such modulations for accuracy.

Number Stroop task

Aggregated data at the subject level are illustrated in Figure 3. Visual inspection of Figure 3 indicates typical conflict effects in RTs (Fig. 3a) and accuracy (Fig. 3b) and a possible, but small, SCE.

RT analysis. For Number Stroop RTs, most metrics (AIC, LRT, and cross-validation error) agreed that the best model included interactions between trial order, prior trial type, current trial type, and mixed use. Model fits are available in Tables S8 and S9 in the Supplemental Material. We scrutinized the model that cross-validated and observed the following pattern of results.

There was a significant effect of current trial type ($b = 69.63$, 95% CI = [62.45, 76.81], $d = 0.41$, $t = 19.01$, $p < .001$),

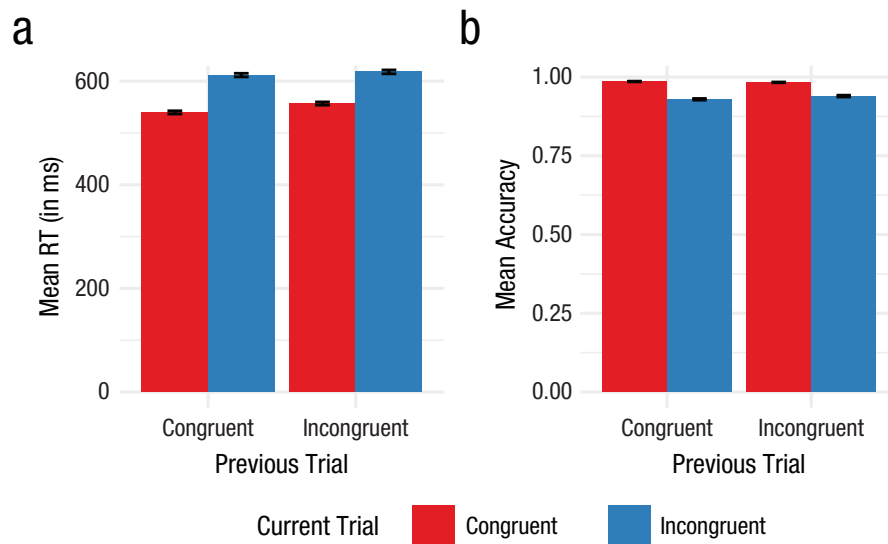


Fig. 3. Subject aggregate (a) reaction times (RTs) and (b) accuracy for the Number Stroop task. Error bars represent standard error of the mean.

indicating that incongruent trials were slower than congruent trials. Current trial type did not interact with prior trial type ($b = -8.67$, 95% CI = $[-23.05, 5.71]$, $d = 0.05$, $t = -1.18$, $p = .237$) or with trial order ($b = 5.55$, 95% CI = $[-0.71, 11.8]$, $d = 0.03$, $t = 1.74$, $p = .082$). Thus, there was no evidence of a sequential congruency effect in RTs to support conflict adaptation at the group level. Crucially, there was a main effect of mixed use ($b = -10.16$, 95% CI = $[-14.45, -5.86]$, $d = 0.06$, $t = -4.64$, $p < .001$) and a three-way interaction between current trial type, trial order, and mixed use ($b = -1.85$, 95% CI = $[-3.51, -0.18]$, $d = 0.01$, $t = -2.17$, $p = .030$). These effects indicated that participants with higher degrees of mixed use were faster on the task. Moreover, for trials that occurred early in the task, participants with high degrees of mixed use exhibited larger conflict effects relative to participants with low degrees of mixed use. However, toward the end of the task, the pattern of results flipped; participants with high degrees of mixed experience reduced conflict effects (for an illustration of the interaction, see Figs. 4 and 5). Full regression results are available in Table S10 in the Supplemental Material.

Accuracy analysis. For Number Stroop accuracy, half of the metrics (AIC and LRT) agreed that the best model included interactions between trial order, prior trial type, current trial type, and L2 AoA, and half (BIC and cross-validation error) agreed on a simpler model without L2 AoA. Model fits are available in Tables S11 and S12 in the Supplemental Material. We scrutinized the simpler model that cross-validated and observed the following pattern of results.

There was a significant effect of current trial type, indicating that incongruent trials were less accurate than congruent trials ($b = -1.52$, 95% CI = $[-1.62, -1.41]$, $z = -29.16$, $p < .001$). Current trial type interacted with prior trial type ($b = 0.36$, 95% CI = $[0.16, 0.57]$, $z = 3.53$, $p < .001$) but not trial order ($b = 0.03$, 95% CI = $[-0.08, 0.13]$, $z = 0.52$, $p = .603$). These effects indicated that conflict effects were smaller after an incongruent trial (i.e., sequential congruency effect), suggesting that participants exhibited conflict adaptation. However, conflict adaptation effects did not build over the task. Full regression results are available in Table S13 in the Supplemental Material. Because the reader may be interested in possible effects of L2 AoA (that did not cross-validate), we provide results in “Effects of L2 AoA on Number Stroop Accuracy” and Table S14 in the Supplemental Material.

Summary. Overall, there was evidence that bilingual experience, measured by mixed use in this experiment, modulated conflict over the task. Participants with high mixed use responded faster and exhibited smaller conflict effects than those with low mixed use, especially as the task progressed. Bilingual experience, as measured in this experiment, did not modulate sequential congruency effects as measured by the Number Stroop task in RTs. Interactions with bilingual experience were specific to RTs; no such effects were observed in accuracy.

Discussion

We modeled reactive inhibitory control performance to test whether individual differences in bilingual

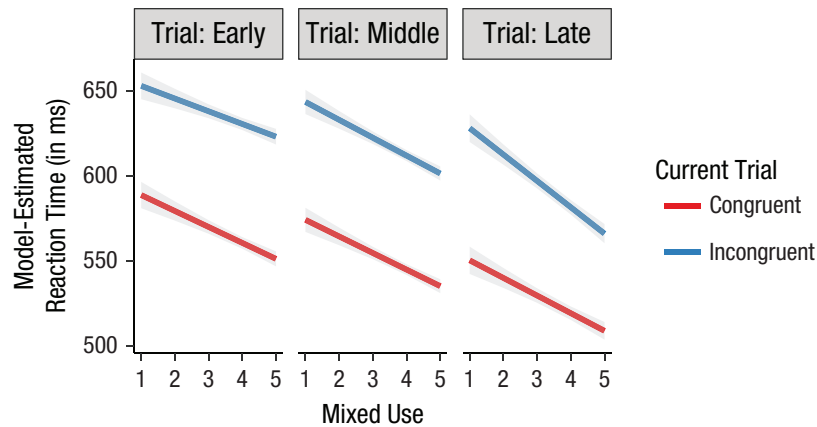


Fig. 4. Illustration of the three-way interaction between current trial type, trial order, and mixed use in the Number Stroop task. This interaction indicates that people with high degrees of mixed use exhibit greater conflict during early trials and less conflict during late trials relative to people with low degrees of mixed use. This suggests that they exhibit a trial order adaptation effect. Error bands represent standard error of the mean.

experience provide signal in predicting out-of-sample data beyond task information. This allowed us to test two broad theoretical perspectives differing in whether linguistic and cognitive systems interact. We modeled Simon Arrows and Number Stroop data with a large sample of bilinguals. Both tasks measured reactive inhibitory control and conflict adaptation but differed in language- and response-related uncertainty. Uncertainty is higher in Number Stroop, which may more precisely simulate control demands experienced by bilinguals (Gullifer & Titone, 2021).

At the group level, we observed engagement of inhibitory control through traditional conflict effects in both

tasks. Incongruent trials were slower and less accurate than congruent trials. We also observed conflict adaptation in both tasks. Conventional sequential congruency effects occurred in RTs and accuracy in Simon Arrows, and in accuracy in Number Stroop. Conflict effects were smaller after incongruent trials than congruent trials, reflecting an increase in control demands from prior conflict (Botvinick et al., 2001) or carryover from the prior application of control (Hubbard et al., 2017; Mansouri et al., 2009; Paap, Myuz, et al., 2019). Conflict effects were also smaller toward the end of the task (an effect restricted to Simon Arrows RTs), reflecting adaptations in conflict management.

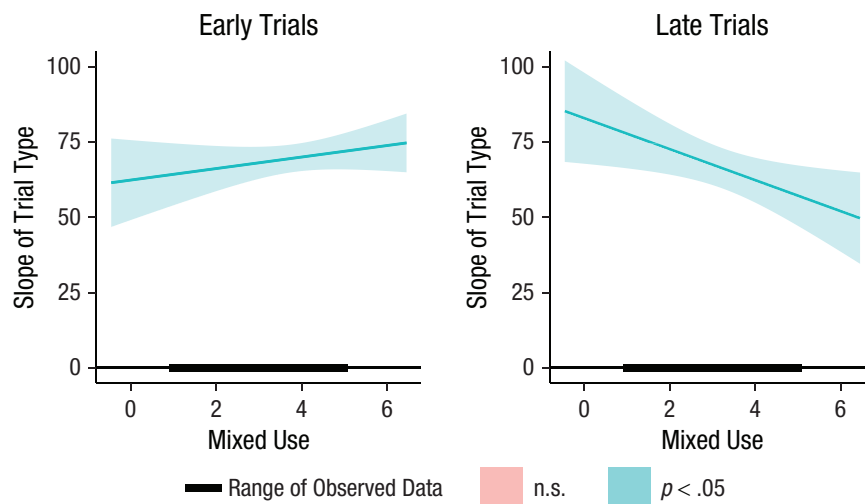


Fig. 5. Johnson-Neyman interval plot illustrating the slope of the trial type effect as a function of trial order and mixed use. This plot illustrates the three-way interaction with respect to conflict.

Crucially, individual differences in bilingual experience modulated performance related to long-term adaptations. We observed robust global speed advantages associated with bilingual experience in Simon Arrows and Number Stroop. In Simon Arrows, participants with higher L2 exposure were faster on all trial types relative to participants with lower L2 exposure. A compatible effect surfaced in Number Stroop as a function of mixed use. Both effects cross-validated, suggesting that they are robust and predictive of new data and relevant to reactive control models. Although this experiment did not explicitly compare monolingual versus bilingual populations, these results are in line with previous literature on conflict monitoring in bilinguals and monolinguals (Bialystok et al., 2004; Costa et al., 2009; Hilchey & Klein, 2011; Paap, Myuz, et al., 2019). They suggest that bilingual language experiences impact the cognitive system, when experiences are measured continuously in accordance with language usage patterns.

In Number Stroop, we also observed a relationship between mixed use, conflict effects, and trial order. People with higher mixed use exhibited larger conflict effects early but smaller conflict effects later relative to people with lower mixed use. This effect was restricted to RTs; accuracy was near ceiling. This suggests that individual differences in bilingual experience relate to efficiency in adapting to conflict between semantic number information. One possibility is that bilinguals with higher mixed use experienced more cross-language competition from the number labels than bilinguals with lower mixed use. These interactions cross-validated, suggesting that they are robust. The task specificity of the interaction might relate to the idea that Number Stroop more readily simulates bilingual language control demands than Simon Arrows.

Overall, these results extend work on proactive control by Gullifer and Titone (2020b) to reactive inhibitory control. There, bilinguals with higher language entropy, reflective of high joint language engagement, exhibited greater engagement of proactive control and active goal maintenance in a nonlinguistic task in which prior context is crucial for performance. Here, bilinguals with higher mixed use, indexing a similar construct (see “Correlational Analysis” in the Supplemental Material), exhibited greater conflict adaptation in Number Stroop compared with bilinguals with lower mixed use. These differences may reflect greater demands placed on interference control (including conflict monitoring and interference suppression) in highly bilingual contexts in which two languages must be regularly controlled (dual language contexts; Green & Abutalebi, 2013; Gullifer & Titone, 2020b).

The results are compatible with work showing that aspects of bilingualism relate to conflict adaptation (Adler et al., 2020; Grundy et al., 2017; Kheder & Kaan, 2021), but with some differences. First, modulations here occurred for long-term adaptation effects but not short-term sequential congruency effects. Thus, sequential congruency effects may not necessarily reflect adaptive attentional processes but instead reflect carryover effects from prior trials, particularly for tasks with equal proportions of trial types (Hubbard et al., 2017; Mansouri et al., 2009; Paap, Myuz, et al., 2019). Alternatively, sequential congruency effects may reflect processes that are adapted by some sociolinguistic contexts (e.g., Toronto, where bilingualism often involves English and another language) but not others (e.g., Montréal, where multilingualism almost always includes French, English, and another language). Methodological differences (e.g., whether monolingual samples were included) may also have contributed to differences between studies. In other words, the results observed here do not invalidate work done in these other locations.

In contrast, long-term adaptation effects observed here for Number Stroop reflect improvements in performance. Taken together, long-term adaptation effects vary within populations and can be accounted for by aspects of social language experience that involve joint language engagement (Adler et al., 2020; Kheder & Kaan, 2021).

Another key difference between this and prior studies is that, besides RT advantages, conflict adaptation effects were not observed in the Simon Arrows task. One possible reason is that these tasks are unreliable in measuring the targeted construct. Other researchers have argued that executive control tasks do not correlate with one another, indicating poor task validity (Paap et al., 2020). However, tasks recruit multiple control processes and may differ in their recruitment strength according to individual and task differences. Thus, it may be too simplistic to expect high between-tasks correlations (Salig et al., 2021). Another possibility is rooted in differences between linguistic and social contexts across studies, impacting transfer between cognitive systems (Barnett & Ceci, 2002; Sala et al., 2019).

Montréal is a uniquely multilingual environment (e.g., Tiv et al., 2021, 2022) relative to other locations such as Florida (Adler et al., 2020; Kheder & Kaan, 2021) and Ontario (Goldsmith & Morton, 2018; Grundy et al., 2017), where English is predominant. Thus, Montréalers who jointly engage multiple languages must frequently manage competing language representations and response options. This coheres with the idea that far transfer occurs for socioculturally relevant cognitive training (Diamond & Ling, 2016). Whereas our cross-validation

results provide evidence that the results will generalize within Montréal, an open question is whether they generalize to other comparable sociolinguistic contexts. Moreover, the generalizability of our results may likely be affected by other factors, such as sampling bias. For example, the sample was overwhelmingly female, and we recruited individuals based on their bilingualism (and willingness to participate in a language study), as opposed to random sampling.

Importantly, the idea and findings that aspects of language usage might impact executive control abilities call into question generalizability of a vast number of studies on various psychological constructs that were restricted to group-level analyses of single populations. A broad assumption in the psychological sciences is that group-level results should generalize to all populations, when this is not thoroughly tested. Going forward, researchers could be more rigorous, not only in terms of methodology but also in exploring the generalizability of our constructs to different populations. The field of bilingualism champions iterative scientific discovery, where theoretical perspectives are reformulated as new data are acquired (Navarro-Torres et al., 2021).

Our results bolster theoretical perspectives that treat language and executive control systems as interactive and adaptive. Although the findings are subtle in terms of effect size and cross-validation error reduction, they highlight the power of treating language experience as a multidetermined spectrum of continuous variables versus an arbitrary dichotomy (e.g., bilinguals vs. monolinguals). Crucially, past studies would have treated this sample as one homogeneous group, when in fact there is a diversity of experiences that systematically predict performance outcomes.

Transparency

Action Editor: Angela Lukowski

Editor: Patricia J. Bauer

Author Contributions

J. W. Gullifer, I. Pivneva, and D. Titone developed the study concept and contributed to the study design. I. Pivneva, V. Whitford, and N. A. Sheikh conducted testing, data collection, and aggregation. J. W. Gullifer and D. Titone analyzed and interpreted the data. J. W. Gullifer drafted the manuscript, with input from D. Titone. V. Whitford and D. Titone provided revisions. All the authors approved the final manuscript for submission.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This work was supported by the Natural Sciences and Engineering Research Council of Canada (Discovery Grant 264146 to D. Titone), the Social Sciences and Humanities

Research Council of Canada (Insight Development Grant 01037 to J. W. Gullifer and D. Titone), and the Centre for Research on Brain, Language and Music.

Open Practices

The studies presented here were not preregistered. Readers may access the code for the analytical routines on OSF (<https://osf.io/axrtv/>). Data will be made available upon request by contacting the corresponding author.

ORCID iD

Jason W. Gullifer  <https://orcid.org/0000-0003-2959-9072>

Supplemental Material

Additional supporting information can be found at <http://journals.sagepub.com/doi/suppl/10.1177/09567976221113764>

References

- Abutalebi, J., & Green, D. W. (2016). Neuroimaging of language control in bilinguals: Neural adaptation and reserve. *Bilingualism: Language and Cognition*, 19(4), 689–698. <https://doi.org/10.1017/s1366728916000225>
- Adler, R. M., Valdés Kroff, J. R., & Novick, J. M. (2020). Does integrating a code-switch during comprehension engage cognitive control? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(4), 741–759. <https://doi.org/10.1037/xlm0000755>
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, 128(4), 612–637.
- Bates, D., Machler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Beatty-Martínez, A. L., Navarro-Torres, C. A., Dussias, P. E., Bajo, M. T., Guzzardo Tamargo, R. E., & Kroll, J. F. (2020). Interactional context mediates the consequences of bilingualism for language and cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(6), 1022–1047. <https://doi.org/10.1037/xlm0000770>
- Beatty-Martínez, A. L., & Titone, D. A. (2021). The quest for signals in noise: Leveraging experiential variation to identify bilingual phenotypes. *Languages*, 6(4), Article 168. <https://doi.org/10.3390/languages6040168>
- Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, 19(2), 290–303. <https://doi.org/10.1037/0882-7974.19.2.290>
- Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences*, 16(4), 240–250. <https://doi.org/10.1016/j.tics.2012.03.001>
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108, 624–652. <https://doi.org/10.1037/0033-295X.108.3.624>
- Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, 1(1), Article 9. <https://doi.org/10.5334/joc.10>

- Chambers, C. D. (2013). Registered reports: A new publishing initiative at Cortex. *Cortex*, 49(3), 609–610. <https://doi.org/10.1016/j.cortex.2012.12.016>
- Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, 113(2), 135–149. <https://doi.org/10.1016/j.cognition.2009.08.001>
- Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, 25(1), 7–29. <https://doi.org/10.1177/0956797613504966>
- de Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive advantage in bilingualism: An example of publication bias? *Psychological Science*, 26(1), 99–107. <https://doi.org/10.1177/0956797614557866>
- Diamond, A., & Ling, D. S. (2016). Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not. *Developmental Cognitive Neuroscience*, 18, 34–48. <https://doi.org/10.1016/j.dcn.2015.11.005>
- Dick, A. S., Garcia, N. L., Pruden, S. M., Thompson, W. K., Hawes, S. W., Sutherland, M. T., Riedel, M. C., Laird, A. R., & Gonzalez, R. (2019). No evidence for a bilingual executive function advantage in the nationally representative ABCD study. *Nature Human Behaviour*, 3(7), 692–701. <https://doi.org/10.1038/s41562-019-0609-3>
- Donnelly, S., Brooks, P. J., & Homer, B. D. (2019). Is there a bilingual advantage on interference-control tasks? A multiverse meta-analysis of global reaction time and interference cost. *Psychonomic Bulletin & Review*, 26(4), 1122–1147. <https://doi.org/10.3758/s13423-019-01567-z>
- Easterbrook, P. J., Gopalan, R., Berlin, J. A., & Matthews, D. R. (1991). Publication bias in clinical research. *The Lancet*, 337(8746), 867–872.
- Francis, G. (2012). Too good to be true: Publication bias in two prominent studies from experimental psychology. *Psychonomic Bulletin & Review*, 19(2), 151–156.
- Gelman, A., & Carlin, J. (2014). Beyond power calculations: Assessing type S (sign) and type M (magnitude) errors. *Perspectives on Psychological Science*, 9(6), 641–651.
- Goldsmith, S. F., & Morton, J. B. (2018). Sequential congruency effects in monolingual and bilingual adults: A failure to replicate Grundy et al. (2017). *Frontiers in Psychology*, 9, Article 2476. <https://doi.org/10.3389/fpsyg.2018.02476>
- Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General*, 121(4), 480–506.
- Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Neuroscience*, 25(5), 515–530. <https://doi.org/10.1080/20445911.2013.796377>
- Green, D. W., & Wei, L. (2014). A control process model of code-switching. *Language Cognition & Neuroscience*, 29, 499–511. <https://doi.org/10.1080/23273798.2014.882515>
- Grosjean, F. (1997). The bilingual individual. *Interpreting*, 2(1), 163–187.
- Grundy, J. G. (2020). The effects of bilingualism on executive functions: An updated quantitative analysis. *Journal of Cultural Cognitive Science*, 4(2), 177–199. <https://doi.org/10.1007/s41809-020-00062-5>
- Grundy, J. G., & Bialystok, E. (2019). When a “replication” is not a replication. Commentary: Sequential congruency effects in monolingual and bilingual adults. *Frontiers in Psychology*, 10, Article 797. <https://www.frontiersin.org/article/10.3389/fpsyg.2019.00797>
- Grundy, J. G., Chung-Fat-Yim, A., Friesen, D. C., Mak, L., & Bialystok, E. (2017). Sequential congruency effects reveal differences in disengagement of attention for monolingual and bilingual young adults. *Cognition*, 163, 42–55. <https://doi.org/10.1016/j.cognition.2017.02.010>
- Gullifer, J. W., Kousaie, S., Gilbert, A. C., Grant, A., Giroud, N., Coulter, K., Klein, D., Baum, S., Phillips, N., & Titone, D. (2021). Bilingual language experience as a multidimensional spectrum: Associations with objective and subjective language proficiency. *Applied Psycholinguistics*, 42(2), 1–34. <https://doi.org/10.1017/S0142716420000521>
- Gullifer, J. W., & Titone, D. (2020a). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*, 23(2), 283–294. <https://doi.org/10.1017/s1366728919000026>
- Gullifer, J. W., & Titone, D. (2020b). Engaging proactive control: Influences of diverse language experiences using insights from machine learning. *Journal of Experimental Psychology: General*, 150(3), 414–430. <https://doi.org/10.1037/xge0000933>
- Gullifer, J. W., & Titone, D. (2021). Bilingualism: A neurocognitive exercise in managing uncertainty. *Neurobiology of Language*, 2(4), 1–43. https://doi.org/10.1162/nol_a_00044
- Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, 50(3), 1166–1186. <https://doi.org/10.3758/s13428-017-0935-1>
- Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review*, 18(4), 625–658. <https://doi.org/10.3758/s13423-011-0116-7>
- Hilchey, M. D., Saint-Aubin, J., & Klein, R. M. (2015). Does bilingual exercise enhance cognitive fitness in traditional non-linguistic executive processing tasks? In J. W. Schwieter (Ed.), *The Cambridge handbook of bilingual processing* (pp. 586–613). Cambridge University Press.
- Hofman, J. M., Watts, D. J., Athey, S., Garip, F., Griffiths, T. L., Kleinberg, J., Margetts, H., Mullainathan, S., Salganik, M. J., Vazire, S., Vespignani, A., & Yarkoni, T. (2021). Integrating explanation and prediction in computational social science. *Nature*, 595(7866), 181–188. <https://doi.org/10.1038/s41586-021-03659-0>
- Hubbard, J., Kuhns, D., Schäfer, T. A. J., & Mayr, U. (2017). Is conflict adaptation due to active regulation or passive carry-over? Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(3), 385–393. <https://doi.org/10.1037/xlm0000306>
- Kheder, S., & Kaan, E. (2021). Cognitive control in bilinguals: Proficiency and code-switching both matter. *Cognition*,

- 209, Article 104575. <https://doi.org/10.1016/j.cognition.2020.104575>
- Kousaie, S., Chai, X. J., Sander, K. M., & Klein, D. (2017). Simultaneous learning of two languages from birth positively impacts intrinsic functional connectivity and cognitive control. *Brain and Cognition*, 117, 49–56. <https://doi.org/10.1016/j.bandc.2017.06.003>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Lehtonen, M., Soveri, A., Laine, A., Jarvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin*, 144(4), 394–425. <https://doi.org/10.1037/bul0000142>
- Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, 25(5), 605–621. <https://doi.org/10.1080/20445911.2013.795574>
- Makoni, S., & Pennycook, A. (2007). Disinventing and (re) constituting languages. In S. Makoni & A. Pennycook (Eds.), *Disinventing and reconstituting languages* (Vol. 2, pp. 137–156). Routledge. https://doi.org/10.1207/s15427595cils0203_1
- Mansouri, F. A., Tanaka, K., & Buckley, M. J. (2009). Conflict-induced behavioural adjustment: A clue to the executive functions of the prefrontal cortex. *Nature Reviews Neuroscience*, 10(2), 141–152. <https://doi.org/10.1038/nrn2538>
- Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, 40, 25–40. <https://doi.org/10.1006/jmla.1998.2602>
- Navarro-Torres, C. A., Beatty-Martínez, A. L., Kroll, J. F., & Green, D. W. (2021). Research on bilingualism as discovery science. *Brain and Language*, 222, Article 105014. <https://doi.org/10.1016/j.bandl.2021.105014>
- Nichols, E. S., Wild, C. J., Stojanoski, B., Battista, M. E., & Owen, A. M. (2020). Bilingualism affords no general cognitive advantages: A population study of executive function in 11,000 people. *Psychological Science*, 31(6), 548–567. <https://doi.org/10.1177/0956797620903113>
- Ortega, L. (2018). Ontologies of language, second language acquisition, and world Englishes. *World Englishes*, 37(1), 64–79. <https://doi.org/10.1111/weng.12303>
- Otheguy, R., García, O., & Reid, W. (2019). A translanguaging view of the linguistic system of bilinguals. *Applied Linguistics Review*, 10(4), 625–651. <https://doi.org/10.1515/applirev-2018-0020>
- Paap, K. R. (2019). The bilingual advantage debate: Quantity and quality of the evidence. In J. W. Schwieter & M. Paradis (Eds.), *The handbook of the neuroscience of multilingualism* (pp. 701–735). Wiley-Blackwell.
- Paap, K. R., Anders-Jefferson, R., Mikulinsky, R., Masuda, S., & Mason, L. (2019). On the encapsulation of bilingual language control. *Journal of Memory and Language*, 105, 76–92. <https://doi.org/10.1016/j.jml.2018.12.001>
- Paap, K. R., Anders-Jefferson, R., Zimiga, B., Mason, L., & Mikulinsky, R. (2020). Interference scores have inadequate concurrent and convergent validity: Should we stop using the flanker, Simon, and spatial Stroop tasks? *Cognitive Research: Principles and Implications*, 5(1), Article 7. <https://doi.org/10.1186/s41235-020-0207-y>
- Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, 69, 265–278. <https://doi.org/10.1016/j.cortex.2015.04.014>
- Paap, K. R., Myuz, H., Anders-Jefferson, R., Mason, L., & Zimiga, B. (2019). On the ambiguity regarding the relationship between sequential congruency effects, bilingual advantages in cognitive control, and the disengagement of attention. *AIMS Neuroscience*, 6(4), 282–298. <https://doi.org/10.3934/Neuroscience.2019.4.282>
- R Core Team. (2021). *R: A language and environment for statistical computing* (Version 4) [Computer software]. R Foundation for Statistical Computing. <http://www.R-project.org>
- Sala, G., Aksayli, N. D., Tatlidil, K. S., Tatsumi, T., Gondo, Y., & Gobet, F. (2019). Near and far transfer in cognitive training: A second-order meta-analysis. *Collabra: Psychology*, 5, Article 18. <https://doi.org/10.1525/collabra.203>
- Salig, L. K., Valdés Kroff, J. R., Slevc, L. R., & Novick, J. M. (2021). Moving from bilingual traits to states: Understanding cognition and language processing through moment-to-moment variation. *Neurobiology of Language*, 2(4), 487–512. https://doi.org/10.1162/nol_a_00046
- Subramaniapillai, S., Rajah, M. N., Pasvanis, S., & Titone, D. (2018). Bilingual experience and executive control over the adult lifespan: The role of biological sex. *Bilingualism: Language and Cognition*, 22(4), 733–751. <https://doi.org/10.1017/s1366728918000317>
- Surraín, S., & Luk, G. (2019). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015. *Bilingualism: Language and Cognition*, 22(2), 401–415. <https://doi.org/10.1017/S1366728917000682>
- Tao, L., Marzecová, A., Taft, M., Asanowicz, D., & Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: The role of age of acquisition. *Frontiers in Psychology*, 2, Article 123. <https://doi.org/10.3389/fpsyg.2011.00123>
- Titone, D., Gullifer, J. W., Subramaniapillai, S., Rajah, N., & Baum, S. (2017). History-inspired reflections on the bilingual advantages hypothesis. In E. Bialystok & M. D. Sullivan (Eds.), *Growing old with two languages: Effects of bilingualism on cognitive aging* (pp. 265–295). John Benjamins Publishing Company.
- Titone, D., & Tiv, M. (2022). Rethinking multilingual experience through a systems framework of bilingualism. *Bilingualism: Language and Cognition*. Advance online publication. <https://doi.org/10.1017/S1366728921001127>
- Tiv, M., Kutlu, E., Gullifer, J. W., Feng, R. Y., Doucerain, M. M., & Titone, D. A. (2022). Bridging interpersonal

- and ecological dynamics of cognition through a systems framework of bilingualism. *Journal of Experimental Psychology: General*. Advance online publication. <https://doi.org/10.1037/xge0001174>
- Tiv, M., Kutlu, E., & Titone, D. (2021). Bilingualism moves us beyond the ideal speaker narrative in cognitive psychology. In W. S. Francis (Ed.), *Bilingualism across the lifespan* (pp. 29–46). Routledge.
- Verbruggen, F., Liefoghe, B., Notebaert, W., & Vandierendonck, A. (2005). Effects of stimulus–stimulus compatibility and stimulus–response compatibility on response inhibition. *Acta Psychologica*, 120(3), 307–326. <https://doi.org/10.1016/j.actpsy.2005.05.003>
- Von Bastian, C. C., Souza, A. S., & Gade, M. (2016). No evidence for bilingual cognitive advantages: A test of four hypotheses. *Journal of Experimental Psychology: General*, 145(2), 246–258.
- Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, 143(5), 2020–2045.