# Language-specific cognitive flexibility is related to code-switching habits and interactional context; domain-general cognitive flexibility is not

**Authors:**

Leah Gosselin (http://orcid.org/0000-0001-5245-0589)

Laura Sabourin (http://orcid.org/0000-0003-1019-020X)

ERP-Ling Lab, Department of Linguistics, University of Ottawa, Ottawa, Canada

**Keywords:** Dual-language; executive functioning; bilingualism; mixing; proactive control

**This is the peer-reviewed, Author-accepted version of the following article:**

Gosselin, L., & Sabourin, L. (2024). Language-specific cognitive flexibility is related to code-switching habits and interactional context; domain-general cognitive flexibility is not. *Journal of Cognitive Psychology*, *36*(6), 688–708. https://doi.org/10.1080/20445911.2024.2365463

**The Version of Record of this manuscript has been published and is available** in *Journal of Cognitive Psychology* <17 Jun 2024> https://www.tandfonline.com/doi/full/10.1080/20445911.2024.2365463.

## Abstract:

Individualistic experience with code-switching has often been found to modulate bilingual executive functioning, though the direction of these effects is variable. The present study investigated whether French-English code-switching in a primarily dual-language context (the environment which requires the most control processes) may lead to increased cognitive flexibility. Sustained (mixing) and transient (switching) cognitive flexibility was examined in a domain-general task and a novel language-specific task (i.e. a cued bilingual lexical decision task). First, mixing and switching effects in the domain-general task were not reliably predicted by the participants' code-switching habits. Second, though the sample displayed minimal switching effects and a mixing benefit in the language-specific task, these were positively predicted by the participants' deliberate code-switching. By contrast, predictors related to dense code-switching were negatively related to the participants' language-specific sustained cognitive flexibility. Altogether, our findings indicate that any training instilled by dual-language code-switching is restricted to language-specific cognitive flexibility.

## 1. INTRODUCTION

As early as sixty years ago, researchers investigating intelligence noted that bilingualism may lead to increased "mental flexibility" (Peal & Lambert, 1962, p. 22). Since then, hundreds of published papers have studied the relationship between bi-/multilingualism and the core executive functions (see Diamond, 2013), leading to the staunchly debated Bilingual Advantage Hypothesis (BAH; see Bialystok, 2017; Paap, 2023 for extensive discussions). This hypothesis posits that the constant and effortful management of two or more co-activated languages (Costa et al., 2000) may lead to increased domain-general executive functioning. Though much empirical evidence supports the existence of bilingualism-related differences (see Bialystok et al., 2012 for a review), countless experiments have yielded null results (e.g. Paap & Greenberg, 2013). Systematic meta-analyses suggest that, if bilinguals indeed exhibit increased executive functioning skills, these effects are modest and inconsistent (e.g. Lehtonen et al., 2018; Nichols et al., 2020; cf. Grundy, 2020). This has pushed researchers to tease apart various dimensions of the bilingual experience, in an effort to better understand the specific circumstances that may contribute to executive functioning differences among bi-/multilingual individuals.

As such, the present paper adopts a within-group approach (i.e. juxtaposing bilinguals with other bilinguals, rather than comparing bilinguals to monolinguals; Soveri et al., 2011) to examine the cognitive effects related to a specific feature of bilingual performance: CODE-SWITCHING, or the act of using multiple languages within a single communicative event (Poplack, 1980). We explore whether code-switching habits impact language-specific and domain-general cognitive flexibility, particularly for individuals who originate from communities which primarily promote dual-language environments, the most cognitively demanding interactional context (Green & Abutalebi, 2013).

### 1.1 Cognitive flexibility 101

In a seminal paper, Miyake and colleagues (Miyake et al., 2000) demonstrated that the three commonly recognised cognitive processes–inhibition, shifting, and information updating–, while certainly interrelated, could also be isolated by the use of specific executive tasks. Of the three core executive functions (EFs), shifting perhaps best matches what was described as "mental flexibility" by Peal and Lambert (1962). Shifting, which is also sometimes called mental set shifting, task switching, or cognitive flexibility, refers to an individual's ability to adapt to the requirements of a given circumstance or task, especially within a changing environment. Individuals with high cognitive flexibility can disengage from a given perspective, reconfigure their mental sets, and efficiently engage in the shifted perspective. This process necessarily implicates inhibitory control, as the original mental set must be suppressed (e.g. Diamond, 2013). Cognitive flexibility appears to be governed by pre-frontal and supplementary motor cortices (Moll et al., 2002); lesions in these areas are sometimes characterised by context-inappropriate behavioural perseverations (Stuss et al., 2001).

Cognitive flexibility has been examined using both language-specific (e.g. number-letter switching, semantic fluency, picture naming) and domain-general (e.g. task/feature switching, dimensional card sorting, trail making) tasks. Within these paradigms, participants are typically asked to identify a specific feature of a target according to a pre-determined cue (e.g. colour). A cue can be consistent within a block (single-cue block), or it can alternate within a block (mixed-cue block). Furthermore, the mixed-cue block is composed of targets from two conditions: repeated-cue targets are trials wherein the previous target ($n$-1) was cued by the same feature (e.g. colour-to-colour); switched-cue targets are trials wherein the previous target was cued by a different feature (e.g. colour-to-shape). Two indices of cognitive flexibility can thus be

computed: (a) the MIXING EFFECT or MIX COST is the difference between single-cue and mixed-cue blocks; (b) the SWITCHING EFFECT or SWITCH COST is the difference between repeated-cue and switched-cue trials. The idea is that mixed-cue blocks, particularly the switched-cue trials within the mixed block, implicate adaptation and mental set shifting, which incurs longer response times (Rogers & Monsell, 1995).

Note that the mixing and switching effects are associated to disparate cognitive mechanisms. Since mixed-cue blocks involve the regulation of antagonistic task sets over a considerable length of time, mixing effects have been linked to more global or sustained monitoring processes, which fall under the umbrella of *proactive* control (Braver, 2012; Braver et al., 2003; Rubin & Meiran, 2005). On the other hand, efficiently switching between cues implies the lack of rigidity or perseveration of a mental set at a given moment. Switching effects are thus said to be related to local and transient "time-sensitive demands" (Prior & Gollan, 2013, p. 624), or to *reactive* control (Braver, 2012; Braver et al., 2003; Diamond, 2013).

Previous researchers have argued that the bilingual experience involves both types of cognitive demands: the maintenance of a given language requires sustained cognitive flexibility, while transient cognitive flexibility allows for dynamic switching between languages according to context (Prior & Gollan, 2013). Conforming to this line of reasoning, some studies have indeed found that bilinguals outperform monolinguals when it comes to tasks tapping into cognitive flexibility (e.g. Prior & MacWhinney, 2010); however, this finding is far from consistent (e.g. Paap et al., 2017). How to explain the incongruous results in the literature? First, it is of primary importance to view bilingualism as a set of complex interacting variables (e.g. age of acquisition, dominance, proficiency, language entropy, typological distance between languages, etc.), rather than an abstract, monolithic entity. The current paper focuses on one such

complex variable, notably bilinguals' switching habits: perhaps bilinguals who have more "training" in mental set shifting via code-switching are the ones who display increased cognitive flexibility.

### 1.2 Code-switching and cognitive flexibility

Within the BAH controversy, usage or performance-based approaches have been picking up speed (Korenar et al., 2023a). Certain linguistic dimensions–notably, language mixing, entropy, and proficiency–are most central to the bilingual experience (Kałamała et al., 2022), and are thus good candidates for explaining individual differences in bilingual executive functioning.

The recent and particular interest in code-switching likely derive from several important psycholinguistic discoveries. First, though code-switching is commonplace among many stable bilingual communities, code-switching often engenders comprehension or processing costs (e.g. increased reaction times [RTs]: Faroqi-Shah & Wereley, 2022; electrophysiological indices: Gosselin & Sabourin, 2021; increased brain activation: Rossi et al., 2021). Second, language switching appears to recruit the same regions of the brain as those involved in domain-general mental set shifting (e.g. De Baene et al., 2015; Garbin et al., 2010). Paired together, these findings suggest that code-switchers are skilled at performing a cognitively difficult act; if this behaviour trains its underlying neural substrates (e.g. Korenar et al., 2023b), it may eventually lead to better performance in general EFs.

Several behavioural studies have examined the relationship between code-switching and cognitive flexibility, and their findings remain mixed. For instance, Barbu and colleagues (2018, 2020) showed that Belgian young adults who code-switch more frequently exhibit overall faster RTs within cognitive flexibility paradigms (see also Lai & O'Brien, 2020; López-Penadés et al.,

2020), but these global effects are not observed among habitual code-switchers in California (Prior & Gollan, 2011). Taken as a whole, research suggests that code-switching may improve *language-specific* shifting (Jylkkä et al., 2020; Lai & O'Brien, 2020; Yim & Bialystok, 2012; Yow & Li, 2015; cf. Jylkkä et al., 2017, 2021). It rarely improves domain-general *sustained* cognitive flexibility (i.e. mixing effect: Paap et al., 2017, 2021; Prior & Gollan, 2011; Yim & Bialystok, 2012; cf. Lai & O'Brien, 2020), but is more commonly (positively) related to domain-general *transient* cognitive flexibility (i.e. switching effect: Prior & Gollan, 2011; Sanchez-Azanza et al., 2020; see also Paap et al., 2017 for a marginal interaction between switch costs and code-switching habits; cf. Lai & O'Brien, 2020; Paap et al., 2021; Yim & Bialystok, 2012).

Though bilinguals are occasionally categorised into discrete groups based on their code-switching frequency (e.g. high-frequency vs. low-frequency code-switchers; Barbu et al., 2018, 2020; Prior & Gollan, 2011), code-switching is typically examined by using continuous indices based on self-reports (though see Lai & O'Brien, 2020; Yim & Bialystok, 2012 for objective measures). One prevalent code-switching questionnaire is the Bilingual Switching Questionnaire (BSwQ; Rodriguez-Fornells et al., 2012), which separates switching behaviours into four subcomponents: switching into the L1, switching into the L2, contextual code-switching (i.e. switching motived by external circumstances) and unintentional switching (i.e. switching unmotivated by external circumstances, lack of switching awareness). Among young adults, it appears that contextual switching or intentional switching tends to be positively linked to cognitive flexibility (Jylkkä et al., 2020; López-Penadés et al., 2020; Sanchez-Azanza et al., 2020; see also Hartanto & Yang, 2016, 2020; Lai & O'Brien, 2020 for inter-sentential switches), and that unintentional switching is negatively associated (Yeh & Lim, 2021; see also Hartanto &

Yang, 2016, 2020; Jylkkä et al., 2021 for intra-sentential or within-speaker switches) though null results are also observed (Jylkkä et al., 2017, experiment 1; Soveri et al., 2011).

These findings seem to suggest that code-switching does not give rise to uniform effects when it comes to bilingual cognitive flexibility. Code-switching may favour mental set shifting if it is externally motivated, but it probably does not if it is internally motivated (see Beatty-Martínez, Navarro-Torres, & Dussias, 2020a for a discussion) or the result of language control failure (Festman et al., 2010). Importantly, these results show that bilinguals' code-switching frequency must be extricated from the circumstances surrounding their code-switching, or their INTERACTIONAL CONTEXT.

### 1.3 Bilingual interactional contexts and cognitive flexibility

According to the Adaptive Control Hypothesis (ACH; Green & Abutalebi, 2013), not all bilingual interactional contexts necessitate the same cognitive demands. Three contexts are defined: in a single-language environment, languages are context-isolated (e.g. speaking French at home and English at school). In a dual-language environment, multiple languages can be used in a given context, but switching is dictated by the requirements of that context (e.g. my co-worker is francophone, but my work documents are in English). Finally, in a dense code-switching context, the use of languages (and switching) is opportunistic since interlocutors can assume common linguistic backgrounds or knowledge with their interlocutors.

The ACH maintains that the dense code-switching context requires few control processes: bilinguals existing in this environment can use any language at any moment, based on ease of retrieval, communicative needs, and the expression of culture or identity (i.e. internally motivated switching; Gollan & Ferreira, 2009; Xu et al., 2021). On the other hand, bilinguals in a

dual-language context (and to a lesser degree, those in a single-language context) must monitor their surroundings, in turn activating the goal-appropriate language and inhibiting any non-pertinent languages (i.e. externally motivated switching). The dual-language context is thus expected to put into practice various control mechanisms, notably goal maintenance, and task engagement/disengagement.

An increasing amount of evidence supports the ACH (though see Paap et al., 2021). Some studies simulate different interactional contexts within their experimental design (e.g. Hofweber et al., 2020a; Jevtović et al., 2020; Zhu et al., 2022). For instance, Han and colleagues (2022) induced contexts by having participants listen to unilingual passages (single-language context), alternating Chinese-English passages (dual-language context), or heavily code-switched passages (dense code-switching context). In a subsequent Flanker task, participants from single- and dual-language conditions displayed faster RTs than those from the dense code-switching condition. Other studies tested groups from regions with more or less language integration, and indeed showed that bilinguals in areas wherein language separation is the norm display some form of increased cognitive control compared to bilinguals in regions bolstering language diversity (e.g. Freeman et al., 2022; see also Hartanto & Yang, 2016; Ooi et al., 2018; Yeh & Lim, 2021). For instance, Beatty-Martínez and colleagues (2020b) found that bilinguals from a primarily dual-language context exhibited better proactive control (i.e. reliance on context) compared to bilinguals in a more integrated linguistic context. More dual-language-like code-switching (i.e. alternation) has also been linked to increased reactive control (Hofweber et al., 2020b; Jiang et al., 2023). Once again, since these mechanisms are presumably implicated in mental set shifting (i.e. sustained cognitive flexibility = proactive control; transient cognitive flexibility = reactive

control), it is possible that dual-language bilinguals exhibit increased cognitive flexibility abilities relative to single-language and dense code-switching counterparts.

Unfortunately, detailed information about participants' sociolinguistic context is often absent, even in bilingualism research (Surrain & Luk, 2019). For this reason, it is not obvious whether the ACH can account for the mixed findings concerning the relationship between code-switching and cognitive flexibility. Many studies examine switching frequency (e.g. self-reported code-switching habits, objective switching frequency) without speculating on the participants' interactional context. Novel experiments utilise continuous indices tapping into interactional contexts (e.g. entropy, balanced usage: Hartanto & Yang, 2020; Li et al., 2021; van den Berg et al., 2022), but do not necessarily provide a baseline for the sociolinguistic context. The baseline is crucial. In a primarily single-language community, increased by-subject entropy may reflect more dual-language-like interactions for that individual (e.g. hypothesised *positive* relationship with cognitive flexibility). By contrast, in a community bolstering language diversity and integration, *decreased* by-subject entropy may reflect more dual-language-like interactions for that individual (e.g. hypothesised *negative* relationship with cognitive flexibility).

Altogether, it appears that a bilingual's interactional context may blur the relationship between their code-switching habits and cognitive flexibility abilities. We attempt to minimise this confound by testing bilinguals living in a stable bilingual community, and in particular, one which promotes dual-language interactions.

### 1.4 Present study

The present study examines whether French-English code-switching in a dual-language context (primarily Ontario, Canada) may lead to increased cognitive flexibility. The interactional

context of the sample was established in our previous work by examining the sociolinguistic, historical, and political background of the sample (Gosselin & Sabourin, 2023, pp. 5–6).

To assess domain-general cognitive flexibility abilities, we administered a colour-shape shifting task. Language-specific abilities were measured using a bilingual lexical decision task. These tasks were designed to assess both sustained cognitive flexibility (mixing effects) and transient cognitive flexibility (switching effects). Hence, we examined four constructs: (1) domain-general sustained cognitive flexibility; (2) domain-general transient cognitive flexibility; (3) language-specific sustained cognitive flexibility; and (4) language-specific transient cognitive flexibility. To our knowledge, we are the first to use a language-cued lexical decision task to test language-specific cognitive flexibility. We also utilise the prevalent BSwQ, under the assumption that L1-to-L2, L2-to-L1, and contextual switches are more typical in the dual-language context (positively linked to cognitive flexibility: Jylkkä et al., 2020; López-Penadés et al., 2020; Sanchez-Azanza et al., 2020), and that unintentional switches are more pervasive in a dense code-switching (negatively associated to cognitive flexibility: Yeh & Lim, 2021; see Gosselin & Sabourin, 2023 for a discussion).

We hypothesised that the cognitive flexibility abilities of our sample would be positively predicted (i.e. reduced mixing and switching effects) by their intentional code-switching habits, since the dual-language context puts into practice control processes implicated in both sustained (goal maintenance) and transient (task engagement/disengagement) cognitive flexibility. Furthermore, we anticipated that this would be the case for both domain-general and language-specific cognitive flexibility, since language switching appears to recruit the same regions of the brain as those involved in domain-general mental set shifting (e.g. De Baene et al., 2015).

## 2. METHODS

### 2.1 Participants

This study was approved by the Research Ethics Board at the University of Ottawa (file #S-01-21-6539). The participants are the same as those reported in Gosselin and Sabourin (2023), who present the findings relating to the sample's inhibitory control abilities. Here, we analyze the data obtained from the two cognitive flexibility tasks (a colour-shape shifting task, and a bilingual lexical decision task).

In short, 266 individuals gave informed consent and completed the study; 41 participants were subsequently excluded as they did not fit the inclusion criteria (e.g. low proficiency in English or French, advanced proficiency in a third language, colour-blindness, etc.). This resulted in a final sample of 225 French-English bilinguals (*M* age = 21.2 years [NOTE 1], *SD* = 7.0 years, *range* = 16-50 years; 190 women, 35 men). In addition to carrying out the experimental tasks, the participants also completed an exhaustive demographic and language background questionnaire (Sabourin et al., 2016), as well as the Bilingual Switching Questionnaire (BSwQ: Rodriguez-Fornells et al., 2012). These sample characteristics are summarised in Table 1. The participants tended to have acquired French before English, currently reported using English more frequently than French, and maintained high proficiency in both languages. A range of code-switching habits were observed, but code-switching was generally quite frequent. For an in-depth description of the participant sample, as well as a thorough explanation of the sociolinguistic context of the sample (i.e. a dual-language environment), see Gosselin and Sabourin (2023, pp. 5–6).

***Table 1*. Overview of the participant sample.**

| | | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| **Highest level of education obtained*** | | 36.9% (83) | 42.2% (95) | 18.2% (41) | 1.8% (4) | 0.9% (2) |
| **Video-game habits**** | | 36.4% (82) | 48.0% (108) | 3.6% (8) | 9.8% (22) | 2.2% (5) |
| | | | | **Ontario** | **Québec** | **Other** |
| **Current province of residence** | | | | 82.2% (185) | 15.5% (35) | 1.8% (4) |
| | | | | **Only French** | **Only English** | **Both** |
| **Self-reported L1** | | | | 54.2% (122) | 19.6% (44) | 26.2% (59) |
| | | | | **Mean** | **SD** | **Range** |
| **French** | ***Age of exposure (years)*** | | | 0.76 | 1.65 | 0 – 7 |
| | ***Current proficiency (out of 5)†*** | | | 4.59 | 0.61 | 2 – 5 |
| | ***Current usage (%)*** | | | 41.8 | 26.1 | 0 – 100 |
| **English** | ***Age of exposure (years)*** | | | 2.14 | 2.87 | 0 – 10 |
| | ***Current proficiency (out of 5)†*** | | | 4.57 | 0.58 | 2 – 5 |
| | ***Current usage (%)*** | | | 57.9 | 26.0 | 0 – 100 |
| **BSwQ‡** | ***EN-to-FR switching (out of 5)*** | | | 2.95 | .73 | 1 – 5 |
| | ***FR-to-EN switching (out of 5)*** | | | 3.18 | .74 | 1 – 5 |
| | ***Contextual switching (out of 5)*** | | | 3.10 | .92 | 1 – 5 |
| | ***Unintended switching (out of 5)*** | | | 2.30 | .89 | 1 – 5 |
| **Bilingual load** | ***Computed dominance (-4 – 4)+*** | | | .35 | 2.77 | -4 – 4 |
| | ***Percent bilingual interactions*** | | | 33.5 | 23.6 | 0 – 94.4 |
| | ***Language entropy statistic*** | | | .76 | .24 | 0 – 1 |

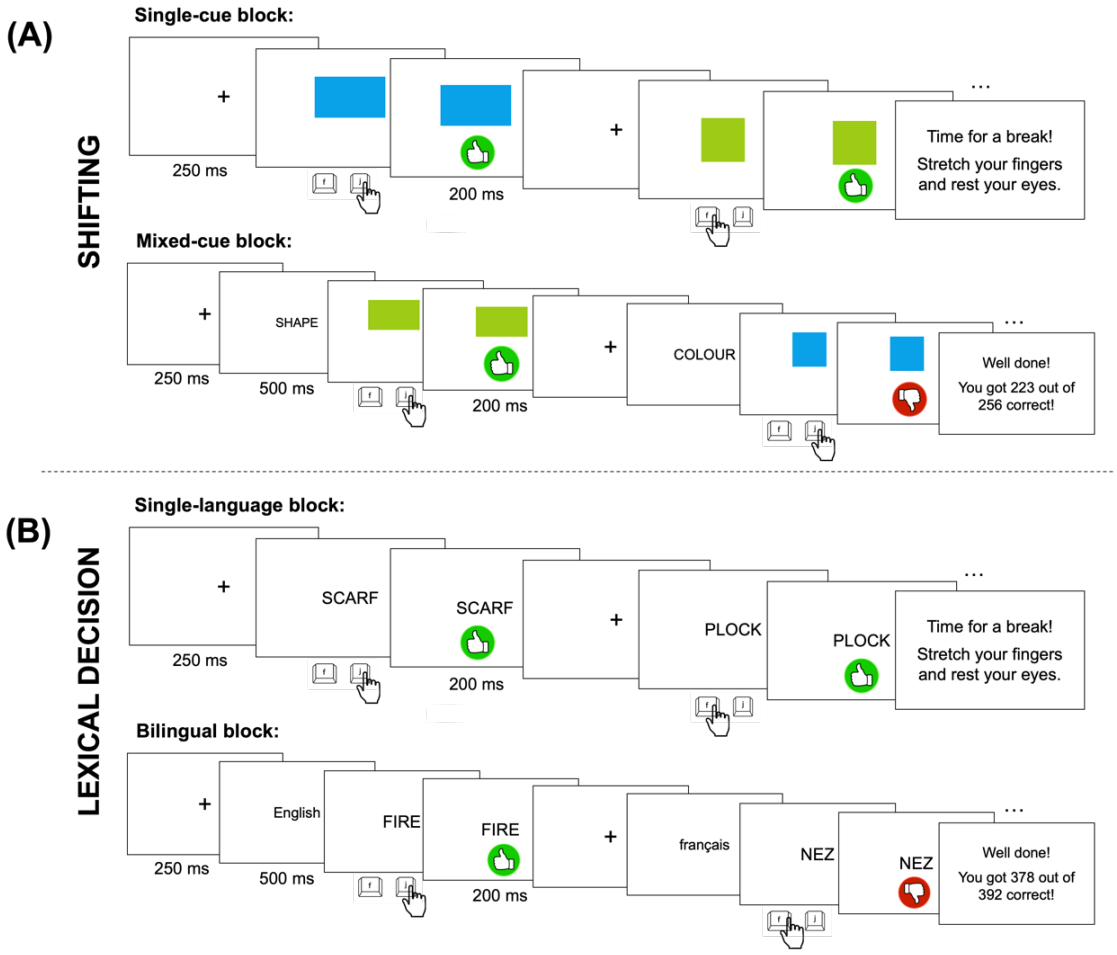
*** Education**: 1=high school; 2=1-2 years post-secondary; 3=3-4 years post-secondary; 4=master's; 5=doctorate.
**** Videogames:** 1=never; 2=once in a while; 3=at least once a week; 4=almost every day; 5=every day.
**†** **Proficiency**: 0=very low; 1=low; 2=intermediate; 3=advanced; 4=near-native; 5=native.
**‡** **BSwQ components**: 1=never; 2=very infrequent; 3=occasionally; 4=frequent; 5=always.
**+ Dominance:** -4 = very English dominant, 0 = balanced dominance, 4 = very French dominant

## 2.2 Procedure and task design

The full study (available for preview and cloning at https://app.gorilla.sc/openmaterials/429412) was administered online, through Gorilla Experiment Builder (Anwyl-Irvine et al., 2020). After giving consent in the language of their choosing (English or French), participants performed four tasks: the two cognitive flexibility

tasks reported in the present study, as well as two inhibitory control tasks (Flanker and bilingual Stroop) reported in Gosselin and Sabourin (2023). There was no set time limit to complete the study; participants could advance through the tasks at their own pace, with the possibility of pausing between each task. The participants ended the study by filling out (in English or in French) the questionnaires (see Section 2.1).

***Figure 1.*** Experimental flow for the shifting task (A) and the lexical decision task (B).




### 2.2.1 Colour-shape shifting task

Feature-shifting tasks (Allport et al., 1994) are often used to assess domain-general cognitive flexibility. We adopted a common variant for feature shifting, wherein participants

were required to indicate either the shape (rectangle, square) or the colour (blue, green) of a geometric figure, according to a preceding cue.

The colour-shape shifting task in the current study was adapted from a sample of this paradigm on Gorilla Experiment Builder (https://app.gorilla.sc/admin/task/17050/editor). It was composed of three blocks. The first two blocks were single-cue blocks. In the "shape" block, there were 32 rectangle trials and 32 square trials (each split evenly in colour). In the "colour" block, there were 32 blue trials and 32 green trials (each split evenly in shape). The order of the single-cue blocks was counterbalanced across participants. The final fixed block was a mixed-cue section: it was composed of 64 repeated-cue trials (half colour-to-colour, half shape-to-shape) and 64 switched-cue trials (half colour-to-shape, half shape-to-colour).

There were five practice trials before each single-cue block and ten practice trials before the mixed-cue block. Participants were given the option to take a break after every block. An additional dummy trial appeared at the beginning of each block; these were not included in the analyses. In total, each participant saw 256 experimental targets (half in the single-cue blocks and half in the mixed-cue block). These trials were pseudo-randomized within blocks into two different orders, which were then counterbalanced across participants.

A single trial from each of the single-cue and mixed cue blocks is depicted in Figure 1. A fixation cross was first presented for 250 ms. For single-cue trials, the target figure appeared immediately after the fixation (on a white screen). For mixed-cue trials, the feature cue (black font in uppercase letters: SHAPE/COLOUR) was presented for 500 ms before the target figure appeared. Participants were instructed to press "f" for a blue or square target, and "j" for a green or rectangular target. An image of a keyboard depicting these responses remained on the screen during the practice trials; this mnemonic disappeared during the experimental trials. The task

flow did not proceed until a keyboard response was selected. Participants received 200 ms feedback after each individual trial. To keep the experiment interactive, participants saw their "score" (proportion of correct responses) at the end of the task. The task took approximately 8–12 minutes to complete.

### 2.2.2 Lexical decision task

To our knowledge, lexical decision tasks (Meyer & Schvaneveldt, 1971) have not been used in executive functioning studies. In our version of this paradigm, participants are required to indicate whether a string is a real word or a non-word, according to a preceding language cue (i.e. French or English). Using this design, the lexical decision paradigm can measure language-specific cognitive flexibility in much of the same way as classical shifting tasks: shifting takes place between languages instead of between features. We opted to utilise a bilinguals lexical decision task, as we wanted to tap into linguistic cognitive flexibility without introducing an additional confound of response modality. Most previous linguistic switching and mixing measures rely on production (e.g. semantic fluency, picture naming: Jylkkä et al., 2021; Lai & O'Brien, 2020; Prior & Gollan, 2011, 2013; Yim & Bialystok, 2012); the novel bilingual lexical decision task has the advantage of reflecting processing unhindered by production.

In the current study, the materials for this task consisted of 50 non-cognate French-English translations pairs, 50 English pseudo-words (e.g. *drack, wetch, chortle, harets, jayl*... ) and 50 French pseudo-words (e.g. *rique, larpe, tousue, malaud, stin*... ). The real English words (and their French translations), as well as the English pseudo-words, were taken from Sabourin and colleagues (2014). The French pseudo-words were selected from the French Lexicon Project (see Ferrand et al., 2010). There were no differences in frequency between the real-words in English

and in French ($t(98) = -0.613$, $p = 0.541$). A full list of this stimuli set including frequency information is available in the Open Science Framework (OSF) repository.

The lexical decision task was composed of three blocks. The first two blocks were single-language blocks. In each of the unilingual French and unilingual English blocks, there were 50 real-words trials and 50 non-words. The order of the single-language blocks was counterbalanced across participants. The final fixed block was bilingual: it was composed of the stimuli from both the French block and the English block. As such, there were 98 repeated-cue trials (half French-to-French, half English-to-English, counterbalanced for real word/nonword status) and 98 switched-cue trials (half French-to-English, half English-to-French, counterbalanced for real word/nonword status).

There were five practice trials before each single-cue block and ten practice trials before the mixed-cue block. Participants were given the option to take a break after each block and half-way through the bilingual block. The first trial at the beginning of each block (and following the break during the mixed block) was a dummy trial; these were not included in the analyses. In total, each participant saw 392 experimental trials (half in the single-cue blocks and half in the mixed-cue block). These trials were pseudo-randomized within blocks into two different orders, which were then counterbalanced across participants. Half the participants answered that a target was a "real word" by using their left hand ("f" key) and a "non-word" by using their right hand ("j" key); this configuration was reversed for the other half of participants.

A single trial from each of the single-language and bilingual blocks is depicted in Figure 1. A fixation cross was first presented for 250 ms. For single-cue trials, the target word (uppercase black font over a white screen) appeared immediately after the fixation. For mixed-cue trials, the language cue (black font in lowercase letters: English/francais)waspresentedfor500msbeforethe

target word appeared. The task was participant-controlled: it did not proceed until a keyboard response was selected. Finally, participants received 200 ms feedback after each individual trial. To keep the experiment interactive, participants saw their "score" (proportion of correct responses) at the end of the task. The task took approximately 10–15 minutes to complete.

### 2.3 Data analysis

#### 2.3.1 Data cleaning

Incorrect trials were not included in the analysis. This resulted in the exclusion of 3392 data points for the shifting task (5.9% of trials) and 5364 data points for the lexical decision task (6.1% of trials). Data cleaning was first implemented within participants, at the trial level (see Chan et al., 2020; Gosselin & Sabourin, 2023; Gullifer et al., 2013; Hartanto & Yang, 2016, 2020; Hofweber et al., 2020a; Jiao et al., 2019, 2020; Jylkkä et al., 2017; Kałamała et al., 2020; López-Penadés et al., 2020; Ooi et al., 2018; Prior & Gollan, 2011; Treffers-Daller et al., 2020 for comparable methods). Absolute minimum (shifting: 150 ms; lexical: 300 ms) and maximum (shifting: 1500 ms; lexical: 2000 ms) thresholds resulted in 1774 trimmed data points for the shifting task (3.1% of trials) and 2517 trimmed data points for the lexical decision task (2.9% of trials). Relative thresholds (± 3SD of each participant's overall mean) were subsequently computed and resulted in the rejection of 1003 data points for the shifting task (1.7% of trials) and 1569 data points for the lexical decision task (1.8% of trials). Participants who did not retain at least 80% of trials after cleaning (shifting: $n = 16$; lexical decision: $n = 16$) were excluded from the final analyses. Last, between-participant cleaning was also implemented (see Jylkkä et al., 2017; Lai & O'Brien, 2020; Yow et al., 2017 for comparable methods). The average mixing effect and switching effect of the sample was calculated for each task (within each language for

the lexical decision task). Participants who displayed effects inferior or superior to three standard deviations of the sample mean effect (shifting: $n = 2$; lexical decision: $n = 8$) were excluded from the analyses. Thus, the shifting task includes a final sample of 207 participants, while the lexical decision includes a final sample of 201 participants.

### 2.3.2 Condition differences

In cognitive flexibility tasks, single-cue blocks are predicted to produce more rapid RTs than trials from a mixed-cue block. Furthermore, in mixed-cue blocks, repetition trials (e.g. shape-to-shape; English-to-English) are expected to generate faster RTs than switched trials (e.g. shape-to-colour; English-to-French). To verify these differences, we submitted the shifting data and the lexical decision data to generalised linear mixed-effect models (using `lme4` packages in R 4.21). The data was fitted using a Gamma distribution and an identity link, one of the suggested best practices for "mental chronometry" data (Lo & Andrews, 2015). This configuration allows us to avoid the RT transformations required for linear mixed models, thereby facilitating model interpretation.

For both tasks, the participants' cleaned RTs was the dependent variable. Fixed effects of CONDITION (Single-cue [NOTE 2], Repetition-cue or Switched-cue) and, for the lexical decision task, LANGUAGE OF THE TRIAL (real-word English or real-word French) were included. By-participant and by-item random intercepts and slopes were first entered into maximal models; due to convergence issues, the random effect structure was simplified until only the by-subject and by-item intercepts remained. Post-hoc comparisons were done via estimated marginal means (`emmeans` package) and were corrected using the Holm adjustment.

### 2.3.3 Operationalising cognitive flexibility

We operationalise the participants' cognitive flexibility abilities by using the two standard metrics: domain-general and language-specific MIXING EFFECTS, as well as domain-general and language-specific SWITCHING EFFECTS. For the lexical decision task, only real-word trials were aggregated to calculate effects. Participants who display reduced (less negative) mixing effects are said to possess better sustained cognitive flexibility, as they can regulate antagonistic task sets over a considerable length of time (e.g. Braver et al., 2003; Rubin & Meiran, 2005; Soveri et al., 2011). Smaller (i.e. less negative) switching effects reflect increased transient cognitive flexibility abilities (i.e. less rigidness or perseveration of the set at a given moment; Braver et al., 2003; Diamond, 2013).

### 2.3.4 Multiple regressions

Similar to much of the previous past research (see Freeman et al., 2022; Gosselin & Sabourin, 2023; Hofweber et al., 2019, 2020a, 2020b; López-Penadés et al., 2020; Navarro et al., 2022; Paap et al., 2017, 2019), we conducted stepwise multiple linear regressions in order to examine whether the participants' cognitive flexibility abilities were predicted by their code-switching habits, and/or their other linguistic and demographic characteristics. Separate regressions were computed according to the task and the effect. As such, there were six dependent variables: three regressions (one for the shifting task, and two for lexical decision task) included the MIXING EFFECT (single-cue block RTs – mixed cue block RTs) and three regressions included the SWITCHING EFFECT (repeated-cue trial RTs – switched-cue trial RTs). In each case, the four BSwQ components (English-to-French switching; French-to-English switching; contextual switching; unintentional switching) were entered as independent factors.

Six additional predictors were also included, as they have been shown (or are theoretically believed) to impact cognitive flexibility: age (Kray & Lindenberger, 2000), education level (Guerra-Carrillo et al., 2017), experience with video games (Colzato et al., 2010), amount of bilingual interactions (Hartanto & Yang, 2016), language entropy (Hartanto & Yang, 2016, 2020), and language dominance [NOTE 3] (Jylkkä et al., 2018). Variables were entered into the null model in stepwise fashion (using the ols_step_both_p function in the `olsrr` package) according to an inclusion criterion of $p < 0.05$ and a removal criterion of $p > 0.10$. We report the final regression equation for each dependent variable.

Note that a visual inspection of standardised residual plots suggests that assumptions for linear regressions were satisfied (maximum standard residual ≤ |3.5|). Casewise diagnostics confirmed that no data point held undue influence in the models (Cook's distance < 0.17). Durbin-Watson statistics fell within the acceptable range (Durbin-Watson = 1.7–2.1) according to the number of regressors and the sample size (Savin & White, 1977), indicating that the independence of errors was satisfied. Furthermore, a correlation matrix did not flag multicollinearity issues between the predictors (all *r*s < 0.70; VIF = 1.05–2.15; Plonsky & Ghanbar, 2018).

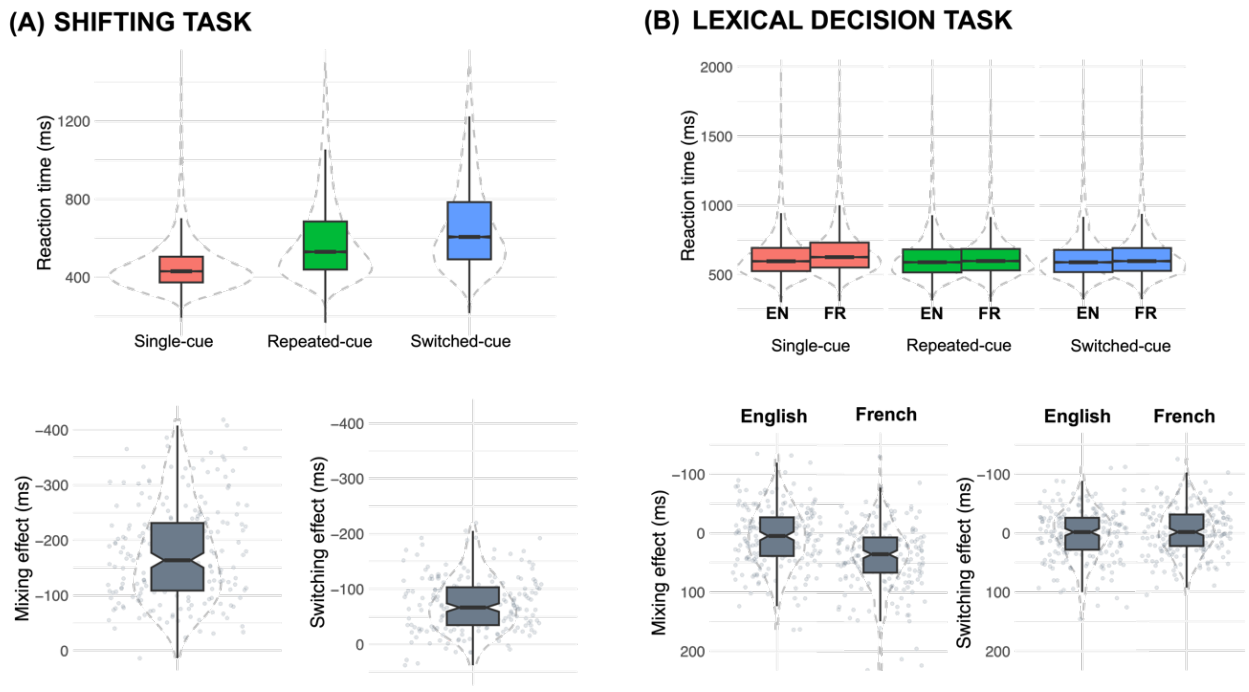
## 3. RESULTS

A full descriptive report and visualisation of the conditions and effects can be seen in Table 2 and Figure 2. The scripts used to clean the data, and to compute generalised linear mixed models and stepwise regressions are available in the OSF repository.

***Table 2.*** Descriptive statistics for the shifting and lexical decision tasks (also separated according to the language of the trial). Means are shown with standard deviations in parentheses.

| | **Single-cue** | **Repeated-cue** | **Switched-cue** | **Mixing** | **Switching** |
|---|---|---|---|---|---|
| **Shifting task** | 456.5 (61.2) | 595.2 (110.8) | 667.3 (126.9) | -174.8 (88.2) | -72.1 (49.4) |
| **Lexical decision** | 649.4 (76.7) | 626.0 (73.9) | 626.9 (72.9) | 23.0 (44.5) | -0.9 (28.1) |
| ***English*** | 631.4 (76.3) | 625.8 (83.0) | 623.5 (79.4) | 6.8 (51.5) | 2.3 (43.4) |
| ***French*** | 667.4 (90.7) | 626.2 (73.5) | 630.2 (80.6) | 39.2 (54.1) | -4.1 (39.7) |

***Figure 2***. *Top*: Reaction times according to CONDITION for the shifting task (A) and lexical decision task (B). *Bottom:* Variation in the Mixing effects (Single-cue block – Mixed-cue block) and Switching effects (Repeated-cue trials – Switched-cue trials) for the shifting task (A) and lexical decision task (B). For each condition/effect, boxplots show the median (thick black line), the inter-quartile range (outline of the box), the maximum and minimum values (whiskers); the dashed violin outline shows the density of the data.



### 3.1 Shifting task

#### 3.1.1 Mixed models

The model for the shifting task yielded expected CONDITION effects (see Table 3). Repeated-cue ($\beta$ = 127.1, $SE$ = 0.97, $p < 0.001$) and Switched-cue trials ($\beta$ = 195.2, $SE$ = 0.78, $p < 0.001$) elicited longer RTs than single-cue trials. Pairwise contrasts also revealed that Repeated-cue trials were significantly faster than Switched-cue trials ($\beta$ = −68.1, $SE$ = 1.2, $p < 0.001$). As such, sizable mixing effects and switching costs were indeed observed (Table 3).

#### 3.1.2 Stepwise regressions

By-participant mixing and switching effects from the shifting task were submitted to stepwise multiple linear regressions (see Table 4). For the mixing effect, only the participants' education level was retained in the final model: each unit-change in the predictor of education resulted in an increased (i.e. became more negative) mixing effect by a magnitude of – 7.5 ms. For the switching effect, the participants' degree of unintentional switching was retained in the final model: as the bilinguals reported more unintended code-switching, the switching effect increased (i.e. became more negative) by a magnitude – 6.9 ms. Note that although the retained predictors withstood the elimination criterion (p<0.10), their influence remained marginal in each respective model (p > 0.05).

***Table 3***. Results from the trial-by-trial mixed models for the shifting task. • p<.10, * $p$<0.05, ** $p$<.01, *** $p$<.001.

| | **Estimate ($\beta$)** | ***SE*** | **95% *CI*** | ***t*-value** |
|---|---|---|---|---|
| **Intercept (Single-cue)** | 471.90 | .88 | 470.2 – 473.6 | 535.4*** |
| **Repetition-cue** | 127.09 | .97 | 125.2 – 129.0 | 131.6*** |
| **Switched-cue** | 195.17 | .78 | 193.6 – 196.7 | 248.7*** |

**Model:** glmer(RT ~ Condition + (1|subject)+(1|item), df, family = Gamma(link = "identity")).

***Table 4***. Summary of the stepwise regression analyses for the shifting task. • p<.10, * *p*<0.05, ** *p*<.01, *** *p*<.001.

| **A. MIXING EFFECT** | | | | | |
|---|---|---|---|---|---|
| **Predictor** | ***R* ($R^2$)** | ***F*** | ***β* (95% CI)** | ***SE*** | ***t*** |
| **Step 1** | .123 (.015) | 3.125• | | | |
| Education level | | | -7.488 (-15.84 – -.86) | 4.236 | -1.768• |
| **B. SWITCHING EFFECT** | | | | | |
| **Predictor** | ***R* ($R^2$)** | ***F*** | ***β* (95% CI)** | ***SE*** | ***t*** |
| **Step 1** | .125 (.016) | 3.255• | | | |
| Unintended switching | | | -6.915 (-14.47 – .64) | 3.832 | -1.804• |

**Predictors**: (i) age, (ii) education level, (iii) video-game experience, (iv) bilingual interactions, (v) language entropy statistic, (vi) language dominance, (vii) French-to-English switching, (viii) English-to-French switching, (ix) contextual switching, (x) unintentional switching.

### 3.2 Lexical decision task

#### 3.2.1 Mixed models

The model for the lexical decision task is included in Table 5. An `anova` summary of the model for the lexical decision task revealed differences according to the variable of CONDITION ($F = 114.23$, $p < 0.001$), LANGUAGE of the trial ($F = 20.94$, $p < 0.001$), and the interaction between these terms ($F = 49.12$, $p < 0.001$). Pairwise contrasts revealed that these effects were not fully in the anticipated direction. First, Repeated-cue trials did indeed elicit shorter RTs than Switched-cue trials ($\beta = -7.99$, $SE = 1.60$, $p < 0.001$). However, single-cue trials were significantly slower than Repeated-cue trials ($\beta = 25.88$, $SE = 1.27$, $p < 0.001$) and Switched-cue trials ($\beta = 17.89$, $SE = 1.29$, $p < 0.001$). Further comparisons showed that this unexpected effect was modulated by the LANGUAGE of the trial, and that French trials generated longer RTs in general ($\beta = 16.01$, $SE = 1.72$, $p < 0.001$). In brief, English trials yielded an expected switching effect (Repeated-cue English < Switched-cue English: $\beta = -13.20$, $SE = 1.51$, $p < 0.001$), but no reliable mixing effect (Single-cue English = Switched-cue English: $\beta = 1.23$, $SE = 1.26$, $p = 1$).

In French, no switching effect was observed (Repeated-cue English = Switched-cue English: $\beta = -2.78$, $SE = 2.09$, $p = 1$), and the mixing effect was counter to our prediction (i.e. mixing "benefit': Single-cue French > Switched-cue French: $\beta = 34.54$, $SE = 1.61$, $p < 0.001$).

***Table 5***. Results from the trial-by-trial mixed models for the lexical decision task. • p<.10, * $p$<0.5, ** $p$<.01, *** $p$<.001.

| | **Estimate ($\beta$)** | ***SE*** | **95% *CI*** | ***t*-value** |
|---|---|---|---|---|
| **Intercept (Single-cue English)** | 647.07 | 1.33 | 644.18 – 649.96 | 438.69*** |
| **Repetition-cue** | -14.43 | 1.15 | -15.24 – -12.62 | -15.56*** |
| **Switched-cue** | -1.23 | 1.26 | -3.60 – 1.14 | .328 |
| **Language of the trial** | 34.76 | 1.55 | 32.47 – 37.04 | 22.41*** |
| **Repetition-cue (French)** | -22.89 | 1.40 | -25.08 – -20.70 | -16.30*** |
| **Switched-cue (French)** | -33.31 | 1.31 | -35.26 – -31.36 | -25.50*** |

**Model:** glmer(RT ~ Condition * Language + (1|subject)+(1|item), df, family = Gamma(link = "identity")).

### 3.2.2 Stepwise regressions

By-participant mixing and switching effects from the lexical decision task were submitted to stepwise multiple linear regressions (see Table 6 and Figure 3). Since the mixed model revealed differences according to the language of the trial (see also Gosselin & Sabourin, 2023), separate regressions were computed for English/French mixing and switching effects.

#### 3.2.2.1 Mixing effect

In English, three predictors were retained in the final model. First, the participants' current language entropy statistic predicted their mixing effect: more integrated language experience (higher entropy scores) resulted in larger (more negative) mixing effects. The participants' self-reported unintentional code-switching habits similarly affected the dependent variable: habitual unintentional switching was related to larger (more negative) mixing effects. By contrast, as the bilinguals reported more (deliberate) frequent French-to-English code-switching, the mixing

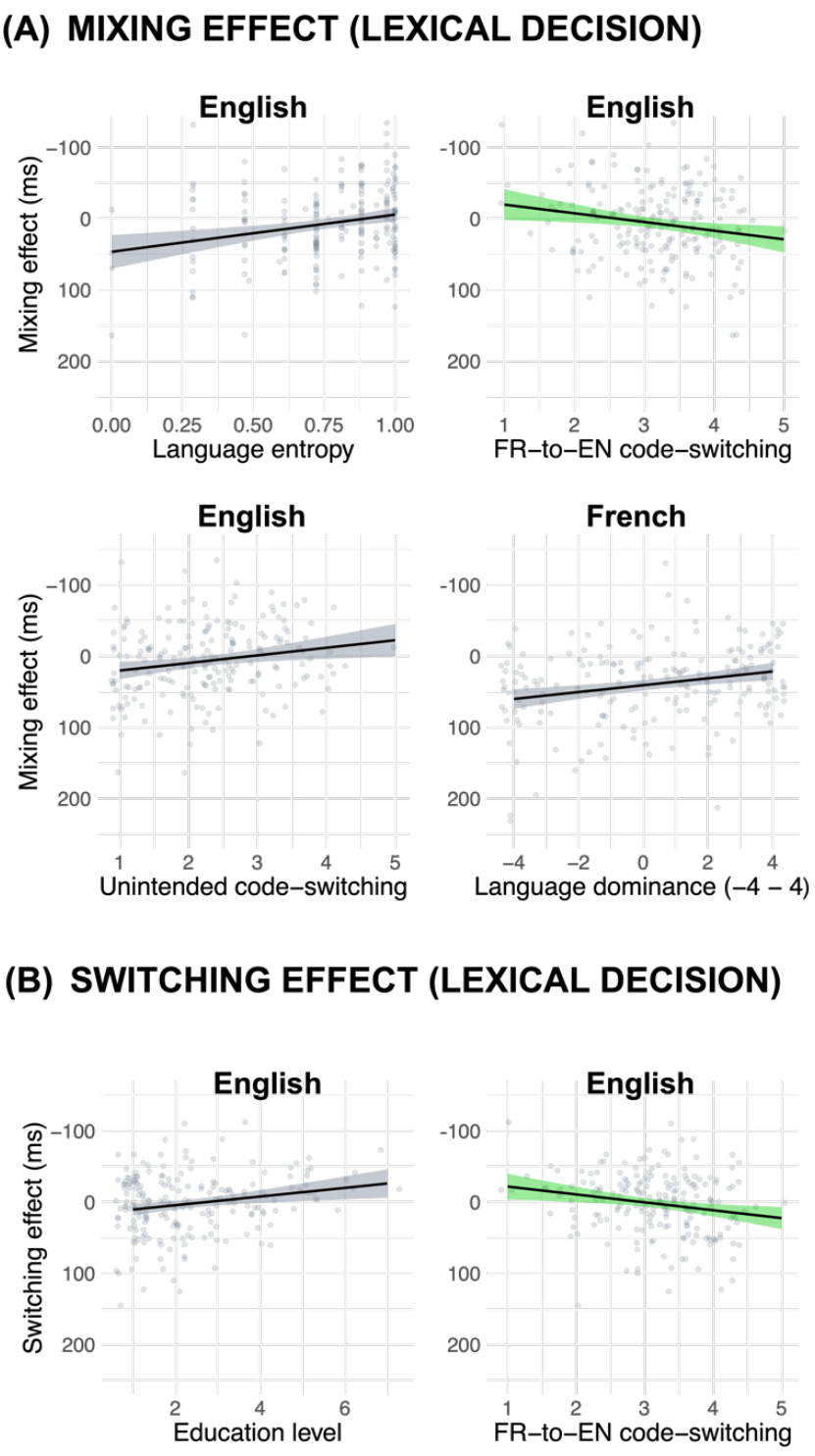
effect decreased (i.e. became less negative) by a magnitude of 14.2 ms per unit. In French, only language dominance was retained in the final model. Recall that language dominance is operationalised on a scale of – 4 to 4, wherein – 4 represents asymmetrically English-dominant bilinguals, 4 represents asymmetrically French-dominant bilinguals, and 0 represents balanced bilingualism. As such, mixing effects in French increased (i.e. became more negative) as dominance scores were skewed towards French.

***Table 6***. Summary of the stepwise regression analyses for the lexical decision task. • p<.10, * $p$<0.5, ** $p$<.01, *** $p$<.001.

| **A. MIXING EFFECT** | | | | | | |
|---|---|---|---|---|---|---|
| | **Predictor** | ***R* ($R^2$)** | ***F*** | ***β* (95% *CI*)** | ***SE*** | ***t*** |
| **English** | **Step 1** | .241 (.058) | 12.255*** | | | |
| | Entropy | | | -52.071 (-81.40 – -22.74) | 14.874 | -3.50*** |
| | **Step 2** | .288 (.083) | 8.977*** | | | |
| | Entropy | | | -48.942 (-78.07 – -19.81) | 14.773 | -3.31*** |
| | FR-to-EN | | | 10.786 (1.66 – 19.92) | 4.630 | 2.33* |
| | **Step 3** | .340 (.116) | 8.578*** | | | |
| | Entropy | | | -38.129 (-67.89 – -8.36) | 15.093 | -2.53* |
| | FR-to-EN | | | 14.160 (4.84 – 23.49) | 4.729 | 3.00** |
| | Unintended | | | -11.100 (-19.25 – -2.95) | 4.132 | -2.69** |
| **French** | **Step 1** | .245 (.060) | 12.701*** | | | |
| | Dominance | | | -4.825 (-7.49 – -2.16) | 1.35 | -3.56*** |
| **B. SWITCHING EFFECT** | | | | | | |
| | **Predictor** | ***R* ($R^2$)** | ***F*** | ***β* (95% *CI*)** | ***SE*** | ***t*** |
| **English** | **Step 1** | .203 (.041) | 8.581** | | | |
| | Education | | | -6.098 (-10.20 – -1.99) | 2.082 | -2.93** |
| | **Step 2** | .255 (.065) | 6.89** | | | |
| | Education | | | -5.063 (-9.23 – -.90) | 2.112 | -2.40* |
| | FR-to-EN | | | 8.996 (1.08 – 16.91) | 4.013 | 2.24* |
| **French** | *ns* | – | – | – | – | – |

**Predictors**: (i) age, (ii) education level, (iii) video-game experience, (iv) bilingual interactions, (v) language entropy statistic, (vi) language dominance, (vii) French-to-English switching, (viii) English-to-French switching, (ix) contextual switching, (x) unintentional switching.

***Figure 3***. Significant predictors for mixing effects (A) and switching effects (B) within the lexical decision task. The shaded areas bordering the regression line illustrate the standard error. Negative is plotted up. Predictors related positively to cognitive flexibility (i.e., smaller/less negative effects) are plotted in green.



#### 3.2.2.2 Switching effect

In English, the participants' educational level as well as their French-to-English code-switching habits were retained in the final model. First, higher education levels resulted in larger (i.e. more negative) switching effects. Second, as the bilinguals reported more (deliberate) frequent French-to-English code-switching, the switching effect decreased (i.e. became less negative) by a magnitude of 9.0 ms per unit. In French, no predictors were retained, and no model was produced.

### 3.3 Reliability, construct validity, and convergence

Trials were split into odd and even trials, and by-participant average mixing and switching effects were calculated for each of these subsets. Intraclass correlation coefficients (type ICC(2,k); Shrout & Fleiss, 1979) were then computed. The mixing effect was found to have excellent internal consistency in the shifting task (*ICC* =0.94) and good internal consistency in the lexical decision task (overall: *ICC* = 0.76; English only: *ICC* = 0.65; French only: *ICC* = 0.70). The switching effect was found to have poor internal consistency in the shifting task (*kICC*=0.28) and in the lexical decision task (overall: *ICC*s = 0). Note that individual conditions (i.e. odd vs. even single-cue trials; odd vs. even repeated-cue trials; odd vs. even switched-cue trials) displayed good or excellent internal consistency in all cases (*ICC*s > 0.80).

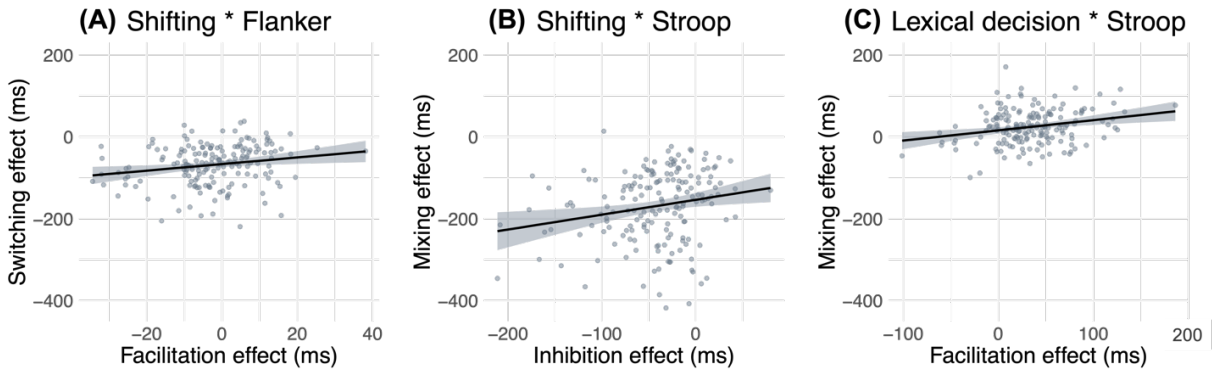
To examine the relation among operationalised cognitive flexibility constructs, Pearson correlations were computed between the mixing and switching effects in each task. In the shifting task, mixing and switching effects were significantly correlated ($r = 0.34$, $p < 0.001$). In the lexical decision task, mixing and switching effects were not correlated ($r$s<|0.03|, $p$s > 0.70). Convergence was also verified by comparing the mixing effect and switching effect across tasks. Neither the mixing effect nor the switching effect were correlated across tasks ($r$s<|0.03|, $p$s > 0.72).

Finally, recall that the sample in the present study also completed domain-general and language-specific inhibitory control tasks, as reported in Gosselin and Sabourin (2023). As inhibitory control is believed to play a part in cognitive flexibility (Diamond, 2013), we computed a correlation matrix between the executive functioning constructs among the participants whose data were retained for all four tasks ($n$=176). Three cross-task correlations were significant (see Table 7). First, domain-general transient cognitive flexibility was positively associated to domain-general monitoring ($r$ = 0.19, $p$ = 0.011; Figure 4A). Second, domain-general sustained cognitive flexibility was positively associated to language specific inhibition ($r$ = 0.20, $p$ = 0.007; Figure 4B). Finally, language-specific sustained cognitive flexibility was positively associated to language-specific monitoring ($r$ = 0.24, $p$ = 0.001; Figure 4C).

***Table 7***. Convergence between cognitive flexibility measures in the present paper and inhibitory control measures from Gosselin and Sabourin (2023). Effects are collapsed across languages for the lexical decision task. Two-tailed Pearson's correlations are reported. *$p$<0.5, **$p$<.01, ***$p$<.001.

| | | Shifting | | Lexical decision | |
|---|---|---|---|---|---|
| | | **mixing** | **switching** | **mixing** | **switching** |
| **Flanker** | **facilitation** | $r$=.101 | $r$=.192* **(A)** | $r$=.010 | $r$=.004 |
| | **inhibition** | $r$=-.032 | $r$=.039 | $r$=.103 | $r$=.130 |
| **Stroop** | **facilitation** | $r$=.034 | $r$=-.021 | $r$=.244*** **(C)** | $r$=.021 |
| | **inhibition** | $r$=.202** **(B)** | $r$=.075 | $r$=.097 | $r$=.127 |

***Figure 4.*** Convergence between cognitive flexibility measures in the present paper and inhibitory control measures from Gosselin and Sabourin (2023). The shaded areas bordering the regression line illustrate the standard error.



### 3.4 Summary

Expected switching and mixing costs were observed for the shifting task, but the magnitude of these effects was not significantly related to the examined predictors, including the critical code-switching variables. The lexical decision task yielded the expected (transient) switching cost, but only in English; the anticipated (sustained) mixing cost was absent in both languages. Nevertheless, variations in the language-specific mixing and switching effects were apparent across the sample. In English, sustained cognitive flexibility was positively predicted by the participants' propensity to switch into English, but negatively related to their linguistic entropy score and their unintentional switching; transient cognitive flexibility was also positively predicted by the participants' propensity to switch into English, but negatively related to their level of education. In French, the sustained cognitive flexibility was impacted by language dominance: more French-dominant bilinguals displayed larger switching effects. Finally, a cross-task correlation matrix showed that indices of cognitive flexibility in the present paper were significantly and positively related to inhibitory control indices from Gosselin and Sabourin (2023).

## 4. DISCUSSION

The current study examined whether code-switching habits influence the sustained and transient cognitive flexibility of bilinguals living in a dual-language environment. Both domain-general (via a colour-shape shifting task) and language-specific (via a bilingual lexical decision task) cognitive flexibility skills were assessed.

As hypothesised, the language-specific cognitive flexibility abilities of our sample were positively predicted by their intentional code-switching habits. That is, bilinguals who reported

frequent code-switching into English displayed better sustained (i.e. reduced mixing effects) and transient (i.e. reduced switching effects) cognitive flexibility, but only for English trials. Predictors related to dense code-switching (i.e. unintentional switching, high language entropy) negatively predicted the participants' language-specific cognitive flexibility skills. However, contrary to our hypotheses, the *domain-general* cognitive flexibility skills of our sample were not reliably associated to their code-switching habits. These main findings are discussed in the following sections.

### 4.1 Language-specific-skills: context matters, particularly for sustained cognitive flexibility

As mentioned, the French-English bilinguals tested in the present study originated from speech communities that stimulate dual-language interactions. This environment constrains individuals to "choose a language" based on the external requirements of a situation (e.g. my co-worker is francophone, but my work documents are in English). The dual-language context (as opposed to the dense code-switching context) entails purposeful switching; we thus expect that deliberate (intentional) code-switching is the default switching behaviour for our sample. Indeed, the participants reported intentional switching, particularly switching into the majority language (English), much more often than unintentional switching ($t(206) = 12.99$, $p < 0.001$), though a range of code-switching habits and entropy scores were reported.

Crucially, our analyses revealed that intentional switching (switching into English) and unintentional switching exhibited directionally-opposite relationships with language-specific cognitive flexibility. More deliberate switching into the majority language was associated to both superior sustained and transient cognitive flexibility in English. Similar results have been observed in previous language-specific tasks targeting sustained cognitive flexibility among

young adults (global RT positively linked to code-switching: Barbu et al., 2018, 2020; mixing effect positively linked to code-switching: Yow & Li, 2015; cf. Jylkkä et al., 2017, 2020, 2021), and–albeit less frequently–for transient cognitive flexibility (switching effect positively linked to code-switching: Jylkkä et al., 2020 for contextual switches; Sanchez-Azanza et al., 2020; cf. Jylkkä et al., 2017, 2021; Yow & Li, 2015). It is worth mentioning that the majority of null or negative results reported in the literature have been obtained from the same population, notably Finish-Swedish or Finish-English bilinguals in Finland, where code-switching is relatively infrequent (on average, 0 to 3 switches every 2 hours: Jylkkä et al., 2020).

Conversely, frequent unintentional switching was detrimental to sustained cognitive flexibility. That is, individuals who reported habitual unintended switching exhibited weaker cognitive flexibility abilities. The participants' language entropy statistic patterned with their unintentional switching habits, suggesting that increased language diversity does not necessarily lead to increased cognitive flexibility. We believe that deliberate switching is a proxy of the dual-language context, and unintentional switching is more usual in a dense code-switching context (e.g. dense code-switchers have more difficulty suppressing automatic responses: Hofweber et al., 2020b). As such, these predictors allow us to speculate about the importance of bilingual interactional contexts when it comes to cognitive flexibility. Overall, our results support the Adaptive Control Hypothesis (Green & Abutalebi, 2013); dual-language-like switching appears to train language-specific cognitive flexibility, but dense-like switching does not. This might also explain why entropy patterns with unintentional switching (cf. Li et al., 2021; van den Berg et al., 2022): higher language diversity is akin to a dense code-switching context, especially if the baseline is a dual-language environment (as is the case for our sample). Perhaps increased

entropy is negatively associated to sustained cognitive flexibility as it eventuates less dual-language-like (i.e. intentional) switching, and more dense (i.e. unintentional) switches.

The picture is complicated by the fact that unintentional code-switching was found to be detrimental for *sustained* cognitive flexibility (mixing effects) but bore no effect on *transient* cognitive flexibility (switching effects). Recall that sustained cognitive flexibility falls under the umbrella of proactive control, while transient cognitive flexibility is associated to reactive control (see Braver, 2012; Braver et al., 2003; Diamond, 2013; Prior & Gollan, 2013; Rubin & Meiran, 2005). Our cross-task convergence analyses support these postulations: the correlation between the language-specific mixing effect and the language-specific facilitation effect (Stroop task from Gosselin & Sabourin, 2023) was significant and positive, and Stroop facilitation effects have been found to be largely reliant on the degree of proactive control imposed by the task (e.g. Goldfarb & Henik, 2007; Kalanthroff et al., 2015).

A recent study utilising an AX Continuance Performance Task (AX-CPT) demonstrated that bilinguals from a primarily dual-language context exhibited better proactive control (i.e. reliance on context) compared to bilinguals in a more integrated linguistic context (i.e. dense code-switching). However, both groups relied on reactive control less than their single-language counterparts (Beatty-Martínez et al., 2020b). Our results fall into line with those of Beatty-Martínez and colleagues (2020b). We speculate that in a dense code-switching context (but not in a dual-language context), proactive control decreases since the internal maintenance of a given language (see Prior & Gollan, 2013) is no longer needed (but see Hofweber et al., 2020a, 2020b; Jiang et al., 2023). However, *reactive* control is likely required in both contexts, as individuals must be ready to process code-switched input (i.e. a co-interlocutor's switches). This would entail that tasks/measures primarily tapping into reactive control (as opposed to those targeting

proactive control) may not yield differences between dual-language and dense code-switching bilinguals.

### 4.2 Domain-general skills are not reliably associated to code-switching or context

We originally hypothesised that the training imparted by the bilinguals' dual-language switching would extend to their domain-general cognitive flexibility skills. In other words, we anticipated that increased intentional switching would predict better sustained and transient cognitive flexibility among our participants. Though robust domain-general mixing and switching costs were observed, these were not reliably associated to the examined predictors.

It is worth mentioning that switching costs trended towards being negatively related to unintentional switching habits; this would suggest that dense code-switching behaviours are detrimental to domain-general transient cognitive flexibility. Furthermore, our cross-task convergence analyses yielded a significant correlation between the domain-general switching effect and the domain-general facilitation effect (flanker task from Gosselin & Sabourin, 2023), once again hinting at the relationship between proactive/reactive control and interactional contexts. However, given that the main findings were at best marginal, we are reluctant to make general claims regarding these effects. Overall, we assert that our prediction was not supported: the training imparted by deliberate code-switching appeared to be restricted to the language-specific domain. Code-switchers are good at shifting linguistic sets, but this does not extend globally to mental set shifting.

Of the studies that have utilised domain-general cognitive flexibility tasks in the code-switching literature, the results remain mixed (Hartanto & Yang, 2016, 2020; Lai & O'Brien, 2020; López-Penadés et al., 2020; Paap et al., 2017, 2021; Prior & Gollan, 2011; Yim &

Bialystok, 2012). Altogether, there is no consistent effect of "code-switching training" when it comes to domain-general cognitive flexibility. Interestingly, however, code-switching behaviours are related far more often to domain-general transient cognitive flexibility (i.e. switching cost) than to domain-general sustained cognitive flexibility (i.e. mixing cost; see Section 1.2). This was also the case in the present study: unintentional switching was marginally related to the switching effect, but not to the mixing effect. We urge future studies to manipulate proactive control processes required during domain-general task switching to pry apart the subtle ways in which code-switching and context may impact bilingual cognitive flexibility.

### 4.3 When language switching reaches peak efficiency

One interesting finding in the present paper is that, while the colour-shape shifting task yielded expected overall mixing and switching costs (−174.8 ms and – 72.1 ms respectively), the bilingual lexical decision task did not. In fact, sample means indicated a mixing *benefit* (23.0 ms) and virtually zero switching effects (−0.9 ms). Only the estimated marginal means of the significant models (which included by-participant and by-item random effects) revealed a moderate switch cost (−13.2 ms) for English trials (and no mixing effect). The switch cost was not detected for French trials, but a significant large mixing *benefit* (34.5 ms) was observed in that language. The question thus stands: why did our language-specific task yield reduced, absent, and reversed effects compared to our domain-general task?

A first option is the design of the bilingual lexical decision task. Each participant saw the bilingual block last (i.e. after the counterbalanced English/ French blocks). The one-hundred words and one-hundred pseudo-words used in the single-language blocks were repeated in the bilingual block. This procedure may have introduced practice effects into the critical final

bilingual block, thus explaining the mixing benefit we observed. However, the design cannot account for the minimal switching costs since this effect is computed uniquely by examining trials from the bilingual block (Repeated-cue trials – Switched-cue trials).

A second option is that the participants in the current study were simply skilled language switchers, by virtue of their code-switching experience and their primary interactional context. Though inter-subject variability was observed (mixing effects from – 99.4 to 44.5 ms; switching effects from –81.6 to 28.1 ms), the overall sample was efficient at mixing and switching between their languages, thus displaying excellent language-specific sustained and transient cognitive flexibility. The task could have been particularly easy for the participants given that it included a concrete "heads-up" about the language schema they needed to select in the bilingual block. Indeed, participants received a 500 ms language cue (*English* or *français*) before each target stimulus, a cue that is unavailable in their day-to-day conversations. Research indicates that interlocutor identity and language selection occur as early as 100 ms (e.g. Blanco-Elorrieta & Pylkkänen, 2017; Molinaro et al., 2023; see also Hui et al., 2022), and that increased "preparation time" (i.e. cue length) during cognitive flexibility and language switching tasks can diminish switch costs (Mosca & Clahsen, 2016; Shen, 2023; Stasenko et al., 2017). Given their linguistic experience, the 500 ms language cue may have given our participants ample time to engage the correct language schema, thereby mitigating switch costs.

A handful of previous studies have also yielded no language-related mixing costs (e.g. Declerck et al., 2019 for French-English language switching task; Grainger & Beauvillain, 1987 for an orthography-specific French-English lexical decision task; Jiao et al., 2019 for a flanker interleaved with Chinese/English sentences; Kechu and Som (2023) for primed translation equivalency and lexical decision tasks). It is possible that conflict adaptation mechanisms are

involved in these cases. Conflict adaptation refers to scenarios in which "conflict detection initiates behavioural regulation that attenuates the cost of resolving subsequent conflict" (Adler et al., 2020, p. 742). For instance, bilinguals may be faster at responding to a Switched-cue trial if it is preceded by a Switched-cue trial, since conflict resolution mechanisms are already initiated. Given that the bilingual block involves the most conflict, participants may have been faster in this block as their strongly deployed cognitive flexibility mechanisms regulated responses to incoming conflict. Again, this points towards skillful proactive control among our dual-language code-switchers.

Keeping in mind models of bilingual adaptation (e.g. DeLuca et al., 2020; Pliatsikas, 2020), we speculate that the neural substrates related to language-specific cognitive flexibility had already reached "peak efficiency" (stage 3 of the Dynamic Restructing Model; Pliatsikas, 2020) among our participants. The dual-language interactional context and experience with language switching have indeed been linked to cortical (Claussenius-Kalman et al., 2021; Marin-Marin et al., 2022) and subcortical (Korenar et al., 2023b) differences. By contrast, our participants were likely much less experienced with the domain-general cognitive flexibility imposed by our colour-shape shifting task; these neural structures were perhaps still at the stage of initial exposure or consolidation (Pliatsikas, 2020). This might explain why, at the behavioural level, our participants exhibited robust domain-general effects, and no such language-specific equivalents. Of course, this would again imply that language-specific training bears only an impact on language-specific abilities and is dissociated from domain-general skills (see also Weissberger et al., 2012 for aging-related evidence).

### 4.4 Limitations: Can code-switching be accurately self-reported

It is important to remember that the main conclusions drawn from this study rely on (1) the validity of the constructs established by the BSwQ, and (2) the accuracy of the participants' self-reports. We would like to end by addressing both of these limitations.

Though the vast majority of studies examining the relationship between code-switching and executive functioning have utilised questionnaires of some kind (though see Yim & Bialystok, 2012; Lai & O'Brien for objective measures), there is no consensus when it comes to the exact questions used to quantify code-switching behaviour (e.g. BSwQ, Language and Social Background Questionnaire, (Revised) Bilingual Interactional Context Questionnaire, Ecological Momentary Assessments, Bilingual Code-Switching Profile, idiosyncratic Likert scales ...) (see Dass et al., 2024). Arguably, questionnaires validated with psychometric evaluations are favourable, as they give researchers information about the internal consistency of measured constructs. In the case of the BSwQ, the sets of questions for each construct were found to have good internal reliability (alpha = 0.72–0.92), and they were correlated with the expected language behaviours (proficiency, usage, AoA, etc.) in the expected direction. This was also true for the "unintentional switching" construct (alpha = 0.72); though unintentional switching intuitively engenders difficult self-evaluation, authors have pointed out that it can also be emotionally salient and therefore memorable (Jylkkä et al., 2020). However, the BSwQ was validated on a (large) sample of Spanish-Catalan bilinguals in Spain. The question stands whether its usage is comparable among other populations, including the sample in the present study.

Going further, in language research (and in social science in general), self-reports have often been criticised for their subjectivity. Some common measure "contaminations" (Howard,

1994, p. 399) may be particularly important when it comes to self-reporting code-switching habits. First, the social desirability bias predisposes participants to report their behaviour based on what is socially permissible (Paulhus, 1984). As code-switching is often stigmatised (including in the region under study; Gosselin & Sabourin, 2023), participants may be underreporting their switching habits (see Muysken, 2000). However, this generalised underreporting arguably affects the sample as a whole, making the recorded continuum relatively accurate, albeit shifted. Second, participants may change their self-report based on their implicit frame of reference (the reference bias: Lira et al., 2022). This is further complexified by the fact that existing questionnaires like the BSwQ do not distinguish between current switching habits and lifetime switching habits (see Olson, 2022 for a questionnaire which aims to fill this gap).

The BSwQ was developed before the ACH (Green & Abutalebi, 2013); it should be noted that its questions (and constructs) were not necessarily designed to map onto interactional contexts. Nevertheless, we maintain that the intentionality of code-switching is a good proxy for interactional contexts (see Gosselin & Sabourin, 2023 for a discussion). If speakers are producing (non-pathological) unintentional switches, it is likely because they are assuming common linguistic knowledge and background with their co-interlocutors, as is the case in a dense code-switching context (but not in a dual-language context). The present study (see also Gosselin & Sabourin, 2023) showed opposite effects for intentional and unintentional switching on language-specific cognitive flexibility. The direction of these effects fall in line with a priori predictions about processing according to the ACH, bolstering the intentional/unintentional distinction.

Finally, some empirical research lends credence to participants' accuracy in self-reporting their switching habits, including the BSwQ itself. For instance, Cox and colleagues (2020) asked

Spanish-English bilinguals to produce speech during an autobiographical memory task; these same bilinguals later self-reported their code-switching habits by completing the BSwQ. The authors found that participants were generally accurate in their self-reports: bilinguals who code-switched more during the oral production task also self-reported more switching. It is worth bearing in mind that the participants were reflecting on their recent code-switching frequency (i.e. from the production task). Self-reports may not be as accurate when participants are asked to evaluate their general behaviour (e.g. Lai & O'Brien, 2020). This being said, future research should attempt to complement self-reported measures with objective measures when time and resources permit.

### 4.5 Conclusion

The current study examined the link between code-switching habits and cognitive flexibility among bilinguals living in a dual-language environment. Mixing effects indexed sustained cognitive flexibility, while switching effects indexed transient cognitive flexibility. First, the domain-general task yielded robust mixing and switching costs, but these were not reliably predicted by the participants' code-switching habits. Second, the sample displayed minimal switching effects and a mixing benefit in the language-specific task. These effects were positively predicted by the participants' deliberate code-switching: bilinguals who frequently code-switched into the majority language displayed better sustained and transient cognitive flexibility in English. Predictors related to dense code-switching (i.e. unintentional switching, high language entropy) negatively predicted the participants' language-specific sustained cognitive flexibility skills, suggesting that proactive control is put into practice during dual-language switching. Our findings point towards the idea that dual-language code-switching trains language-specific cognitive flexibility, but that this training does not extend to domain-general

abilities. In fact, the peak switching efficiency imparted by the participants' interactional context was powerful enough to lead to absent or reversed language-specific mixing/ switching costs.

**Notes:**

**[1]** Ten participants included in the final sample did not report their age and were thus attributed the mean age of the sample (21.2 years).

**[2]** The "Colour" and "Shape" blocks are collapsed into the single averaged condition, "single-cue block".

**[3]** Dominance scores were computed according to the procedure in Gosselin and Sabourin (2023) (i.e., by examining age of exposure, proficiency, current usage and self-reported first language). These scores range from −4 (very English-dominant) to +4 (very French-dominant).

**Disclosure statement.** No potential conflict of interest was reported by the author(s).

Funding. This work was supported by Ontario Graduate Scholarship; Social Sciences and Humanities Research Council [Joseph-Armand Bombardier Scholarship].

**Author contributions.** LG and LS conceptualised and designed the study. LG collected the data, conducted the analyses, and wrote the manuscript. LS supervised these tasks. Both authors revised and approved the final version of the manuscript.

**Data availability statement.** The full study design (including consent form, questionnaires, and tasks) is available for preview and cloning at https://app.gorilla.sc/openmaterials/429412. The materials, R scripts, and anonymized data (by-trial and aggregated) are available on OSF at https://osf.io/qdsv4/?view_only=a2566d611fb44cdb89ab895591a338e5.

**Author's note.** Preliminary results obtained by analyzing initial subsets of the data were also presented at the *XV International Symposium of Psycholinguistics* (June 2021) and the *31st Canadian Society for Brain, Behaviour and Cognitive Science* meeting (June 2021). This study was not pre-registered.

## References

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://psycnet.apa.org/doi/10.1037/xlm0000755

Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance 15: Conscious and nonconscious information processing* (pp. 421–452). The MIT Press.à

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods, 52*, 388-407. https://doi.org/10.3758/s13428-019-01237-x

Barbu, C. A., Gillet, S., & Poncelet, M. (2020). Investigating the effects of language-switching frequency on attentional and executive functioning in proficient bilinguals. *Frontiers in Psychology, 11*, 1078. https://doi.org/10.3389/fpsyg.2020.01078

Barbu, C., Orban, S., Gillet, S., & Poncelet, M. (2018). The impact of language switching frequency on attentional and executive functioning in proficient bilingual adults. *Psychologica Belgica, 58*(1), 115-127. https://doi.org/10.5334/pb.392

Beatty-Martínez, A. L., Navarro-Torres, C. A., & Dussias, P. E. (2020a). Codeswitching: A bilingual toolkit for opportunistic speech planning. *Frontiers in Psychology, 11*, 1699. https://doi.org/10.3389/fpsyg.2020.01699

Beatty-Martínez, A. L., Navarro-Torres, C. A., Dussias, P. E., Bajo, M. T., Guzzardo Tamargo, R. E., & Kroll, J. F. (2020b). Interactional context mediates the consequences of bilingualism for language and cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(6), 1022-1047. https://psycnet.apa.org/doi/10.1037/xlm0000770

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

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

Blanco-Elorrieta, E., & Pylkkänen, L. (2017). Bilingual language switching in the laboratory versus in the wild: The spatiotemporal dynamics of adaptive language control. *Journal of Neuroscience, 37*(37), 9022-9036. https://doi.org/10.1523/JNEUROSCI.0553-17.2017

Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends in Cognitive Sciences, 16*(2), 106-113. https://doi.org/10.1016/j.tics.2011.12.010

Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron, 39*(4), 713-726. https://doi.org/10.1016/S0896-6273(03)00466-5

Chan, C. G., Yow, W. Q., & Oei, A. (2020). Active bilingualism in aging: Balanced bilingualism usage and less frequent language switching relate to better conflict monitoring and goal maintenance ability. *The Journals of Gerontology: Series B, 75*(9), e231-e241. https://doi.org/10.1093/geronb/gbaa058

Claussenius-Kalman, H. L., Vaughn, K. A., Archila-Suerte, P., & Hernandez, A. E. (2021). Highly proficient, balanced bilingualism is related to thinner cortex in two cognitive control regions. *Annals of the New York Academy of Sciences, 1485*(1), 83-94. https://doi.org/10.1111/nyas.14491

Colzato, L. S., Van Leeuwen, P. J., Van Den Wildenberg, W., & Hommel, B. (2010). DOOM'd to switch: superior cognitive flexibility in players of first person shooter games. *Frontiers in Psychology, 1*, 8. https://doi.org/10.3389/fpsyg.2010.00008

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1283-1296. https://psycnet.apa.org/doi/10.1037/0278-7393.26.5.1283

Cox, J. G., LaBoda, A., & Mendes, N. (2020). “I’m gonna Spanglish it on you”: Self-reported vs. oral production of Spanish–English codeswitching. *Bilingualism: Language and Cognition*, *23*(2), 446–458. https://doi.org/10.1017/ S1366728919000129

Dass, R., Smirnova-Godoy, I., McColl, O., Grundy, J. G., Luk, G., & Anderson, J. A. (2024). A Content Overlap Analysis of bilingualism questionnaires: Considering diversity. *Bilingualism: Language and Cognition*, 1-7. https://doi.org/10.1017/S1366728923000767

De Baene, W., Duyck, W., Brass, M., & Carreiras, M. (2015). Brain circuit for cognitive control is shared by task and language switching. *Journal of Cognitive Neuroscience, 27*(9), 1752-1765. https://doi.org/10.1162/jocn_a_00817

Declerck, M., Koch, I., Duñabeitia, J. A., Grainger, J., & Stephan, D. N. (2019). What absent switch costs and mixing costs during bilingual language comprehension can tell us about language control. *Journal of Experimental Psychology: Human Perception and Performance, 45*(6), 771-789. https://psycnet.apa.org/doi/10.1037/xhp0000627

DeLuca, V., Segaert, K., Mazaheri, A., & Krott, A. (2020). Understanding bilingual brain function and structure changes? U bet! A unified bilingual experience trajectory model. *Journal of Neurolinguistics, 56*, 100930. https://doi.org/10.1016/j.jneuroling.2020.100930

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*, 135-168. https://doi.org/10.1146/annurev-psych-113011-143750

Faroqi-Shah, Y., & Wereley, S. (2022). Investigation of code-switching cost in conversation and self-paced reading tasks. *International Journal of Bilingualism, 26*(3), 308-333. https://doi.org/10.1177/13670069211056438

Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., Augustinova, M., & Pallier, C. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods, 42*, 488– 496. https://doi.org/10.3758/BRM.42.2.488

Festman, J., Rodriguez-Fornells, A., & Münte, T. F. (2010). Individual differences in control of language interference in late bilinguals are mainly related to general executive abilities. *Behavioral and Brain Functions*, *6*(1), 1–12. https://doi.org/10.1186/1744-9081-6-5

Freeman, M. R., Robinson Anthony, J. J., Marian, V., & Blumenfeld, H. K. (2022). Individual and sociolinguistic differences in language background predict Stroop performance. *Frontiers in Communication, 7*, 105. https://doi.org/10.3389/fcomm.2022.865965

Garbin, G., Sanjuan, A., Forn, C., Bustamante, J. C., Rodríguez-Pujadas, A., Belloch, V., Hernandez, M., Costa, A., & Ávila, C. (2010). Bridging language and attention: Brain basis of the impact of bilingualism on cognitive control. *Neuroimage*, *53*(4), 1272–1278. https://doi.org/10.1016/j.neuroimage.2010.05.078

Goldfarb, L., & Henik, A. (2007). Evidence for task conflict in the Stroop effect. *Journal of Experimental Psychology: Human Perception and Performance, 33*(5), 1170–1176. https://doi.org/10.1037/0096-1523.33.5.1170

Gollan, T. H., & Ferreira, V. S. (2009). Should I stay or should I switch? A cost–benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(3), 640-665. https://psycnet.apa.org/doi/10.1037/a0014981

Gosselin, L., & Sabourin, L. (2021). Lexical-semantic processing costs are not inherent to intra-sentential code-switching: The role of switching habits. *Neuropsychologia, 159*, 107922. https://doi.org/10.1016/j.neuropsychologia.2021.107922

Gosselin, L., & Sabourin, L. (2023). Language athletes: Dual-language code-switchers exhibit inhibitory control advantages. *Frontiers in Psychology, 14*, 1150159. https://doi.org/10.3389/fpsyg.2023.1150159

Grainger, J., & Beauvillain, C. (1987). Language blocking and lexical access in bilinguals. *The Quarterly Journal of Experimental Psychology Section A, 39*(2), 295-319. https://doi.org/10.1080/14640748708401788

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

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

Guerra-Carrillo, B., Katovich, K., & Bunge, S. A. (2017). Does higher education hone cognitive functioning and learning efficacy? Findings from a large and diverse sample. *PloS one, 12*(8), e0182276. https://doi.org/10.1371/journal.pone.0182276

Gullifer, J. W., Kroll, J. F., & Dussias, P. E. (2013). When language switching has no apparent cost: Lexical access in sentence context. *Frontiers in Psychology, 4*, 278. https://doi.org/10.3389/fpsyg.2013.00278

Han, X., Wei, L., & Filippi, R. (2022). Modulating effects of interactional contexts on bilinguals' cognitive control: Evidence for the Adaptive Control Hypothesis. *International Journal of Bilingualism*. https://doi.org/10.1177/13670069221102011

Hartanto, A., & Yang, H. (2016). Disparate bilingual experiences modulate task-switching advantages: A diffusion-model analysis of the effects of interactional context on switch costs. *Cognition, 150*, 10-19. https://doi.org/10.1016/j.cognition.2016.01.016

Hartanto, A., & Yang, H. (2020). The role of bilingual interactional contexts in predicting interindividual variability in executive functions: A latent variable analysis. *Journal of Experimental Psychology: General*, *149*(4), 609–633. https://doi.org/10.1037/xge0000672

Hofweber, J., Marinis, T. & Treffers-Daller, J. (2019). Predicting executive functions in bilinguals using ecologically valid measures of code-switching behaviour. In I. A. Sekerina, L. Spradlin & V. Valian (Eds.), *Bilingualism, Executive Functions, and Beyond. Questions and Insights* (pp. 181–205). Amsterdam: Benjamins. https://dx.doi.org/10.1075/sibil.57

Hofweber, J., Marinis, T., & Treffers-Daller, J. (2020a). How different code-switching types modulate bilinguals' executive functions: A dual control mode perspective. *Bilingualism: Language and Cognition*, *23*(4), 909–925. https://doi.org/10.1017/S1366728919000804

Hofweber, J. E., Marinis, T., & Treffers-Daller, J. (2020b). Experimentally induced language modes and regular code-switching habits boost bilinguals' executive performance: evidence from a within-subject paradigm. *Frontiers in Psychology*, *11*, 542326. https://dx.doi.org/10.3389/fpsyg.2020.542326

Howard, G. S. (1994). Why Do People Say Nasty Things About Self-Reports? *Journal of Organizational Behavior, 15*(5), 399–404. https://doi.org/10.1002/job.4030150505

Hui, N. Y., Fong, M. C. M., & Wang, W. S. (2022). Bilingual Prefabs: No Switching Cost Was Found in Cantonese–English Habitual Code-Switching in Hong Kong. *Languages, 7*(3), 198. https://doi.org/10.3390/languages7030198

Jevtović, M., Duñabeitia, J. A., & de Bruin, A. (2020). How do bilinguals switch between languages in different interactional contexts? A comparison between voluntary and mandatory language switching. *Bilingualism: Language and Cognition, 23*(2), 401-413. https://doi.org/10.1017/S1366728919000191

Jiang, S., Ma, L., & Chen, B. (2023). Dynamic engagement of cognitive control in intra-sentential code-switching during comprehension. *Bilingualism: Language and Cognition, 26*(1), 62-77. https://doi.org/10.1017/S1366728922000323

Jiao, L., Grundy, J. G., Liu, C., & Chen, B. (2020). Language context modulates executive control in bilinguals: Evidence from language production. *Neuropsychologia, 142*, 107441. https://doi.org/10.1016/j.neuropsychologia.2020.107441

Jiao, L., Liu, C., Liang, L., Plummer, P., Perfetti, C. A., & Chen, B. (2019). The contributions of language control to executive functions: From the perspective of bilingual comprehension. *Quarterly Journal of Experimental Psychology, 72*(8), 1984-1997. https://doi.org/10.1177/1747021818821601

Jylkkä, J., Laine, M., & Lehtonen, M. (2021). Does language switching behavior rely on general executive functions?. *Bilingualism: Language and Cognition, 24*(3), 583-595. https://doi.org/10.1017/S1366728920000619

Jylkkä, J., Lehtonen, M., Lindholm, F., Kuusakoski, A., & Laine, M. (2018). The relationship between general executive functions and bilingual switching and monitoring in language production. *Bilingualism: Language and Cognition*, *21*(3), 505–522. https://doi.org/10.1017/S1366728917000104

Jylkkä, J., Soveri, A., Laine, M., & Lehtonen, M. (2020). Assessing bilingual language switching behavior with Ecological Momentary Assessment. *Bilingualism: Language and Cognition, 23*(2), 309-322. https://doi.org/10.1017/S1366728918001190

Jylkkä, J., Soveri, A., Wahlström, J., Lehtonen, M., Rodríguez-Fornells, A., & Laine, M. (2017). Relationship between language switching experience and executive functions in bilinguals: an Internet-based study. *Journal of Cognitive Psychology, 29*(4), 404-419. https://doi.org/10.1080/20445911.2017.1282489

Kałamała, P., Chuderski, A., Szewczyk, J., Senderecka, M., & Wodniecka, Z. (2022). Bilingualism caught in a net: A new approach to understanding the complexity of bilingual experience. *Journal of Experimental Psychology: General*. https://psycnet.apa.org/doi/10.1037/xge0001263

Kałamała, P., Szewczyk, J., Chuderski, A., Senderecka, M., & Wodniecka, Z. (2020). Patterns of bilingual language use and response inhibition: A test of the adaptive control hypothesis. *Cognition, 204*, 104373. https://doi.org/10. 1016/j.cognition.2020.104373

Kalanthroff, E., Avnit, A., Henik, A., Davelaar, E.J., Usher, M. (2015). Stroop proactive control and task conflict are modulated by concurrent working memory load. *Psychonomic Bulletin & Review, 22*, 869–875. https://doi.org/10.3758/s13423-014-0735-x

Kechu, O., & Som, B. (2023). Is task specific control among balanced bilinguals a result of nature of acquisition and language use?. *Journal of Cultural Cognitive Science, 7*(1), 23-42. https://doi.org/10.1007/s41809-022-00114-y

Korenar, M., Treffers-Daller, J., & Pliatsikas, C. (2023a). Two languages in one mind: Insights into cognitive effects of bilingualism from usage-based approaches. *Naše řeč, 106*(1), 24-46.

Korenar, M., Treffers-Daller, J., & Pliatsikas, C. (2023b). Dynamic effects of bilingualism on brain structure map onto general principles of experience-based neuroplasticity. *Scientific Reports, 13*(1), 3428. https://doi.org/10.1038/s41598-023-30326-3

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging, 15*(1), 126-147. https://psycnet.apa.org/doi/10.1037/0882-7974.15.1.126

Lai, G., & O'Brien, B. A. (2020). Examining language switching and cognitive control through the adaptive control hypothesis. *Frontiers in Psychology, 11*, 1171. https://doi.org/10.3389/fpsyg.2020.01171

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin, 144*(4), 394–425. https://doi.org/10.1037/bul0000142

Li, X., Ng, K. K., Wong, J. J. Y., Lee, J. W., Zhou, J. H., & Yow, W. Q. (2021). Bilingual language entropy influences executive functions through functional connectivity and signal variability. *Brain and Language, 222*, 105026. https://doi.org/10.1016/j.bandl.2021.105026

Lira, B., O'Brien, J.M., Peña, P.A., Galla, B.M., D'Mello, S., Yeager, D.S., Kautz, T., Munkacsy K., Duckworth, A.L. (2022). Large studies reveal how reference bias limits policy applications of self-report measures. *Scientific Reports, 12,* 19189. https://doi.org/10.1038/s41598-022-23373-9

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology, 6*, 1171. https://doi.org/10.3389/fpsyg.2015.01171

López-Penadés, R., Sanchez-Azanza, V. A., Aguilar-Mediavilla, E., Buil-Legaz, L., & Adrover-Roig, D. (2020). Associations of natural language switching with executive control in early bilinguals. *Bilingualism: Language and Cognition, 23*(5), 1008-1019. https://doi.org/10.1017/S1366728920000012

Marin-Marin, L., Costumero, V., Ávila, C., & Pliatsikas, C. (2022). Dynamic effects of immersive bilingualism on cortical and subcortical grey matter volumes. *Frontiers in Psychology, 13*, 886222. https://doi.org/10.3389/fpsyg.2022.886222

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*(2), 227–234. https://doi.org/10.1037/h0031564

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49-100. https://doi.org/10.1006/cogp.1999.0734

Molinaro, N., Nara, S., & Carreiras, M. (2023). Early language dissociation in the bilingual mind: Sensor-level MEG evidence. *bioRxiv*. https://doi.org/10.1101/2023.03.22.533825

Moll, J., Oliveira-Souza, R. D., Moll, F. T., Bramati, I. E., & Andreiuolo, P. A. (2002). The cerebral correlates of set-shifting: an fMRI study of the trail making test. *Arquivos de Neuro-psiquiatria, 60*, 900-905. https://doi.org/10.1590/S0004-282X2002000600002

Mosca, M., & Clahsen, H. (2016). Examining language switching in bilinguals: The role of preparation time. *Bilingualism: Language and Cognition, 19*(2), 415-424. https://doi.org/10.1017/S1366728915000693

Muysken, P. (2000). *Bilingual speech: A typology of code-mixing*. Cambridge University Press.

Navarro, E., DeLuca, V., & Rossi, E. (2022). It takes a village: Using network science to identify the effect of individual differences in bilingual experience for Theory of Mind. *Brain Sciences, 12*(4), 487. https://doi.org/10.3390/brainsci12040487

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*(5), 548-567.

Olson, D. J. (2022). The Bilingual Code-Switching Profile (BCSP): Assessing the reliability and validity of the BCSP questionnaire. *Linguistic Approaches to Bilingualism*, *14*(3), 400-433. https://doi.org/10.1075/lab.21039.ols

Ooi, S. H., Goh, W. D., Sorace, A., & Bak, T. H. (2018). From Bilingualism to Bilingualisms: Bilingual experience in Edinburgh and Singapore affects attentional control differently. *Bilingualism: Language and Cognition*, *21*(4), 867–879. https://doi.org/10.1017/S1366728918000020

Paap, K. R. (2023). *The bilingual advantage in executive functioning hypothesis : how the debate provides insight into psychology's replication crisis* (1st ed.). Routledge. https://doi.org/10.4324/9781003308027

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., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology, 66*(2), 232-258. https://doi.org/10.1016/j.cogpsych.2012.12.002

Paap, K. R., Mason, L., & Anders-Jefferson, R. (2021). Predictions about the cognitive consequences of language switching on executive functioning inspired by the adaptive control hypothesis fail more often than not. *Brain Sciences, 11*(9), 1217. https://doi.org/10.3390/brainsci11091217

Paap, K. R., Myuz, H. A., Anders, R. T., Bockelman, M. F., Mikulinsky, R., & Sawi, O. M. (2017). No compelling evidence for a bilingual advantage in switching or that frequent language switching reduces switch cost. *Journal of Cognitive Psychology, 29*(2), 89-112. https://doi.org/10.1080/20445911.2016.1248436

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*(3), 598–609. https://doi.org/10.1037/0022-3514.46.3.598

Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs: General and Applied, 76*(27), 1-23. https://psycnet.apa.org/doi/10.1037/h0093840

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

Plonsky, L., & Ghanbar, H. (2018). Multiple regression in L2 research: A methodological synthesis and guide to interpreting R2 values. *The Modern Language Journal*, *102*(4), 713–731. https://doi.org/10.1111/modl.12509

Poplack, S. (1980). Sometimes I' ll start a sentence in Spanish y termino en español: Toward a typology of code-switching. *Linguistics, 18,* 581‑618. https://doi.org/10.1515/ling.1980.18.7-8.581

Prior, A., & Gollan, T. H. (2011). Good language-switchers are good task-switchers: Evidence from Spanish–English and Mandarin–English bilinguals. *Journal of the International Neuropsychological Society, 17*(4), 682-691. https://doi.org/10.1017/S1355617711000580

Prior, A., & Gollan, T. H. (2013). The elusive link between language control and executive control: A case of limited transfer. *Journal of Cognitive Psychology, 25*(5), 622-645. https://doi.org/10.1080/20445911.2013.821993

Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and Cognition, 13*(2), 253-262. https://doi.org/10.1017/S1366728909990526

Rodriguez-Fornells, A., Kramer, U., Lorenzo-Seva, U., Festman, J., & Münte, T. F. (2012). Self-assessment of individual differences in language switching. *Frontiers in Psychology*, *2*, 388. https://doi.org/10.3389/fpsyg.2011.00388

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*(2), 207-231. https://psycnet.apa.org/doi/10.1037/0096-3445.124.2.207

Rossi, E., Dussias, P. E., Diaz, M., van Hell, J. G., & Newman, S. (2021). Neural signatures of inhibitory control in intra-sentential code-switching: Evidence from fMRI. *Journal of Neurolinguistics, 57*, 100938. https://doi.org/10.1016/j.jneuroling.2020.100938

Rubin, O., & Meiran, N. (2005). On the origins of the task mixing cost in the cuing task-switching paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(6), 1477-1491. https://psycnet.apa.org/doi/10.1037/0278-7393.31.6.1477

Sabourin, L., Brien, C., & Burkholder, M. (2014). The effect of age of L2 acquisition on the organization of the bilingual lexicon: Evidence from masked priming. *Bilingualism: Language and Cognition, 17*(3), 542-555. https://doi.org/10.1017/S1366728913000643

Sabourin, L., Leclerc, J.-C., Lapierre, M., Burkholder, M. & Brien, C. (2016). The language background questionnaires in L2 research: Teasing apart the variables. In L. Hracs (Ed.), *Proceedings of the 2016 Annual Conference of the Canadian Linguistics Association*. Calgary, AB

Sanchez-Azanza, V. A., López-Penadés, R., Aguilar-Mediavilla, E., & Adrover-Roig, D. (2020). Latent variable evidence on the interplay between language switching frequency and executive control in Spanish-Catalan bilinguals. *International Journal of Bilingualism, 24*(5-6), 912-930. https://doi.org/10.1177/1367006920902525

Savin, N. E., & White, K. J. (1977). The Durbin-Watson test for serial correlation with extreme sample sizes or many regressors. *Econometrica: Journal of the Econometric Society*, 1989-1996.

Shen, Q. (2023). Investigating the modulation of active preparation and passive dissipation on inhibitory control processes in the language switching paradigm. *Frontiers in Psychology, 14*, 1065268. https://doi.org/10.3389/fpsyg.2023.1065268

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Soveri, A., Rodriguez-Fornells, A., & Laine, M. (2011). Is there a relationship between language switching and executive functions in bilingualism? Introducing a within group analysis approach. *Frontiers in Psychology, 2*, 183. https://doi.org/10.3389/fpsyg.2011.00183

Stasenko, A., Matt, G. E., & Gollan, T. H. (2017). A relative bilingual advantage in switching with preparation: Nuanced explorations of the proposed association between bilingualism and task switching. *Journal of Experimental Psychology: General, 146*(11), 1527-1550. https://psycnet.apa.org/doi/10.1037/xge0000340

Stuss, D. T., Bisschop, S. M., Alexander, M. P., Levine, B., Katz, D., & Izukawa, D. (2001). The trail making test: A study in focal lesion patients. *Psychological Assessment, 13*(2), 230–239. https://doi.org/10.1037/1040-3590.13.2.230

Surrain, 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

Treffers-Daller, J., Ongun, Z., Hofweber, J., & Korenar, M. (2020). Explaining individual differences in executive functions performance in multilinguals: The impact of code-switching and alternating between multicultural identity styles. *Frontiers in Psychology, 11*, 561088. https://doi.org/10.3389/fpsyg.2020.561088

van den Berg, F., Brouwer, J., Tienkamp, T. B., Verhagen, J., & Keijzer, M. (2022). Language entropy relates to behavioral and pupil indices of executive control in young adult bilinguals. *Frontiers in Psychology, 13*, 1760. https://doi.org/10.3389/fpsyg.2022.864763

Weissberger, G. H., Wierenga, C. E., Bondi, M. W., & Gollan, T. H. (2012). Partially overlapping mechanisms of language and task control in young and older bilinguals. *Psychology and Aging, 27*(4), 959-974. https://psycnet.apa.org/doi/10.1037/a0028281

Xu, Q., Markowska, M., Chodorow, M., & Li, P. (2021). Modeling bilingual lexical processing through code-switching speech: A network science approach. *Frontiers in Psychology*, *12*, 662409. https://doi.org/10.3389/fpsyg.2021.662409

Yeh, L. H., & Lim, J. Y. (2021). Multilinguals in a Monolingual Community: The Influence of Non-linguistic Switching Cost and Awareness of Language Environment on Intentional and Unintentional Language Switch. *Revista Linguagem em Foco, 13*(4), 78-95. https://doi.org/10.46230/2674-8266-13-7354

Yim, O., & Bialystok, E. (2012). Degree of conversational code-switching enhances verbal task switching in Cantonese–English bilinguals. *Bilingualism: Language and Cognition, 15*(4), 873-883. https://doi.org/10.1017/S1366728912000478

Yow, W. Q., & Li, X. (2015). Balanced bilingualism and early age of second language acquisition as the underlying mechanisms of a bilingual executive control advantage: Why variations in bilingual experiences matter. *Frontiers in Psychology, 6*, 164. https://doi.org/10.3389/fpsyg.2015.00164

Yow, W. Q., Oei, A. C., & Li, X. (2017). Novel Evidence for the Bilingual Advantage: Effects of Language Control on Executive Function in Balanced and Unbalanced Dual-Language Users. *39th Annual Meeting of the Cognitive Science Society*.

Zhu, J. D., Blanco-Elorrieta, E., Sun, Y., Szakay, A., & Sowman, P. F. (2022). Natural vs forced language switching: Free selection and consistent language use eliminate significant performance costs and cognitive demands in the brain. *NeuroImage, 247*, 118797. https://doi.org/10.1016/j.neuroimage.2021.118797