Journal of Experimental Psychology: Learning, Memory, and Cognition

Bilingualism and the Executive Function Trade-Off: A Latent Variable Examination of Behavioral and Event-Related Brain Potentials

Mercedes Fernandez, Jonathan B. Banks, Samantha Gestido, and Maria Morales

Online First Publication, October 6, 2022. http://dx.doi.org/10.1037/xlm0001186

Journal of Experimental Psychology:
Learning, Memory, and Cognition

ISSN: 0278-7393

https://doi.org/10.1037/xlm0001186

Bilingualism and the Executive Function Trade-Off: A Latent Variable Examination of Behavioral and Event-Related Brain Potentials

Mercedes Fernandez, Jonathan B. Banks, Samantha Gestido, and Maria Morales
Department of Psychology and Neuroscience, Nova Southeastern University

The impact of bilingualism on the executive functioning constructs of inhibition, shifting, and updating remains unclear, with prior findings yielding inconsistent results. Several explanations for the lack of congruency have been suggested, including the dependence on observed variables, the impact of test modality on performance, and the need to examine the density of dual languages in the environment. To address these concerns, the current study examined differences between a large group of monolingual (n = 109) and bilingual (n = 152) college students on nonlinguistic behavioral and neural measures of inhibition, shifting, and updating using a latent variable approach. We investigated the impact of test modality by presenting each task in the auditory and visual modalities. Additionally, we examined the effects of language balance and language switching in daily life on the measures of executive functioning. Results revealed greater neural expenditure (i.e., higher ERP amplitude) and weaker performance on tasks assessing response inhibition and shifting abilities in bilinguals. Further, although a neural marker of memory updating did not reveal group differences, performance was stronger in monolinguals. These findings were consistent across test modality. Last, language balance was a stronger predictor of behavioral and neural measures than language switching frequency. Our findings highlight the importance of examining differences at the latent level and exploring the influence of linguistic balance.

Keywords: bilingualism, event-related brain potentials, executive function, structural equations modeling

People masterfully modify their behaviors on an ongoing basis to meet the changing demands of their environment. Similarly, people who speak more than one language switch between languages effortlessly and flawlessly to meet environmental linguistic demands. According to the Inhibitory Control (IC) model, when bilinguals speak in one language, their other language is actively suppressed (Green, 1998). The ability to shift between languages and between other behaviors is believed to be carried out in the frontal lobes by inhibitory processes and, more generally, is associated with executive function (EF; Abutalebi et al., 2012). Understanding how inhibitory processes may be strengthened will reveal not only how multiple languages are controlled in the brain but has the potential to guide research and interventions for disorders linked to deficient neural inhibitory processes such as attention-deficit disorder and obsessive-compulsive disorder.

Speculatively, because bilinguals engage inhibitory processes more frequently (i.e., not only to modify behaviors, but also to control language) than monolinguals, this frequent engagement leads to enhanced EF. However, after decades of research, a clear pattern supporting superior EF abilities in bilinguals has not emerged, and some researchers indicate that most study findings do not support such a claim (Lehtonen et al., 2018; Paap et al., 2014, 2015; von Bastian et al., 2016). In fact, some argue that language control in bilinguals is contained within a language mechanism (Paap et al., 2021), whereas others call into question the idea of general inhibition as a construct (Rey-Mermet et al., 2018).

Others, however, maintain that more work needs to be done before one can more confidently speak to the presence or absence of a bilingual advantage (Valian, 2015; van den Noort et al., 2019). Specifically, researchers suggest that the context in which bilinguals speak their languages may impact the relationship between bilingualism and EF and recommend controlling language context. Others suggest that inconsistent results may be related to task impurity and recommend latent variable analysis. Yet other concerns involve the lack of consistent findings within studies related to test modality. To advance the field, Cespón and Carreiras (2020) recommend the addition of event-related brain potential (ERP) measures. ERPs are electrical potentials that are time-locked to a cognitive event and are generated

Samantha Gestido is now at the Department of Clinical and School Psychology, Nova Southeastern University. Maria Morales is now at the Department of Psychology, Florida International University.

We thank all students who worked tirelessly recruiting and testing participants on this project. This work would not have been possible without you. We are also thankful to Alejandra Quintero and Abraham Yacaman for their superb work and contribution to all aspects of this project. A special thanks to Morgan Musgrove for training and supervising students on data collection early on in the project. To Matthew S. Welhaf, we are grateful for your assistance with the design of the data analysis. This project was funded by the National Science Foundation (Award 1632377) awarded to Mercedes Fernandez.

Correspondence concerning this article should be addressed to Mercedes Fernandez, Department of Psychology and Neuroscience, Nova Southeastern University, 3301 College Avenue, Fort Lauderdale, FL 33314, United States. Email: mf934@nova.edu

by populations of neurons within milliseconds after the event. Because of their temporal resolution, ERPs are considered the gold standard for observing neural activity across time.

In response to these concerns, the current study recorded behavioral and ERP activity while bilingual and monolingual participants completed nonlinguistic EF tasks. We used a latent variable approach, included a language context questionnaire, and we administered tasks in two modalities. Below we review the literature, discuss prior work from our laboratory, and present the current study design.

Operationalizing the Language Experience

Researchers state that one of the challenges to this line of research is the difficulty of defining what is meant by bilingualism. For example, a person who is fluent in two languages but mostly speaks in one language would be classified as bilingual just the same as a person who is fluent in two languages and speaks both languages regularly.

Gullifer and Titone (2020) state that a failure to capture the diversity of language experience may explain the absence of a bilingual advantage. Green and Wei (2014) described different contexts in which bilinguals speak their two languages and explain that these different experiences may differentially affect EF abilities: (a) a single context bilingual speaks one language in one context (e.g., at home) and their other language in a different context (e.g., work); (b) the dual language bilingual speaks both languages in the same context (e.g., at home); and (c) a dense code switching bilingual switches between languages within a single conversation for no apparent reason. In support of this theory, a study using a self-report questionnaire on frequency of language switching in different contexts found that dual language context predicted the latent variable of shifting, whereas dense code-switching predicted inhibition and goal maintenance (Hartanto & Yang, 2020).

Studies to date, however, do not provide strong support for the hypothesis that language experience impacts EF. In one study (Paap et al., 2020), participants were carefully screened to select groups of pure single context, dual context, dense code-switching, and monolinguals. Participants were then compared on nine measures of EF generated from nonverbal tasks (Simon, spatial and vertical Stroop, flanker, switch, visual search, ambiguous figures). Results failed to reveal a link between language context and performance. The researchers noted that most studies fail to reveal any such relationship and questioned whether such a link exists.

Another study (Kałamała et al., 2020) evaluated the link between language context and performance on EF tests of response inhibition (Stroop, Go/NoGo, stop-signal and antisaccade) and again failed to reveal a link between language context and performance.

Additionally, studies in which language switching is objectively measured fail to reveal a relationship between language switching and shifting EF. For instance, in one study, participants completed switch tasks and verbal fluency tasks (letter or category fluency; Paap et al., 2017). They also completed a variant of the verbal fluency task—participants switched between letters or between categories to generate exemplars. Results did not reveal any relationship between switch task performance and language switching ability.

Similarly, Woumans et al. (2019) found no relationship between shifting abilities and language switching. Language switching was assessed by asking participants to switch between languages to generate exemplars within a semantic category. In an earlier study, the same research group (Woumans et al., 2015) investigated the link between language switching and inhibition (Simon and attentional network task [ANT]). They found a positive relationship (r = .530, p = .002) between language switching proficiency and inhibition (smaller Simon effect) in their balanced bilingual group but not in their unbalanced group or in a group of interpreters in training. Given that high degree of proficiency is necessary to efficiently switch between languages, it is plausible that this relationship is driven by second language proficiency more so than by language context. Below (under Prior Work from Our Laboratory) we report studies from our laboratory which link neural inhibition to second language proficiency. Summarily, most of these studies do not favor the idea that the context in which bilinguals speak their two languages impacts EF abilities of inhibition or shifting.

Task Impurity

Another concern raised in the literature is the failure to replicate results on the bilingual advantage. For instance, Costa et al. (2008) found a bilingual advantage in young adults, our target population, on the flanker task. Notably, the advantage was only observed on the first two of three trial blocks. If a bilingual advantage results from stronger EF abilities, it would not be expected to disappear after two blocks of trials. Other studies also fail to replicate the findings on the flanker or other tasks (Simon and Stroop) of inhibition (Kousaie & Phillips, 2012; Paap & Greenberg, 2013; Paap et al., 2019, 2020). These replication failures raise the question of whether studies may be capturing task-specific effects unrelated to EF abilities. If so, it would explain why some studies reveal a bilingual advantage while others do not.

More recently, one study revealed a lack of relationship between self-reported measures of self-control and impulsivity and performance on inhibition tasks (flanker, Simon and spatial Stroop) raising the question of the suitability of the tasks used to capture general inhibition (Paap et al., 2021).

Friedman (2016) recommends using a latent model approach (e.g., structural equations modeling, [SEM]) to examine possible effects of bilingualism on EF abilities. In this approach, researchers administer two or more tasks known to capture the EF ability of interest. SEM separates the nonshared variance attributed to task-specific factors (e.g., motor speed) from the shared variance (variance common to all tasks), to reveal an "uncontaminated" measure of EF. By removing task-specific "noise," this approach is likely to yield results that more accurately reflect EF abilities and lead to more consistent findings across studies.

Modality Specific Effects

Inconsistent findings in studies that use tasks presented in different modalities raise yet other concerns about the assumption of a bilingual advantage. That is, if general inhibitory processes are enhanced in bilinguals, test modality should not impact performance. Yet, Luo et al. (2013) found stronger spatial but poorer verbal performance when comparing bilinguals to monolinguals on working memory (WM) tasks. Similarly, we compared bilinguals with monolinguals (discussed in detail below) on nonlinguistic visual and auditory inhibition tasks and found greater amplitude on a neural inhibition marker in bilinguals, but only in the auditory modality (Fernandez et al., 2014). Determining whether bilingual performance on EF tests is tied to test modality will shed light on the role of general inhibition in language control.

Prior Work From Our Laboratory

Studies assessing inhibition usually use versions of the Simon, Stroop, or Flanker task. Indeed, a recent study by Paap et al. (2020) compared bilinguals to monolinguals on these tasks, and their findings did not reveal a bilingual advantage. These tasks require inhibitory control of attention because distracting, but task-irrelevant, stimulus features capture attention and interfere with task performance. Another task that also requires inhibition, the Go/NoGo task, is less frequently used to compare language groups. This task requires inhibitory control to suppress behavioral responses. Noteworthy, research suggests that response suppression in the Go/NoGo task and language switching in bilinguals share similar neural inhibitory mechanisms. Specifically, the amplitude of the N2 ERP component (a negative going wave in the 200–300 ms post-stimulus time window, with maximum amplitude over frontal-central sites) has been linked to inhibition in monolinguals (Falkenstein et al., 1999; Thomas et al., 2009) and to language switching in bilinguals (Jackson et al., 2001).

We speculated that if indeed the neural mechanism engaged to inhibit behavioral responses in the Go/NoGo task is the same one that inhibits the nontarget language when a bilingual speaks, the experience of controlling language would strengthen response inhibition in bilinguals. We further hypothesized that response inhibition may develop in tandem with second language proficiency to control increasing competition between languages. If this hypothesis is correct, one would expect the highest linguistic competition to occur when bilinguals reach language balance. In this case, the highest level of inhibition would be needed to control competition between equally strong languages. Conversely, if language control and response inhibition do not share similar inhibitory mechanisms, neural activity would not distinguish language groups on the Go/NoGo task.

Studies from our laboratory reveal greater N2 amplitude on the Go/NoGo task in bilinguals and link inhibition to second language proficiency. In our first study (Fernandez et al., 2013), we used an auditory Go/NoGo task in which participants responded to target tone pairs with a button press and withheld a motor response to nontarget pairs. We recorded neural activity and behavioral responses while participants performed the task, and we objectively measured language proficiency. Equiprobable Go and NoGo stimuli were presented binaurally. We presented an equal number of Go and NoGo trials because when an unequal number is presented, it cannot be determined whether the neural response on NoGo trials is due to response inhibition or to the relative novelty of the less frequent NoGo stimulus. Thus, to avoid the influence of stimulus probability, we presented Go and NoGo trials with equal frequency. Notably, research shows that equiprobable auditory Go/NoGo tasks elicit control processes to inhibit the Go response and link neural inhibition (N2 ERP component) to performance (Fogarty et al., 2018).

Compared with monolinguals, bilinguals showed greater NoGo N2 amplitude. Additionally, second language proficiency scores were correlated with the NoGo N2 amplitude such that higher second language proficiency was correlated with greater inhibition. To our knowledge this was the first study to link neural inhibition to second language proficiency. Behavioral data revealed similar reaction times (RTs) between groups and errors on NoGo trials were too few to analyze (5 participants in each group incorrectly responded with a button press on one or two trials).

In a follow-up study, we used the same auditory task and added a nonlinguistic visual Go/NoGo task (Fernandez et al., 2014). The two tasks were identical except that the tones were replaced with geometric figures in the visual task. Again, bilinguals showed stronger auditory NoGo N2 amplitude compared with monolinguals. The correlation between the auditory NoGo N2 amplitude and second language proficiency approached but did not reach statistical significance (p = .068) in this study. Unexpectedly, the visual NoGo trials did not reveal N2 amplitude differences between the groups. We speculated that lack of group differences in neural inhibition on visual trials could be due to lack of power (fewer than 20 participants per group). Another possible explanation was that inhibition in bilinguals is tied to the auditory modality since the experience of being bilingual relates to auditory processes. Indeed, research shows that compared with monolinguals, bilinguals "fine tune" their auditory system and are more efficient at processing sounds (Krizman et al., 2012). As in the prior study, behavioral data revealed similar RTs between groups and errors on NoGo trials were too few to analyze.

Together, these studies linked neural inhibition to second language proficiency and support the claim that language switching and response inhibition share similar neural inhibitory mechanisms. Additionally, this work raised the question of whether inhibition in bilinguals is linked to the auditory modality. Because accuracy rates in both groups were near perfect, our findings did not reveal group performance differences in either study. We could not determine whether lack of differences was because none exists or because our tasks were too easy to capture such differences. These unanswered questions are addressed in the design of the current study.

The Current Study

Our goal was to replicate our prior studies on inhibition, extend our research to other components of EF, and address concerns raised in the literature. We used the three-factor (inhibition, shifting, updating) EF model proposed by Miyake et al. (2000) as the foundation for this study. We used the Go/NoGo task and recorded response bias as our behavioral measure and the N2 ERP component as our neural marker of inhibition. Just as in our prior study (Fernandez et al., 2014), we measured the amplitude of the N2 as the difference between the mean P3 and mean N2 amplitude (henceforth referred to as the N2/P3 complex in this article), because these two components have very similar frontal-central scalp distribution and occur close in time in the Go/NoGo task. To assess shifting abilities, we used a switch task and measured speed and accuracy as well as the N2 ERP component of interest. The amplitude of the N2 ERP is larger on switch relative to no-switch trials and has been shown to increase along with improved performance after cognitive training (Gajewski et al., 2017; Gajewski, Ferdinand, et al., 2018). To date, most studies comparing bilinguals with monolinguals on switch tasks do not support the contention that bilinguals have stronger shifting abilities (Paap et al., 2017). Last, as our memory updating task, we used the n-back which requires participants to indicate whether the current stimulus matches the one presented *n*-positions before. We measured speed and accuracy as well as the P3 ERP. This component has a central-parietal scalp distribution and has been shown to be sensitive to memory capacity. Studies reveal that the P3 amplitude increases with increasing WM demands. Daffner et al. (2011) showed that as task demands increased, the P3 amplitude increased in a high WM capacity group. However, in a low WM capacity group, the P3 amplitude decreased as task demands increased. The increase in the P3 amplitude was interpreted as reflecting the additional resources of the high WM capacity group. Notably, ERP studies comparing bilinguals to monolinguals on the *n*-back task do not reveal a bilingual advantage (Kousaie & Phillips, 2017; Morrison et al., 2019).

Our first goal was to reveal neural and behavioral differences between bilinguals and monolinguals at a latent level. We hypothesized that if indeed the experience of managing two languages enhances EF, bilinguals would show better performance and greater neural activity than monolinguals on tasks of inhibition, shifting and updating. Alternatively, if language control and EF do not share the same control mechanism, then bilinguals and monolinguals would show similar performance and neural activity across EF tasks. Our second goal was to reveal the impact of test modality on EF abilities. We hypothesized that if a general inhibitory mechanism controls both language and EF, then group differences would be independent of test modality. Alternatively, if language control is modality specific, then only the auditory modality would reveal group differences. Our third goal was to reveal the relationship between language balance and language switching to each EF factor. We hypothesized that language balance would be a better predictor than language switching of our EF measures.

Method

Participants

English speaking monolinguals (n = 109) and Spanish/English bilinguals (n = 152) between the ages of 18 and 30 years were recruited from Nova Southeastern University (NSU). We selected Spanish-speaking bilinguals because as a Hispanic-Serving Institution, NSU has a large Spanish-speaking student body. Right-handed participants with normal hearing, normal or corrected-to-normal vision, and intact color vision were tested. Participants were excluded if they reported neurological or psychiatric conditions that affect cognition, or if they were taking prescription medications that affect performance. Additionally, participants who only spoke Spanish or who spoke languages other than English and Spanish were excluded.

Measures

Demographics

A demographic questionnaire was administered to obtain information on participant age and sex, parental education, country of birth of participant and of parents, and household income.

Language

A language questionnaire was administered to determine age of second language acquisition, dominant language, and frequency of language usage in different settings (home, school, social settings).

Music Ability

A music abilities questionnaire was administered to determine whether the participant played a musical instrument, years of music training and frequency of playing an instrument. Because music training has been shown to improve WM, we recorded this information and controlled for this variable (George & Coch, 2011; Kausel et al., 2020).

Mind Wandering

A mind wandering questionnaire was administered at the end of each computer task to quantify mind wandering frequency, valence, intention, and awareness during each task. We did not report the results of this questionnaire because it is not pertinent to the current study.

Bilingual Verbal Ability Test

The Bilingual Verbal Ability Test (BVAT; Munoz-Sandoval et al., 1998) contains three subtests (Picture Vocabulary, Oral Vocabulary, Verbal Analogies) and generates an index of second language proficiency and of general verbal ability. In bilinguals, items failed in English are administered in the person's native language and scores are combined to determine overall language ability. We administered all subtests in English and Spanish to all participants, and the sum total of correct responses to the three subtests was used to quantify proficiency in each language. When a participant failed eight consecutive items, that subtest was discontinued, and the next subtest was administered. When a participant scored zero on Picture Vocabulary, the easiest subtest, the other two subtests were not administered, and the participant was assigned a score of zero.

Raven's Standard Progressive Matrices

Raven's Standard Progressive Matrices (Raven et al., 2003), a standardized and culture fair test unaffected by linguistic or ethnic background, is widely used to assess nonverbal fluid intelligence, and we used this test to match groups on nonlinguistic abilities.

Executive Functioning Tasks and Behavioral Measures

Task selection was based on three requirements: (a) the tasks had to tap the target EF components (inhibition, shifting, and updating); (b) the tasks had to be nonverbal because we were testing linguistically diverse participants; (c) equivalent versions of the tasks had to be presented in the auditory and visual modalities. Each participant completed six nonverbal tasks. We created comparable stimuli in the two modalities for each task and maintained similar trial parameters across modalities within task. Each task had two or more levels of difficulty. Order of task presentation and task modality was counterbalanced. Modalities were presented sequentially within task. Participants were given short breaks between tasks and between blocks of trials within task. Descriptive statistics for the tasks can be found in Table 1.

Tasks were created using STIM2 software (Compumedics NeuroScan, Germany). During auditory task presentation, a stationary X in the center of a computer monitor was continuously displayed as a fixation point. Stimuli for all visual tasks were presented in the center of the monitor. A four-button response pad

Table 1
Descriptive Statistics for Language, Fluid Intelligence, Behavioral and ERP Measures by Group

Measure	Monolinguals			Bilinguals		
	M (*SD*)	Skewness	Kurtosis	*M* (*SD*)	Skewness	Kurtosis
Language and fluid intelligence						
BVAT (English) total score	93.06 (9.10)	−0.31	0.13	86.71 (12.64)	−0.11	−0.28
BVAT (Spanish) total score	12.15 (7.69)	0.62	−0.32	56.85 (21.24)	−0.32	−0.77
English-Spanish BVAT difference	−80.91 (10.07)	0.39	0.37	−32.75 (22.26)	−0.69	−0.59
Raven's SPM total score	48.89 (5.70)	−0.83	0.74	47.52 (6.86)	−1.12	1.38
Behavioral responses						
Inhibition – Response Bias						
Auditory Go/NoGo no pressure	−0.19 (0.91)	−0.04	−0.85	0.13 (1.04)	0.05	0.09
Auditory Go/NoGo pressure	0.00 (0.91)	−0.25	−0.39	0.00 (1.06)	1.04	5.26
Visual Go/NoGo no pressure	−0.21 (1.01)	0.29	−0.96	0.15 (0.97)	−0.08	−1.37
Visual Go/NoGo pressure	−0.07 (1.03)	−0.25	−0.30	0.05 (0.98)	−0.29	−0.45
Shifting – Switch cost						
Auditory dimension change	0.29 (0.85)	0.30	−0.34	−0.19 (1.05)	0.13	−0.71
Auditory dimension and hand change	0.33 (0.88)	0.37	−0.72	−0.21 (1.02)	0.05	−0.59
Visual dimension change	0.24 (0.88)	−0.46	−0.30	−0.15 (1.04)	−0.60	−0.26
Visual dimension and hand change	0.24 (0.84)	−0.28	−0.47	−0.16 (1.06)	−0.70	−0.14
Updating – Trade-off						
Auditory 2-back	0.30 (1.48)	−0.65	0.73	0.16 (1.37)	−0.90	0.79
Auditory 3-back	−0.01 (1.44)	−0.42	0.33	−0.11 (1.27)	−0.45	0.01
Visual 2-back	0.38 (1.26)	−1.14	4.05	0.15 (1.17)	−0.83	1.16
Visual 3-back	0.02 (1.31)	−1.59	6.45	−0.17 (1.07)	−0.60	0.56
ERP amplitudes						
Inhibition – N2/P3 complex						
Auditory Go/NoGo no pressure	4.07 (2.37)	0.39	0.58	5.38 (2.54)	1.17	1.81
Auditory Go/NoGo pressure	3.45 (2.3)	0.58	−0.15	3.95 (2.17)	0.63	0.37
Visual Go/NoGo no pressure	6.25 (2.56)	0.38	0.85	7.12 (3.56)	0.38	0.40
Visual Go/NoGo pressure	5.05 (2.66)	0.02	0.17	5.21 (2.68)	0.57	0.83
Shifting – N2 ERP						
Auditory Switch trials	0.58 (1.77)	−1.34	6.85	0.15 (1.43)	0.04	3.96
Auditory No-switch trials	1.00 (2.45)	−1.10	5.31	0.70 (1.87)	1.55	12.08
Visual Switch trials	1.16 (1.75)	0.85	1.55	0.55 (1.60)	0.74	2.71
Visual No-switch trials	1.20 (1.65)	0.34	0.20	0.59 (1.60)	0.78	3.70
Updating – P3 ERP						
Auditory 0-back	3.39 (4.53)	0.47	4.26	2.53 (3.40)	0.68	1.97
Auditory 2- and 3-back combined	0.69 (2.98)	0.28	3.06	1.03 (2.64)	−0.16	0.78
Visual 0-back	8.39 (4.66)	0.16	−0.52	6.82 (3.72)	0.18	0.03
Visual 2- and 3-back combined	5.54 (3.68)	0.35	−0.38	5.36 (3.64)	0.77	1.54

Note. BVAT = Bilingual Verbal Ability Test; SPM = Standard Progressive Matrices; ERP = event-related brain potential.

was used to record behavioral responses. Headphones were used to present auditory stimuli and to reduce environmental noise during visual task presentation.

Inhibition

The auditory Go/NoGo task consisted of high (1,100 Hz) and low-pitch (1,000 Hz) tones (80-ms duration, 5-ms rise and fall times; intensity 70 dB SPL). Each trial consisted of two tones separated by 1,200 ms. For each trial, when a target tone was followed by another target tone (target/target, 36% of trials), participants pressed a response button to the second tone. When the target tone was followed by a nontarget tone (target/nontarget, 36% trials), participants withheld their response. The remaining trials, which started with the nontarget tone (nontarget/target 14%; nontarget/nontarget 14%), were not analyzed. The Go/NoGo task consisted of 200 trials, divided into four blocks of 50 trials, with an intertrial interval (ITI) of 1,800 ms. To increase task difficulty, an auditory signal (300 ms I kHz, 60 dB SPL tone burst) was sounded if the participant did not respond within 600 ms after stimulus onset. This time pressure was introduced after the first 100 trials. Participants focused on the fixation point, responded as quickly as possible to target tone pairs (Go trials), and withheld responding on NoGo trials. The task began after participants read the instructions on the monitor and practiced the task. After the second block of trials, participants were trained on the task with the added time pressure (tone burst), after which the remaining two blocks of trials were presented. Participants were instructed to respond quickly to avoid the tone burst.

The visual Go/NoGo task was similar to the auditory task except that the tones were replaced with red and green circles presented against a black background subtending a visual angle of 2.9°.

Response bias was computed as our measure of inhibition on auditory and visual Go/NoGo tasks. We used the response bias formula proposed in Snodgrass and Corwin (1988), [(C_L = .5 [ln {[(1 − FA) (1 − H)]/[(H) (FA)]}], where ln is the natural log, H the hit rate on Go trials, and FA the false alarm rate on NoGo trials. Response bias reflects response tendencies across all (Go and NoGo) trials or the extent to which one response is more probable than another. Positive values reflect greater tendencies toward withholding a response, indicative of stronger response inhibition. Negative values reflect greater tendencies toward responding. Values close to zero indicate no bias toward responding or withholding a response and reflect better inhibitory control. Response bias was separately computed for trials with and without time pressure, and scores were *z*-transformed for latent variable analyses.

Shifting

The auditory switch task consisted of tones that differed along two dimensions, pitch (high [1,500 Hz] and low-pitch [500 Hz]) and duration (short [80 ms] or long [200 ms]). Each trial began with a 300-ms cue, to indicate target dimension, followed by a 1,500-ms cue-to-trial interval (CTI). The target was presented for 80 or 200 ms. The interval between the start of the target and the beginning of the next trial was 2,400 ms. Four blocks of 50 trials were presented after the practice trials. Switch and no-switch trials were equiprobable and occurred in a pseudorandom order with similar frequency within block. Target response hand was also equiprobable. Participants responded by pressing a right or left-hand response-pad button to indicate high or low pitch and short or long duration.

The visual switch task was similar to the auditory task except that stimuli, both targets and cue, were visual, and the targets were presented for 150 ms because duration was not variable as in the auditory task. Targets (circles) differed along two dimensions, color (red/green) and size (small/large) subtending a visual angle of 2.9° or 6.2°, respectively.

We computed switch cost by combining both accuracy rates and RT. We followed the binning procedure described in Draheim et al. (2016), to generate scores for Dimension Switch Trials and Dimension and Hand Switch Trials. In this procedure, the mean RT on accurate No-switch trials is calculated and subtracted from the RT of each individual accurate switch trial. Notably, this metric generates a score that represents how fast a participant responded on a switch trial relative to his or her own mean RT on No-switch trials. The scores for all subjects combined are rank ordered into deciles and assigned a bin value of 1 to 10. Incorrect responses are assigned a bin value of 20. Bin scores are then summed to generate a single score per subject. We multiplied scores by -1 so that higher scores (lower switch cost) reflect better performance. We computed bin scores separately for trials in which the target dimension switched between trials and for trials in which both the target dimension and the response hand switched between trials. Scores were subsequently z-transformed for latent variable analyses.

Updating

The auditory *n*-back task consisted of rhythmic three-tone melodies obtained from Schneiders et al. (2012). The *n*-back requires participants to indicate whether the current stimulus is the same as the one presented *n* positions before. Participants responded by pressing a right or left-hand button on the response pad to indicate a match or no match, respectively. For the 0-back position, nontarget stimuli were melodies, and the target stimulus was a single tone. The cue indicating the target *n*-back position (0, 2, 3) was presented for 3000 ms, followed by a 500 ms black screen. The stimulus followed and was presented for 500 ms followed by a 2,500-ms ITI. A total of 195 trials (54 targets) were presented, separated into three sets of 65 trials. Each set consisted of three blocks, one per *n*-back position.

The visual *n*-back task was identical to the auditory task except that the stimuli were black and white patterned squares, subtending a visual angle of 6.2°. The stimulus for the 0-back position was a black and white patterned square with a gray dot in the center. A plus sign appeared in the center of the monitor as a fixation point during the 2,500 ms ITI.

For both auditory and visual *n*-back tasks, we computed Balance Integration Scores (BIS, referred to as Trade-off Scores) which are derived by combining RT and accuracy rates (zscorePC [percent correct] – zscoreRT) into one metric, (Liesefeld & Janczyk, 2019). This metric gives equal weight to speed (RT) and accuracy. Thus, this score reflects a combination of RT and accuracy rates, and higher scores on this task indicate better combined performance. For example, a person who is 1 *SD* above the mean for accuracy (i.e., a Zscore of 1) and .5 *SD* above the mean for speed (i.e., a *z* score of $-.5$; a negative value because it reflects a value smaller than the mean) will obtain a BIS score of 1.5. A participant whose accuracy was .5 *SD* and RT was 1 *SD* above the mean would also have a BIS score of 1.5. We computed BIS for 2- and 3-back trials, the trials with highest demand on memory load.

To confirm that the task manipulations (time pressure/switching/memory load) yielded the expected effect, we examined the raw RT within tasks based on level of manipulation prior to transforming the scores. To evaluate the pressure manipulation in the Go/NoGo task, we compared RT for Go trials with no pressure manipulation to trials with the pressure manipulation. For the auditory Go/NoGo task, pressure trial responses were faster under the pressure manipulation (M = 360.24, SD = 83.28) than the no pressure trials (M = 443.48, SD = 82.55), $t(244) = 13.85$, $p < .0001$; $d = 1.00$. Similarly, pressure trials were faster (M = 305.42, SD = 43.83) than no pressure trials (M = 362.10, SD = 70.14) for the visual Go/NoGo task, $t(249) = 16.15$, $p < .0001$; $d = .97$. We did not compare Go trial RTs with NoGo trial RTs owing to the smaller number of NoGo trial RT responses. For the auditory switch task, trials with no switching were faster (M = 550.60, SD = 83.80) than trials with either a dimension switch (M = 566.98, SD = 80.26), $t(234) = 3.54$, $p < .001$, $d = .20$, or dimension and hand switch (M = 571.97, SD = 82.83), $t(234) = 4.60$, $p < .0001$; $d = .26$. Similarly, for the visual switch task, trials with no switching were faster (M = 474.95, SD = 71.40) than trials with either a dimension switch (M = 483.48, SD = 78.63), $t(253) = 3.64$, $p < .001$, $d = .11$, or dimension and hand switch (M = 480.02, SD = 77.09), $t(253) = 2.25$, $p = .025$, $d = .07$. For the auditory *n*-back task, 0-back trial responses were significantly faster (M = 812.66, SD = 274.40) than both two-back trial responses (M = 1148.36, SD = 280.35), $t(210) = 15.86$, $p < .0001$; $d = 1.21$, or three-back trial responses (M = 1128.81, SD = 299.82), $t(209) = 13.62$, $p < .0001$; $d = 1.10$. Similarly, for the visual *n*-back task, 0-back trials responses were significantly faster (M = 426.81, SD = 109.99) than both two-back trials responses (M = 564.03, SD = 215.81), $t(249) = 10.58$, $p < .0001$; $d = .80$, or three-back trial responses (M = 564.69, SD = 232.29), $t(246) = 9.94$, $p < .0001$; $d = .76$.

Electrophysiological Recoding, Processing, and ERP Measures

The continuous EEG was recorded with a lycra cap fitted with 64 Ag/AgCl sintered electrodes (Quick-Cap), amplified with a

Neuvo 64-channel amplifier, and sampled at 500 Hz (Compumedics U.S.A. Inc., Charlotte, NC). Eye movement was recoded with four electrodes placed above and below the left eye and on the outer canthus of each eye. Reference electrodes were placed on the right and left mastoid. Electrode impedance was maintained at $<$10 kΩ, and most were under 5 kΩ. After recording, the EEG data were processed offline with Curry 8 software (Compumedics U.S.A. Inc.). Offline, the EEG was rereferenced to the common average reference and filtered (high-pass filter set to .10 Hz, slope = .2; low-pass filter set to 30 Hz, slope = 6.0; 60 Hz notch filter, slope = 1.5). Eyeblinks exceeding $\pm 75 \mu V$ were corrected using the covariance method. Stimulus-locked trials (−140 to 800 ms) were then extracted from the ongoing EEG and baseline (−140 to 0 ms) corrected. The noise statistic was applied to automatically reject contaminated trials. Noise was computed over the baseline period and trials that exceeded the average noise level were automatically rejected. Only trials with correct responses were averaged together by trial type and exported for analysis.

Our Go/NoGo inhibition ERP of interest was the NoGo N2/P3 complex which shows maximum amplitude over frontal and central sites and is linked to response inhibition (Ramautar et al., 2006). Linked to nonmotor, cognitive inhibition in monolinguals (Thomas et al., 2009) and to language switching in bilinguals (Jackson et al., 2001), the N2 ERP was our measure of shifting abilities during the Switch task. Last, the P3 ERP was selected as the memory capacity measure during the *n*-back task. This ERP shows maximum amplitude over posterior sites and is linked to memory updating (Daffner et al., 2011).

We computed ERPs at two separate levels of difficulty for each task to enter in the model: The Go/NoGo task had the Pressure and No-pressure trials. The Switch task had the Switch and No-switch trials. Because the *n*-back had three trial levels, we used the 0-back for one and combined the 2- and 3-back trials into the other level. Combining the two highest levels increased the number of trials contributing to the average. Notably, the *n*-back task had the fewest number of target trials, 18 per *n*-back level.

ERP amplitudes were measured at three electrode strips: Frontal-Central Strip (FC), Central Strip (C), and Central-Parietal (CP). For each electrode strip, we averaged across three electrodes on the left, midline, and right side of the head to generate a total of nine averaged sites: FC5, FC3, FC1 (FC left); FC1, FCz, FC2 (FC midline), FC2, FC4, FC6 (FC right); C5, C3, C1 (C left); C1, Cz, C2 (C midline), C2, C4, C6 (C right); CP5, CP3, CP1 (CP left); CP1, CPz, CP2 (CP midline), CP2, CP4, CP6 (CP right). For analysis, we selected the ERP amplitudes at one of the nine (averaged) sites to enter in the model. We first checked that the ERP of interest followed the expected scalp distribution (for example, maximum amplitude toward the back of the head for the updating P3, and maximum amplitude at frontal and central sited for the inhibition N2/P3 complex and shifting N2. We then chose the site (left, midline, right) that revealed maximum group differences.

Inhibition N2/P3 ERP

We computed a difference wave by subtracting the Go trial ERP from the NoGo trial ERP, and we used this difference wave to measure the NoGo N2/P3 complex. We computed N2/P3 complex as the difference between the amplitude of the P3 and N2. We defined the N2 ERP as the most negative peak in the 150–240 ms time window after stimulus onset and the P3 as the most positive peak in the 290–410 ms time window. Time windows for all ERPs were selected based on visual inspection of grand averages. N2 and P3 amplitudes were computed by taking the mean of a 14-ms time window centered on the peak amplitude in the corresponding time window (i.e., local peak method, Luck, 2014). We used the same procedure and time windows for auditory and visual modality trials. The N2/P3 complex was measured separately for trials with and without time pressure.

Shifting N2 ERP

We measured the N2 amplitude, which we computed separately for Switch and No-switch trials. Trials in which the target dimension or the target dimension and response hand changed from the previous trial were averaged together and comprised Switch trials. The N2 was computed as the average voltage in the 200–320 ms time window for the auditory modality and in the 190- to 314-ms time window for the visual modality.

Updating P3 ERP

The P3 amplitude was computed separately for 0-back trials and for the combined 2- and 3-back trials. The *n*-back task contained the fewest number of target trials, 18 targets per condition. Because of the low number of trials, we averaged across 2- and 3-back trials to improve the signal to noise ratio in these more challenging trials. The P3 was computed as the average voltage in the 340- to 640-ms time window for the auditory modality and in the 280- to 540-ms time window for the visual modality.

Procedure

Participants who met study criteria were tested after reading and signing a consent form approved by the University's Institutional Review Board. Participants first completed the Demographic, Language, and Music questionnaires. The EEG cap was then affixed, and the six computerized EF tasks were administered. Each task started with the training trials, during which time the EEG was not recorded. After the training trials, the researcher checked with the participant that the instructions were understood and started to record the EEG. After completion of each task, the Mind Wandering questionnaire was administered. Participants took short breaks between blocks of trials during each task and after completing each task. The researcher checked EEG impedance between tasks and set the computer for the next task. After completing all EF tasks, the EEG cap was removed, and the participant completed the BVAT (order of Spanish and English was counterbalanced) and Raven's Standard Progressive Matrices. Testing took approximately 4–4.5 hr to complete, and participants were compensated for their time.

Results

Demographic Information

A total of 268 participants were tested and divided into two language groups based on self-classification as monolingual (n = 109) or bilingual (n = 152). Seven participants were excluded from the analysis because they spoke a language other than English or Spanish. Most monolinguals were born in the United States (n = 95 [88%]). More than half of the bilinguals were also born in the United States (n = 86 [56%]), others (n = 61 [40%]) were born in Central or South America. As a group, monolinguals (M = 19.03, SD = 1.71) were younger than bilinguals by one year, on average (M = 20.01, SD = 3.25), $t(259) = 2.883$, $p = .004$. However, age was not correlated with any latent variable in the subsequent analysis. The ratio of males to females was similar between groups (monolinguals 25/84, bilinguals 30/122), $\chi^2(1, N = 261) = .477$, $p = .490$.

Participants indicated household income with a checkmark next to the appropriate salary range (<5K, 5–10K, 11–20K, 21–40K, 41–60K, 61–80K, 81–100K, >100K). Between-group comparison did not reveal group differences $\chi^2(8, N = 261) = 14.152$, $p = .078$. Participants also reported highest educational attainment for each parent by checking one of the following (<high school; high school/GED; some college/AA degree/technical school; college degree; graduate degree). Neither mother nor father educational level comparison revealed between-group differences (mother $\chi^2[4, N = 259] = 4.58$, $p = .333$; father $\chi^2[4, N = 240] = 9.05,1$ p = .060).

Language Proficiency and Nonverbal Intelligence Scores

Descriptive statistics for English and Spanish BVAT and Raven's SPM can be found in Table 1. Between-group comparison on English BVAT scores revealed a Monolingual group advantage, $t(254)$ = 4.424, $p < .001$, $d = .58$. As expected, the Bilingual group scored higher than the Monolingual group on the Spanish BVAT, $t(254)$ = 20.724, $p < .001$, $d = 2.80$. A difference score between English and Spanish BVAT scores was computed as a measure of language balance. Because a smaller difference reflects more equivalency between languages or greater degree of language balance, for consistency, scores were multiplied by −1, so that higher scores reflect higher degree of balance. Between-group comparison on language balance revealed a bilingual advantage, $t(254) = 20.394$, $p < .001$, $d = 2.74$. Lastly, both groups obtained similar scores on the Raven's SPM, $t(254) = 1.684$, $p = .093$, $d = .22$.

Language Usage Questionnaire

Bilinguals acquired their second language by 7.8 years of age (SD = 5.3, range 1–23 years). Most bilinguals learned Spanish first (n = 101), a few reported learning English and Spanish simultaneously (n = 4), the remainder (n = 47) learned English as their first language. Additionally, most reported being English language dominant (n = 81), a few reported being Spanish dominant (n = 19), the remainder reported similar proficiency across both languages (n = 51). One participant did not report his dominant language.

Participants were asked to report what language or languages they spoke at home, in social settings, and in the classroom, and to report percent of time (90/10, 75/25, 50/50) they spoke each language in each setting. This information was used as a measure of language switching. Participants who reported speaking only one language were assigned a "language switching score" of 0; participants who reported 90/10 split between their languages were assigned a score of .5; those who reported a 75/25 split were assigned a value of 1.0; and those who reported 50/50 split were assigned a value of 2. Thus, higher scores reflect more language switching.

Most bilinguals reported switching at home and in social settings, but a subset (n = 17) reported no language switching. Additionally, a few monolinguals (n = 8) reported language switching, mostly in social settings, even though they reported not speaking a second language and scored low in the Spanish BVAT. Given how ubiquitous Spanish is in south Florida, this most likely reflects limited vocabulary knowledge used in informal settings.

Language switching scores were summed across settings to compute a "total language switching" score. We predicted that balanced bilinguals (i.e., those who obtained similar scores between their English and Spanish BVAT) would report more switching between their two languages. Indeed, a Pearson correlation revealed a strong relationship between language balance and language switching, $r(254) = .720$, $p < .001$.

Music Abilities

The two groups were similar in the number of participants who played a musical instrument, Monolinguals (n = 29), Bilinguals (n = 38), $\chi^2(1, N = 261) = .135$, $p = .714$.

Latent Variable Analysis

We used R (Version 3.5.1) with the Lavaan package (Rosseel, 2012) to estimate latent variables with missing data. Missing data was estimated using maximum likelihood functions in Lavaan. We report χ^2 tests for model fit, but because they may be significant with larger samples, we also report χ^2/df, and interpret values ≤ 2 as indicating adequate fit. Further, we report comparative fit indices (CFI), Tucker-Lewis indices (TLI), which ≥ .90 indicate adequate fit, and ≥ .95 indicate good fit; and root mean square error of approximation (RMSEA; with a 90% CI) and standardized root mean square residual (SRMR), which values ≤ .08 indicate adequate fit and values ≤ .05 indicate good fit (Schermelleh-Engel et al., 2003). We conducted multigroup analysis to determine whether the models tested fit for both monolinguals and bilinguals.

Three-Factor Model

Behavioral Data

We first tested a three-factor model, consistent with the Miyake et al. (2000) model, with inhibition tasks loading on a general inhibition factor, updating tasks loading on a general updating factor, and shifting tasks loading on a general shifting factor. Descriptive statistics for the tasks can be found in Table 1. As shown in Table 2, this three-factor model provided a poor fit to the data.

ERP Data

We tested the same three-factor model on the ERP data. Similar to the behavioral model findings, as seen in Table 2, this three-factor ERP model also provided a poor fit to the data. Descriptive statistics for the ERP can be found in Table 1. Table 1 shows the auditory and visual NoGo N2/P3 complex amplitude at electrode FC right (average of FC2, FC4, FC6), the auditory and visual n-back P3 amplitude at CP midline (average of CP1, CPz, CP2), and the auditory

Table 2
Fit Statistics for Latent Variable Models

Model	$\chi^2(df)$	χ^2/df	χ^2 Monolingual	χ^2 Bilingual	CFI	TLI	RMSEA [90% CI]	SRMR
Behavioral data								
Three-factor model	490.54 (102)	4.81	178.86	311.69	.691	.600	.171 [.156, .187]	.093
Six-factor model	115.24 (78)	1.48	52.26	65.98	.970	.950	.060 [.035, .083]	.050
Six-factor model collapsed across language groups	62.23 (39)	1.59	—	—	.982	.970	.048 [.023, .069]	.030
Hierarchical three-factor model	128.99 (90)	1.43	59.22	64.77	.973	.960	.054 [.027, .076]	.054
Hierarchical three-factor model collapsed across language groups	66.70 (45)	1.48	—	—	.983	.975	.043 [.018, .064]	.034
Hierarchical three-factor model with language switching	77.26 (54)	1.43	—	—	.982	.974	.041 [.017, .060]	.035
Hierarchical three-factor model with language balance	87.593 (54)	1.62	—	—	.975	.964	.049 [.029, .067]	.040
ERP data								
Three-factor model	239.490 (104)	2.30	130.45	109.05	.841	.798	.100 [.083, .117]	.102
Six-factor model	113.169 (82)	1.38	65.29	47.88	.963	.941	.054 [.026, .077]	.061
Six-factor model collapsed across language groups	65.906 (41)	1.61	—	—	.971	.953	.048 [.025, .068]	.050
Hierarchical three-factor model	135.859 (96)	1.49	74.00	61.857	.953	.936	.057 [.032, .077]	.073
Hierarchical three-factor model collapsed across language groups	78.166 (48)	1.63	—	—	.965	.951	.049 [.028, .068]	.061
Hierarchical three-factor model with language switching	78.015 (56)	1.39	—	—	.975	.965	.038 [.013, .057]	.053
Hierarchical three-factor model with language balance	84.309 (56)	1.51	—	—	.968	.955	.043 [.022, .062]	.058

Note. CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual.

and visual Switch task N2 amplitude at C left (average of C5, C3, C1) and C right (average of C2, C4, C6), respectively.

Six-Factor Model

Behavioral Data

Owing to the distinct sensory modality of the tasks within each EF construct, we tested a six-factor modality specific model, splitting each EF into a visual and auditory factor, resulting in the following latent variables: visual inhibition, auditory inhibition, visual updating, auditory updating, visual shifting, and auditory shifting. This model provided adequate fit for the data, as shown in Table 2. Factor loadings and correlations between factor scores for each group can be seen in Table 3. Correlations between observed variables entered in the confirmatory factor analysis (CFA) by group can be seen in Table 4. We examined measurement invariance to determine if the models fit the data for the monolinguals and bilinguals in the same manner. No differences were observed between groups on loadings, $\chi^2(6) = 1.82$, $p = .935$, or intercepts, $\chi^2(6) = 4.51$, $p = .608$, but a difference was observed for factor means, $\chi^2(6) = 18.52$, $p = .005$. Because no difference was observed between the models, we conducted a CFA collapsing across language groups (see Figure 1a). This allows for greater confidence and stability in the model owing to the larger sample size for the collapsed model. As shown in Table 2, this model provided a good fit for the data. To test the first goal of examining the differences between monolinguals and bilinguals on mean factor scores and our second goal of examining the impact of test modality on EF differences between language groups, we extracted factor scores for the six factors observed in the prior analysis. Unlike task group means, comparing mean factor scores should better reflect a measure of the construct than by comparing task scores which include task-specific effects (e.g., motor speed). Mean factor scores by group can be seen in Table 5, and correlations between observed variables collapsed across groups can be seen in Table 6. Factor loadings for each task were similar across demand levels, suggesting a similar contribution of each difficulty level to the latent variable. A series of *t* tests were conducted to examine differences between language groups. As shown in Table 5, a significant effect of language group was observed such that bilinguals evidenced stronger Visual and Auditory Inhibition, but weaker Visual Updating, and Visual and Auditory Shifting. Language groups did not differ on Auditory Updating. Because one of our goals was to compare groups on measures that directly captures motor response inhibition, we further examined the Go/NoGo task performance for both sensitivity and Go trial RTs. Group means and *t* tests can be found in Table 7. Stronger motor inhibition was observed in bilinguals, resulting in poorer performance on Go trials as revealed by sensitivity differences which showed a higher hit rate in the monolingual group. Further, Go trial RTs were generally similar across groups.

ERP Data

The six-factor, modality-specific model was then tested with the ERP data, splitting each EF into a visual and auditory factor. This model provided an adequate fit for the data, as shown in Table 2. Factor loadings and correlations between factor scores for each group can be seen in Table 8. Correlations between observed variables entered in the CFA by group can be seen in Table 9. As with the behavioral data, we tested the six-factor modality specific

Table 3
Factor Loadings and Interfactor Correlations for Six-Factor Model by Group for Behavioral Data

Monolinguals	Visual inhibition	Auditory inhibition	Visual updating	Auditory updating	Visual shifting	Auditory shifting
Visual no pressure	.72 (.49)					
Visual pressure	.75 (.44)					
Auditory no pressure		.63 (.60)				
Auditory pressure		.61 (.63)				
Visual 2-back			.79 (.38)			
Visual 3-back			.79 (.37)			
Auditory 2-back				.87 (.24)		
Auditory 3-back				.74 (.46)		
Visual dimension change					.96 (.08)	
Visual dimension & hand change					.89 (.22)	
Auditory dimension change						.95 (.13)
Auditory dimension & hand change						.93 (.11)
Interfactor correlations						
Visual inhibition	—					
Auditory inhibition	.77**	—				
Visual updating	−.43**	−.22	—			
Auditory updating	−.45**	−.27	.66**	—		
Visual shifting	−.44**	−.40**	.57**	.39**	—	
Auditory shifting	−.35**	−.32*	.46**	.50**	.65**	—
Bilinguals	**Visual inhibition**	**Auditory inhibition**	**Visual updating**	**Auditory updating**	**Visual shifting**	**Auditory shifting**
Visual no pressure	.90 (.18)					
Visual pressure	.63 (.60)					
Auditory no pressure		.89 (.21)				
Auditory pressure		.42 (.82)				
Visual 2-back			.65 (.58)			
Visual 3-back			.67 (.55)			
Auditory 2-back				.87 (.24)		
Auditory 3-back				.70 (.51)		
Visual dimension change					.99 (.01)	
Visual dimension & hand change					.88 (.22)	
Auditory dimension change						1.00 (−.003)
Auditory dimension & hand change						.90 (.19)
Interfactor correlations						
Visual inhibition	—					
Auditory inhibition	.47**	—				
Visual updating	−.24	−.09	—			
Auditory updating	−.23*	−.10	.57**	—		
Visual shifting	−.37**	−.38**	.25*	.34**	—	
Auditory shifting	−.33**	−.33*	.25*	.22*	.45**	—

Note. The loading for the Auditory dimension change variable is greater than 1 and the estimated error variance is negative. These parameters are maximum likelihood estimates with standard errors around those estimates, and the theoretical maximum loading and minimum error variance are both within the standard error of these estimates. Fixing the loadings of the path from Auditory dimension change to the Auditory latent variable did not change the model fit or the correlations between the latent variables, so we allowed these parameters to be freely estimated.
* $p < .05$. ** $p < .01$.

model collapsing across groups. As shown in Table 2, this model provided a slight improvement in model fit based on CFI, TLI, RMSEA, and SRMR scores. A test of differences between the grouped and collapsed model, based on changes in χ^2 did not reach statistical significance, $\chi^2(48) = 57.693$, $p = .159$, but based on parsimony and improvements in other fit statistics, the collapsed model was used to extract factor scores to compare groups (see Figure 1b). Mean factor scores by group can be seen in Table 10, and correlations between observed variables collapsed across groups can be seen in Table 11. To test our first goal of examining language group differences on ERP measures and our second goal of examining differences in test modality of EF abilities between language groups, we conducted a series of *t* tests. As shown in Table 10, a significant effect of language group was observed such that bilinguals evidenced stronger auditory inhibition and visual shifting. Group differences favored bilinguals for visual inhibition but did not reach statistical significance ($p = .054$). Monolinguals showed stronger visual but not auditory updating.

Hierarchical Three-Factor Model

Behavioral Data

Based on the prior work by Miyake et al. (2000), which served as the foundation for this study, we tested a hierarchical model in which the six modality-specific factors load onto three latent variables, see Figure 2a. As seen in Table 2, this model provided adequate fit for the data. We then compared the fit statistics, AIC and BIC, of the hierarchical three-factor model with the six-factor modality specific model. The six-factor modality specific model had slightly poorer fit (AIC = 7485.350, BIC = 7848.539) compared

Table 4

Correlations Between Observed Variables in CFA Models, Monolingual (Bottom of Diagonal) and Bilingual (Top of Diagonal) for Behavioral Data

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. Visual Go/NoGo no pressure	—	.57***	.38***	.14	−.13	−.13	−.19*	−.14	−.32***	−.32***	−.31***	−.29***
2. Visual Go/NoGo pressure	.53***	—	.24**	.35***	−.18*	−.07	−.09	−.20*	−.22*	−.27**	−.16	−.15
3. Auditory Go/NoGo no pressure	.43***	.32**	—	.37***	−.06	−.04	−.09	.05	−.33***	−.26**	−.31***	−.26**
4. Auditory Go/NoGo pressure trials	.27**	.36***	.38***	—	−.20*	.02	−.10	−.12	−.23**	−.26**	−.05	−.10
5. Visual 2-back	−.14	−.24*	−.02	−.17	—	.43***	.31***	.22**	.24**	.21*	.22*	.20*
6. Visual 3-back	−.25*	−.26*	−.05	−.20	.63***	—	.34***	.27**	.11	.07	.14	.13
7. Auditory 2-back	−.31**	−.29**	−.14	−.16	.49***	.41***	—	.61***	.29***	.25**	.21*	.23**
8. Auditory 3-back	−.22*	−.14	−.17	−.12	.42***	.47***	.65***	—	.24**	.29**	.13	.18*
9. Visual dimension change	−.17	−.31**	−.17	−.39***	.30**	.32**	.31**	.10	—	.88***	.44***	.39***
10. Visual dimension & hand change	−.14	−.36***	−.18	−.29**	.24*	.36***	.27*	.13	.83***	—	.38***	.35***
11. Auditory dimension change	−.14	−.30**	−.11	−.33**	.27*	.25*	.41***	.17	.57***	.50***	—	.90***
12. Auditory dimension & hand change	−.08	−.21*	−.09	−.24*	.27*	.24*	.41***	.30**	.53***	.50***	.87***	—

Note. CFA = confirmatory factor analysis.
* $p < .05$. ** $p < .01$. *** $p < .001$.

with the hierarchical three-factor model (AIC = 7470.098, BIC = 7790.560). We retained the hierarchical model for the remaining analyses. Factor loadings by language group can be seen in Table 12.

As in the six-factor modality specific model, a measurement invariance analysis revealed no difference based on loadings, $\chi^2(9) = 6.76$, $p = .662$, or intercepts, $\chi^2(3) = 2.97$, $p = .397$, but again, a difference was observed for factor means, $\chi^2(9) = 18.47$, $p = .030$. Based on the lack of difference between models on loadings and intercepts, we conducted a hierarchical CFA collapsing across groups to extract factor scores for inhibition, shifting, and updating, see Figure 2a. This collapsed model provided a good fit for the data and the best fit of any model tested (AIC = 7447.11, BIC = 7607.34; see Table 2). To investigate language-group differences, we extracted factor scores for the inhibition, shifting, and updating factors. As shown in Table 5, inhibition was stronger in bilinguals whereas updating and shifting were stronger in monolinguals. To ensure that differences observed between language groups were not attributable to age, SES, parental education, or fluid intelligence, we conducted a series of ANCOVA with these factors included as covariates. The inclusion of these covariates into the analyses examining group performance differences on the inhibition, shifting, and updating factors did not alter the findings.

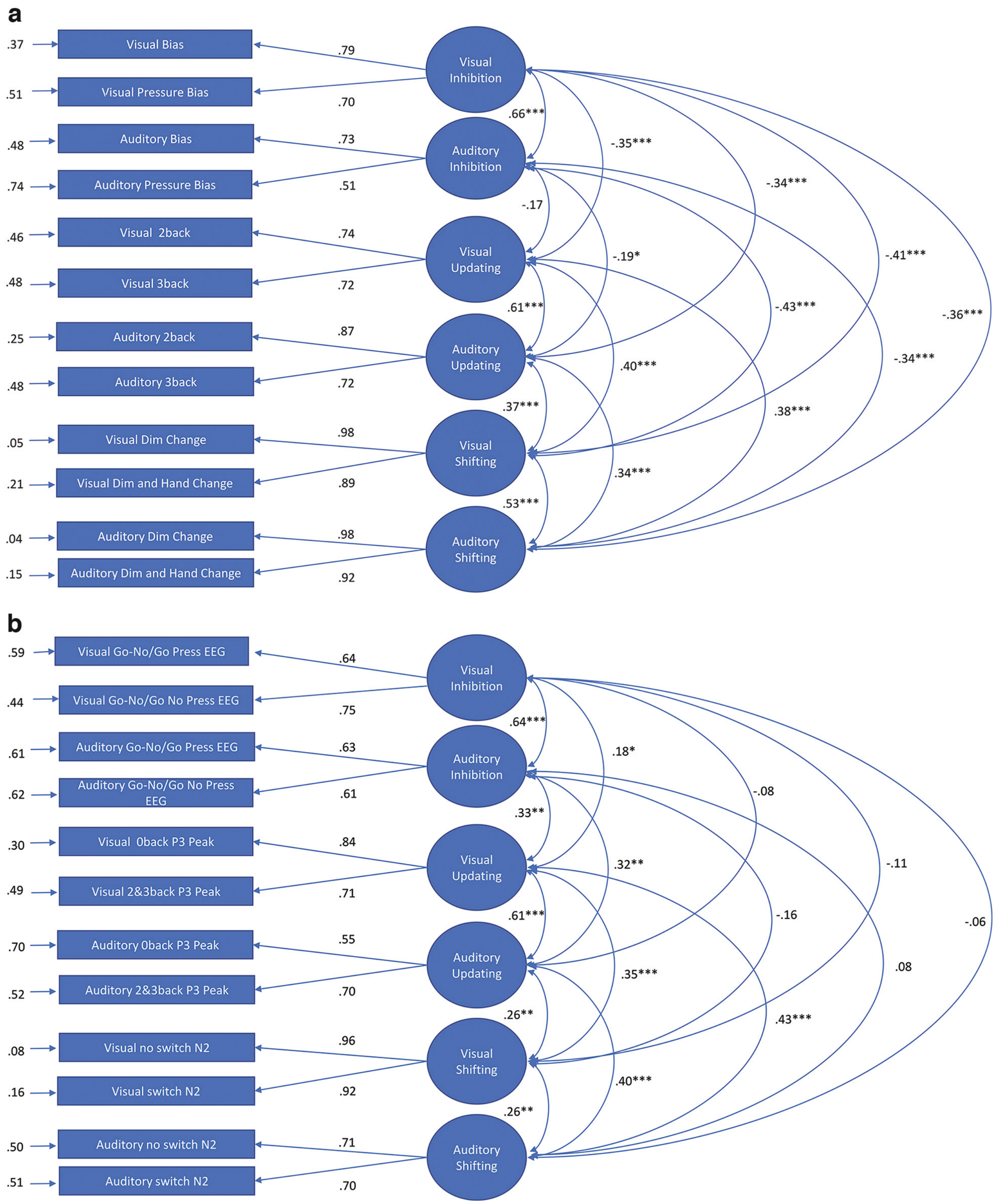
We then examined whether the relationship between EF components was different between groups. To determine whether the correlation between the latent variables differed between monolinguals and bilinguals, we conducted a CFA examining the hierarchical three-factor model by language condition, constraining the relationships between the latent variables to be equal across groups. This model produced a good fit, $\chi^2(99) = 140.17$, $p = .004$, $\chi^2/df = 1.42$, CFI = .967, TLI = .956, RMSEA = .057, 90% CI [.033, .077], SRMR = .064. A comparison of the model with constraints on relationships between the latent variables and a model with no constraints revealed no significant difference between the models using a chi-square difference test, $\chi^2(3) = 7.02$, $p = .071$. The lack of a significant difference between the models indicates that the relationships between the latent variables did not differ between monolinguals and bilinguals.

ERP Data

We followed the same procedure with the ERP data and tested a hierarchical model in which the six modality-specific factors load onto three latent variables, see Figure 2b. As shown in Table 2, similar to the behavioral data model, the ERP model provided adequate fit for the data. We then compared the fit statistics, AIC and BIC, of the hierarchical three-factor model with the six-factor modality specific model. The six-factor modality specific model had slightly poorer fit (AIC = 12714.664, BIC = 13063.607) compared with the hierarchical three-factor model (AIC = 12688.486, BIC = 12994.705). We retained the hierarchical model for the remaining analyses. Factor loadings by language group can be seen in Table 13.

As we did before with the behavioral data, we tested a hierarchical three-factor model collapsing across groups. As shown in Table 2, this model provided a slight improvement in model fit based on CFI, TLI, RMSEA, and SRMR scores compared with the hierarchical three-factor model split by group. This collapsed model

Figure 1
Six-Factor Model



Note. Six-factor modality–specific CFA model collapsed across language groups for (a) behavioral data, and (b) ERP data. CFA = confirmatory factor analysis; ERP = event-related brain potential. See the online article for the color version of this figure.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 5

Descriptive and Inferential Statistics for Latent Variables in CFA Models by Group for Behavioral Data

Variable	Monolinguals				Bilinguals							
	M	*SD*	Skewness	Kurtosis	*M*	*SD*	Skewness	Kurtosis	*t*	*p*	*d*	*n*
Six-factor model												
Visual inhibition	−0.14	0.70	0.22	−0.67	0.10	0.66	−0.15	−1.05	2.72	.007	.35	259
Auditory inhibition	−0.12	0.55	0.02	−0.73	0.08	0.60	0.05	−0.31	2.72	.007	.35	259
Visual updating	0.11	0.84	−1.53	6.55	−0.08	0.68	−0.54	0.54	2.07	.039	.25	259
Auditory updating	0.10	1.15	−0.59	0.78	−0.07	1.03	−0.80	−0.92	1.27	.206	.03	259
Visual shifting	0.18	0.83	−0.31	−0.42	−0.13	0.96	−0.63	−0.10	2.67	.0,004	.35	259
Auditory shifting	0.22	0.77	0.30	−0.16	−0.16	0.91	0.12	−0.54	3.54	.0,004	.45	259
Hierarchical three-factor model												
Inhibition	−0.12	0.45	0.24	−0.67	0.08	0.46	−0.09	−0.99	3.08	.002	.44	259
Updating	0.10	0.65	−1.01	3.07	−0.07	0.54	−0.55	0.33	2.35	.019	.28	259
Shifting	0.17	0.61	−0.32	0.30	−0.12	0.63	−0.13	−0.37	3.46	.0,006	.47	259

Note. CFA = confirmatory factor analysis.

provided good fit for the data (see Table 2). To investigate language-group differences, we extracted factor scores for the inhibition, shifting, and updating factors. As shown in Table 10, inhibition (larger N2/P3 complex) and shifting (more negative N2 ERP) were stronger in bilinguals whereas updating (larger P3) was not different between the groups. To ensure that differences observed between language groups were not due to age, SES, parental education, or fluid intelligence, we conducted a series of ANCOVA with these factors included as covariates. The inclusion of these covariates into the analyses examining group differences on the inhibition, shifting, and updating factors did not alter the findings.

We then examined whether the relationship between EF components was different between groups. To determine whether the correlation between the latent variables differed between monolinguals and bilinguals, we conducted a CFA examining the hierarchical three-factor model by language condition, constraining the relationships between the latent variables to be equal across groups. This model produced a good fit, $\chi^2(97) = 113.57$, $p = .120$, $\chi^2/df = 1.17$, CFI = .981, TLI = .974, RMSEA = .036, 90% CI [.000, .031], SRMR = .066. A comparison of the model with constraints on relationships between the latent variables and a model with no constraints revealed no significant difference between the models using a chi-square difference test, $\chi^2(47) = 51.84$, $p = .291$. The lack of a significant difference between the models indicates that the relationships between the latent variables did not differ between monolinguals and bilinguals.

Language Switching and Language Balance

Behavioral Data

To test the hypothesis that the context in which bilinguals use their two languages impacts EF abilities, we used the total language switching score (calculated from participants self-reported language switching frequency) and language balance scores (calculated as the difference between Spanish BVAT scores and English BVAT scores) to predict the three EF latent variables from the hierarchical model collapsed across language groups. Specifically, we tested a series of models in which language switching and language balance separately predicted inhibition, shifting, and updating. We hypothesized that language balance would better predict EF measures than language switching.

For the language switching analysis, a single indicator latent variable was created from the language switching variable. As can be seen in Table 2, the model provided a good fit to the data. Figure 3 shows that language switching frequency was correlated to inhibition and shifting, but not updating.

Table 6

Correlations Between Observed Variables in CFA Models Collapsed Across Groups for Behavioral Data

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. Visual Go/NoGo no pressure bias	—											
2. Visual Go/NoGo pressure bias	.55***	—										
3. Auditory Go/NoGo no pressure bias	.41***	.28***	—									
4. Auditory Go/NoGo pressure bias	.18**	.35***	.37***	—								
5. Visual 2-back trials trade-off scores	−.16*	−.21**	−.06	−.19**	—							
6. Visual 3-back trade-off scores	−.20**	−.16*	−.06	−.06	.53***	—						
7. Auditory 2-back trade-off scores	−.25***	−.18**	−.12	−.12	.39***	.38***	—					
8. Auditory 3-back trade-off scores	−.18**	.18**	−.05	−.12	.32***	.37***	.63***	—				
9. Visual dimension change switch cost	−.29***	−.26***	−.30***	−.28***	.28***	.20**	.31***	.20**	—			
10. Visual dimension & hand change switch cost	−.28***	−.31***	−.26***	−.26***	.24***	.19**	.28***	.24***	.87***	—		
11. Auditory dimension change switch cost	−.28***	−.22***	−.27***	−.14*	.26***	.20**	.29***	.16***	.50***	.45***	—	
12. Auditory dimension & hand change switch cost	−.25***	−.19**	−.24***	−.14*	.26***	.20**	.31***	.23**	.46***	.43***	.90***	—

Note. CFA = confirmatory factor analysis.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 7

Go/NoGo Task Sensitivity and RTs Measures by Group

Measure	Monolinguals *M* (*SD*)	Bilinguals *M* (*SD*)	*t* Test
Sensitivity			
Auditory no pressure trials	4.98 (1.44)	4.59 (1.56)	$t(243) = 1.96$, $p = .051$
Auditory pressure trials	5.15 (2.00)	4.39 (2.02)	$t(243) = 2.87$, $p = .005$
Visual no pressure trials	7.40 (1.62)	6.89 (1.73)	$t(248) = 2.33$, $p = .021$
Visual pressure trials	6.75 (1.46)	6.36 (1.46)	$t(248) = 2.08$, $p = .039$
Go trial RT			
Auditory no pressure trials	434.89 (77.08)	449.40 (85.89)	$t(243) = 1.35$, $p = .177$
Auditory pressure trials	314.65 (33.18)	306.56 (28.84)	$t(243) = 2.02$, $p = .044$
Visual no pressure trials	356.21 (70.40)	366.29 (69.89)	$t(248) = 1.12$, $p = .263$
Visual pressure trials	292.43 (33.86)	290.18 (28.52)	$t(248) = 0.57$, $p = .571$

Note. RT = reaction time.

Next, we used language balance scores to predict inhibition, shifting, and updating. The model provided a good fit to the data (see Table 2). As shown in Figure 4, language balance was related to inhibition and shifting, but not updating.

Owing to the strong relationship between language balance and language switching, $r(254) = .720$, $p < .001$, and the similar relationships between these constructs and the EF latent variables, we examined the unique associations that language balance and

Table 8

Factor Loadings and Interfactor Correlations for Six-Factor Model by Group for ERP Data

Monolinguals	Visual inhibition	Auditory inhibition	Visual updating	Auditory updating	Visual shifting	Auditory shifting
Visual Go/NoGo pressure	.48 (.13)					
Visual Go/NoGo no pressure	.96 (.21)					
Auditory Go/NoGo pressure		.62 (.09)				
Auditory Go/NoGo no pressure		.68 (.09)				
Visual 0-back			.90 (.07)			
Visual 2 & 3-back combined			.70 (.07)			
Auditory 0-back				.84 (.07)		
Auditory 2- and 3-back combined				.63 (.05)		
Visual no-switch					.99 (.02)	
Visual switch					.90 (.02)	
Auditory no-switch						.84 (.09)
Auditory switch						.83 (.09)
Interfactor correlations						
Visual inhibition	—					
Auditory inhibition	.60**	—				
Visual updating	−.20	−.50***	—			
Auditory updating	−.04	−.30*	.56***	—		
Visual shifting	−.03	−.13	.39***	.36**	—	
Auditory shifting	−.11	−.07	.39***	.32**	.33**	—
Bilinguals	**Visual inhibition**	**Auditory inhibition**	**Visual updating**	**Auditory updating**	**Visual shifting**	**Auditory shifting**
Visual Go/NoGo pressure	.78 (.10)					
Visual Go/NoGo no pressure	.67 (.10)					
Auditory Go/NoGo pressure		.50 (.11)				
Auditory Go/NoGo no pressure		.67 (.13)				
Visual 0-back			.76 (.08)			
Visual 2- and 3-back combined			.75 (.08)			
Auditory 0-back				.46 (.08)		
Auditory 2- and 3-back combined				.55 (.09)		
Visual no-switch					.94 (.02)	
Visual switch					.93 (.02)	
Auditory no-switch						.87 (.20)
Auditory switch						.47 (.13)
Interfactor correlations						
Visual inhibition	—					
Auditory inhibition	.53***	—				
Visual updating	.17	.28*	—			
Auditory updating	−.32	.39*	.69***	—		
Visual shifting	−.08	−.001	.27**	.19	—	
Auditory shifting	−.14	.08	.39**	.56**	.10	—

Note. ERP = event-related brain potential.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 9

Correlations Between Observed Variables in CFA Models, Monolingual (on Bottom of Diagonal) and Bilingual (on Top of Diagonal) for ERP Data

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. Visual Go/NoGo pressure	—	.51***	.19*	.24**	.05	.14	−.08	−.17*	−.10	.03	−.19*	−.02
2. Visual Go/NoGo no pressure	.46***	—	.20*	.29***	.15	.11	−.09	−.05	−.10	−.05	.01	.09
3. Auditory Go/NoGo pressure	.38***	.38***	—	.34***	.08	.15	.12	.02	−.15	−.13	.06	.05
4. Auditory Go/NoGo no pressure	.12	.38***	.43***	—	.15	.12	.19*	.14	.05	.09	.02	.09
5. Visual 0-back	.12	.14	.23*	.35***	—	.57***	.36***	.21*	.18*	.18*	.33***	.21*
6. Visual 2- and 3-back combined	−.01	.19	.18	.34***	.64***	—	.22*	.27***	.24**	.24**	.18*	.14
7. Auditory 0-back	.13	.10	.19	.25*	.61***	.48***	—	.26**	.14	.09	.26**	.18
8. Auditory 2- and 3-back combined	.08	.00	.17	.11	.35***	.23*	.59***	—	.10	.07	.21*	.11
9. Visual no-switch	−.08	−.04	−.11	−.07	.34***	.29**	.23*	.32**	—	.87***	.11	.13
10. Visual switch	−.09	−.02	−.12	−.08	.29**	.27**	.10	.24*	.89***	—	.03	.14
11. Auditory no-switch	.06	−.10	.01	.04	.24*	.24*	.19	.18	.27**	.16	—	.42***
12. Auditory switch	.10	−.05	.01	.02	.34***	.16	.20	.27**	.26**	.22*	.65***	—

Note. CFA = confirmatory factor analysis.
* $p < .05$. ** $p < .01$. *** $p < .001$.

language switching had with inhibition and shifting. We extracted factor scores for inhibition and shifting from the hierarchical three-factor model collapsed across language group. Two regression analyses were conducted with language balance and language switching entered into the model simultaneously as predictors of inhibition and then shifting. The first regression—predicting inhibition—provided a significant model, $F(2, 252) = 12.34$, $p < .0001$; adjusted $R^2 = .08$, with only language balance serving as a significant predictor, $\beta = .314$, $t = 3.64$, $p = .0003$. Similarly, a significant model was found predicting shifting, $F(2, 252) = 13.14$, $p < .0001$; adjusted $R^2 = .09$, with only language balance serving as a significant predictor, $\beta = -.35$, $t = 4.12$, $p < .0001$. To ensure that the impact of language balance on these constructs was not attributable to the age of second language acquisition, we examined the correlation between these factors. The correlation between age at second language acquisition and language balance was nonsignificant, $r(153) = .14$, $p = .081$, suggesting that age of acquisition did not serve as a confounding variable.

ERP Data

We replicated the analyses examining the impact of language switching and language balance on EF abilities, examining the impact on ERP measures. As shown in Table 2, the models with both language shifting and language balance provided good fits to the data. As shown in Figure 5, language switching was related to updating and shifting but not inhibition. Similarly, as shown in Figure 6, language balance was related to updating and shifting but not inhibition.

Consistent with the behavioral models, we extracted factor scores for the ERP measures of inhibition, shifting, and updating from the hierarchical three-factor model collapsed across language groups. Two regression analyses were conducted with language balance and language switching entered into the model simultaneously as predictors of updating and then shifting. The first regression—predicting updating—provided a significant model, $F(2, 252) = 4.00$, $p = .019$, adjusted $R^2 = .02$, but neither of the predictors was significant, suggesting the shared variance between language balance and shifting is related to updating. The second regression—predicting shifting—provided a significant model, $F(2, 252) = 7.38$, $p < .001$, adjusted $R^2 = .05$, with only language balance serving as a significant predictor, $\beta = -.21$, $t = 2.39$, $p = .019$.

Discussion

The first goal of this study was to investigate neural and behavioral differences in EF between bilinguals and monolinguals at the latent level. As predicted, results revealed higher ERP amplitude in bilinguals during inhibition and shifting tasks, see Figure 7. Behaviorally, however, bilinguals performed more poorly than monolinguals on these tasks. Neural activity did not distinguish the groups on WM and performance favored monolinguals. Together, these findings do not support the claim that bilingualism enhances EF abilities. A second goal of this study was to reveal the link between language balance and EF abilities. Our findings linked neural activity during inhibition and shifting tasks to language balance

Table 10
Descriptive and Inferential Statistics for Latent Variables in CFA Models by Group for ERP Data

Variable	Monolinguals				Bilinguals							
	M	*SD*	Skewness	Kurtosis	*M*	*SD*	Skewness	Kurtosis	*t*	*p*	*d*	*n*
Six-factor model												
Visual inhibition	−0.26	1.64	0.18	0.85	0.16	1.82	0.52	0.62	−1.94	.054	.24	259
Auditory inhibition	−0.25	1.18	0.29	0.07	0.15	1.10	0.70	0.33	−2.81	.005	.35	259
Visual updating	0.46	3.42	0.28	−0.30	−0.31	2.80	0.59	0.77	1.99	.047	.25	259
Auditory updating	0.07	1.83	0.71	2.70	−0.05	1.39	0.25	0.85	0.59	.558	.07	259
Visual shifting	0.36	1.52	0.53	0.47	−0.23	1.42	0.84	3.72	3.18	.002	.40	259
Auditory shifting	0.10	1.12	−0.63	2.57	−0.06	0.94	0.25	0.69	1.24	.215	.16	259
Hierarchical three-factor model												
Inhibition ERP	−0.23	1.14	0.27	0.07	0.15	1.06	0.67	0.31	−2.75	.006	.35	259
Updating ERP	0.31	2.88	0.39	−0.12	−0.21	2.28	0.69	1.06	1.64	.103	.21	259
Shifting ERP	0.13	0.58	0.60	0.71	−0.09	0.48	0.51	0.61	3.29	.001	.41	259

Note. CFA = confirmatory factor analysis. ERP = event-related brain potential.

such that more balanced bilinguals demonstrated greater neural activity than less balanced bilinguals and monolinguals. These results support the claim that linguistic and nonlinguistic behaviors in bilinguals share a common inhibitory control mechanism. The third goal of this study was to investigate whether executive control in bilinguals is modality specific. SEM revealed that although the tasks contributed modality specific variance, inhibitory control is independent of test modality.

Inhibition

Our study revealed greater neural (N2/P3 amplitude) and behavioral (response bias) inhibition in bilinguals on the Go/NoGo task. Moreover, response bias was positively correlated with language balance. Thus, the most balanced bilinguals demonstrated the strongest tendency toward response suppression. Because our goal was to compare groups on a measure that *directly* captures inhibition of a motor response, we used the Go/NoGo task. We computed response bias, a measure of the extent to which one response tendency is more probable than another (e.g., a more probable NoGo response than a Go response). Ultimately, stronger motor inhibition in bilinguals led to poorer performance on Go trials as revealed by sensitivity differences which showed a higher hit rate in the monolingual group. Unlike bias and sensitivity, Go trial RTs were generally similar across groups, suggesting that the observed differences in performance were not driven by general differences in speed but rather a greater response suppression tendency.

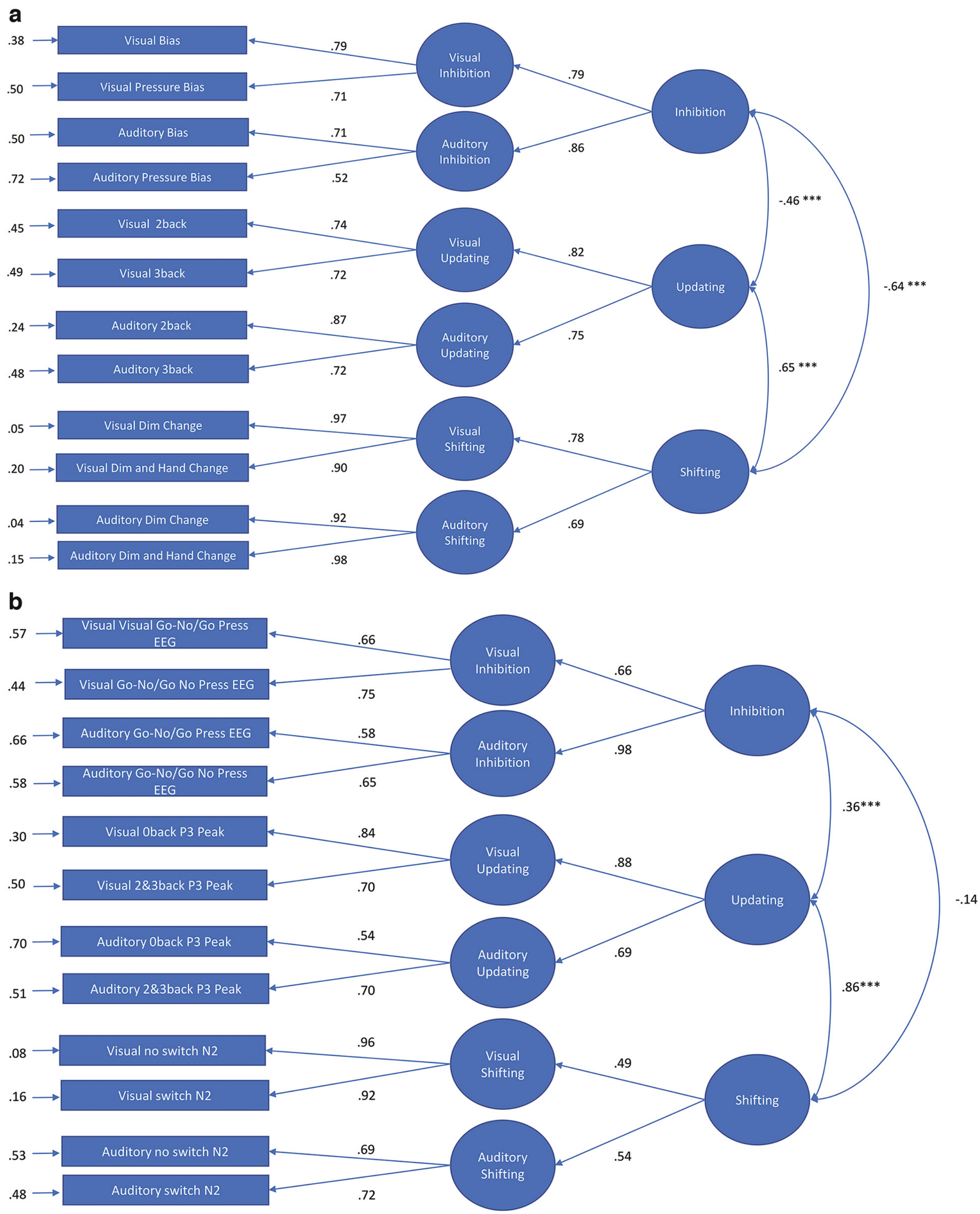
We know of one other study (Shulley & Shake, 2016) which measured response bias, but unlike the current study, language groups performed similarly. Differences between their results and ours may be attributable to differences in methodology. Half the trials in our tasks were performed under time pressure. Notably, we used the same task on two prior studies without time pressure trials and bilinguals and monolinguals performed equally well. Additionally, linguistic proficiency in their study was based on a self-report measure, and it is unknown how precisely self-reports measures capture second language proficiency and language balance. This is important because our study showed a positive correlation between language balance and response inhibition. Bilinguals in their study may not have been as balanced in their two languages as bilinguals in the current study and therefore inhibitory control was not as developed in their participants. Noteworthy, Spanish is ubiquitous in south Florida and most Spanish/English bilinguals speak and hear their two languages frequently. Indeed, our bilinguals reported speaking Spanish and English regularly. Their bilingual population was very diverse in the languages they spoke. In addition

Table 11
Correlations Between Observed Variables in CFA Models Collapsed Across Groups for ERP Data

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. Visual Go/NoGo pressure N2/P3 complex	—											
2. Visual Go/NoGo no pressure N2/P3 complex	.47***	—										
3. Auditory Go/NoGo pressure N2/P3 complex	.30***	.27***	—									
4. Auditory Go/NoGo no pressure N2/P3 complex	.21**	.33***	.39***	—								
5. Visual 0-back P3 amplitude	.04	.13*	.09	.23***	—							
6. Visual 2- and 3-back combined P3 amplitude	.07	.14*	.14*	.21***	.59***	—						
7. Auditory 0-back P3 amplitude	−.01	−.02	.11	.20**	.49***	.33***	—					
8. Auditory 3-back trade-off scores	−.05	−.04	.11	.13	.26***	.24***	.42***	—				
9. Visual dimension change switch cost	−.12	−.08	−.17**	−.03	.28***	.26***	.19**	.17**	—			
10. Visual dimension & hand change switch cost	−.04	−.04	−.16*	.00	.25***	.26***	.11	.14*	.88***	—		
11. Auditory dimension change switch cost	−.08	−.04	.05	.03	.28***	.20**	.23**	.21**	.16*	.08	—	
12. Auditory dimension & hand change switch cost	−.03	.03	−.02	.03	.29***	.16*	.17*	.17*	.21***	.19**	.47***	—

Note. CFA = confirmatory factor analysis.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Figure 2
Hierarchical Three-Factor Model



Note. Hierarchical three-factor CFA model collapsed across language groups for (a) behavioral data, and (b) ERP data. CFA = confirmatory factor analysis; ERP = event-related brain potential. See the online article for the color version of this figure.
*** $p < .001$.

Table 12
Factor Loadings (With Standard Errors) and Interfactor Correlations for Hierarchical Three-Factor Model by Group for Behavioral Data

Monolinguals	Inhibition	Updating	Shifting
Visual inhibition	1.00 (.21)		
Auditory inhibition	.76 (.17)		
Visual updating		.86 (.12)	
Auditory updating		.77 (.10)	
Visual shifting			.82 (.09)
Auditory shifting			.80 (.09)
Interfactor correlations			
Inhibition	—		
Updating	−.52***	—	
Shifting	−.52***	.74***	—
Bilinguals	**Inhibition**	**Updating**	**Shifting**
Visual inhibition	.71 (.15)		
Auditory inhibition	.70 (.12)		
Visual updating		.71 (.16)	
Auditory updating		.78 (.17)	
Visual shifting			.73 (.09)
Auditory shifting			.62 (.08)
Interfactor correlations			
Inhibition	—		
Updating	−.33*	—	
Shifting	−.76***	.54***	—

* $p < .05$. *** $p < .001$.

to English, their bilinguals spoke Arabic, Bosnian, Chinese or Mandarin, Farsi, Haitian Creole, Korean, Portuguese, Spanish, Swahili, Swedish, Taiwanese, and Twi. It is unknown whether bilinguals in their study were exposed to or had the opportunity to engage their two languages as frequently as bilinguals in our study. Either one or both factors may account for the different results.

In general, few studies to date have used ERP measures to investigate EF in bilinguals. Studies using the Go/NoGo task, unlike studies using other inhibition tasks (Stroop, Simon, flanker), consistently reveal greater neural inhibition (higher N2 ERP amplitude) in bilinguals (for a review, see Cespón & Carreiras, 2020). One plausible explanation for these language group differences is that the control mechanism that suppresses one action in favor of another action, for instance, suppressing the action to go in favor of the action to not go, is the same mechanism that suppresses one language when a bilingual speaks in their other language. Because bilinguals engage this control mechanism more often than monolinguals, response inhibition is more developed in bilinguals.

Not all researchers, however, agree that the NoGo N2 ERP represents response inhibition. Indeed, Nieuwenhuis et al. (2003) propose that the NoGo N2 reflects conflict between two competing responses. They explain that the NoGo N2 shares similarities with the N2 ERP elicited by tasks such as the flanker, which is associated with response conflict (Nieuwenhuis et al., 2003, p. 18). We propose that if the NoGo N2 reflects conflict monitoring, then bilinguals should show a similar neural signature on other conflict monitoring tasks (flanker, Stroop, Simon). As stated above, the Go/NoGo task, but not other tasks, consistently elicits greater neural activity in bilinguals and distinguishes language groups.

Costa et al. (2009) argue that when the frequency of congruent and incongruent trials is similar, this increases monitoring demands and reveals language group differences favoring bilinguals. Our findings do not support this hypothesis. We presented an equal number of Go and NoGo trials, but our bilinguals showed relatively weaker performance because they had more omission errors on Go trials than the monolingual group. Thus, our findings do not support the claim that increasing monitoring demands reveals superior bilingual performance.

Consistent with our hypothesis, one study presented equiprobable stimuli and manipulated level of conflict on three tasks, Go/NoGo, simple RT, and a Stroop-like task (Gonzalez-Rosa et al., 2013). Based on results of three different measures, behavioral, ERP, and fMRI, the authors concluded that the frontal NoGo N2 and NoGo P3 ERP reflect response inhibition rather than response conflict.

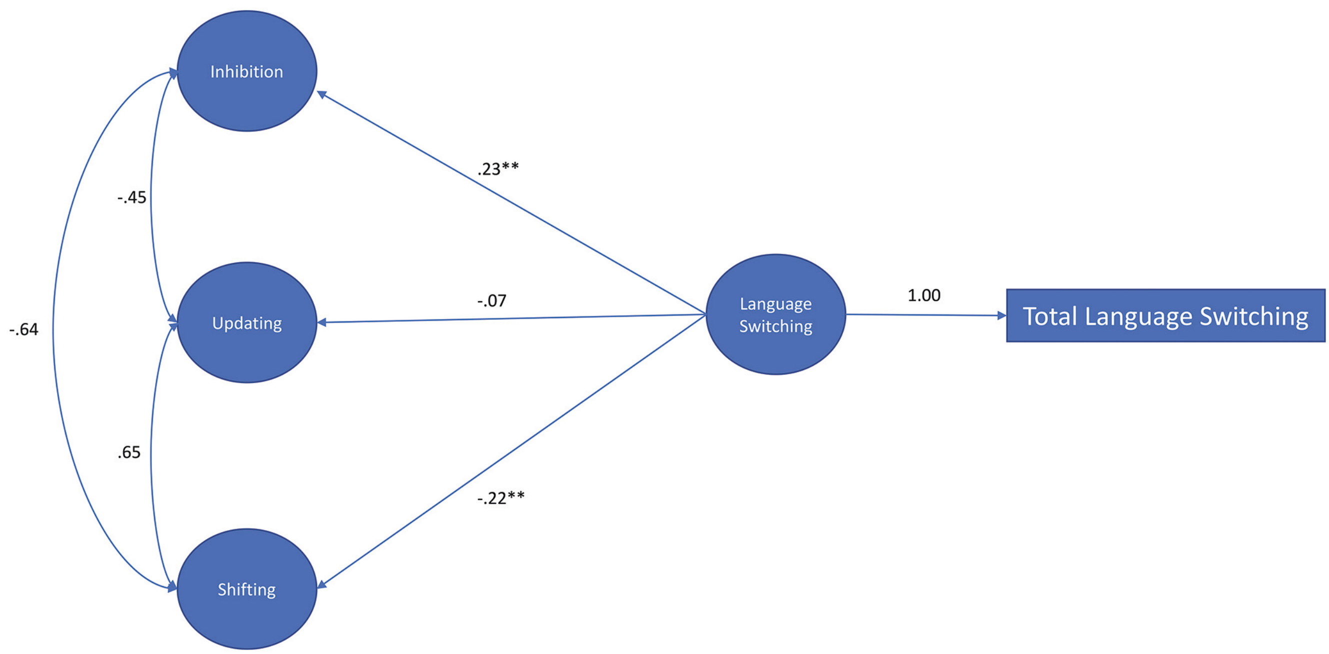
Summarily, our findings reveal both greater neural activity and greater tendency toward response suppression in bilinguals and link linguistic to nonlinguistic behaviors. Our results support the theory that the ability to control language as well as other behaviors is carried out in the frontal lobes by inhibitory processes, (Abutalebi et al., 2012). However, our results do not support the hypothesis that the experience of controlling languages leads to a bilingual advantage (i.e., better performance) on

Table 13
Factor Loadings (With Standard Errors) and Interfactor Correlations for Hierarchical Three-Factor Model by Group for ERP Data

Monolinguals	Inhibition	Updating	Shifting
Visual inhibition	.75 (.12)***		
Auditory inhibition	.95 (.14)***		
Visual updating		.86 (.11)***	
Auditory updating		.67 (.12)***	
Visual shifting			.54 (.12)***
Auditory shifting			.62 (.12)***
Interfactor correlations			
Inhibition	—		
Updating	.49**	—	
Shifting	−.12	.79***	—
Bilinguals	**Inhibition**	**Updating**	**Shifting**
Visual inhibition	.52 (.10)***		
Auditory inhibition	1.00 (.20)***		
Visual updating		.81 (.13)***	
Auditory updating		.85 (.20)***	
Visual shifting			.55 (.21)**
Auditory shifting			.25 (.12)*
Interfactor correlations			
Inhibition	—		
Updating	.44***	—	
Shifting	−.06	.87***	—

Note. ERP = event-related brain potential. The loading for the Auditory inhibition variable is greater than 1 and the estimated error variance is negative. These parameters are maximum likelihood estimates with standard errors around those estimates, and the theoretical maximum loading and minimum error variance are both within the standard error of these estimates. Fixing the loadings of the path from Auditory dimension change to the Auditory latent variable did not change the model fit or the correlations between the latent variables, so we allowed these parameters to be freely estimate.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Figure 3
SEM Predicting Inhibition, Updating, and Shifting From Language Switching for Behavioral Data



Note. SEM = structural equations modeling. See the online article for the color version of this figure.
** $p < .01$.

nonlinguistic tests of EF. Although we found higher neural activity and greater tendency toward response suppression in bilinguals, stronger inhibition resulted in poorer performance. Our findings support our contention that neural inhibition develops in tandem with second language proficiency to control increasing competition between languages. It may be that when deployed to control nonlinguistic behaviors, such as on the Go/NoGo task, behavioral responses are sometimes mistakenly inhibited. This explains in our study why bilinguals generally exhibited stronger response bias and weaker performance, with balanced bilinguals exhibiting the strongest bias and weakest performance.

Shifting

Similar to the findings of the Go/NoGo task, bilinguals showed higher switch costs and greater amplitude on our switch task ERP marker. Our behavioral findings are consistent with most other studies which reveal that bilinguals either perform similarly or worse than monolinguals (for a review see Paap et al., 2017).

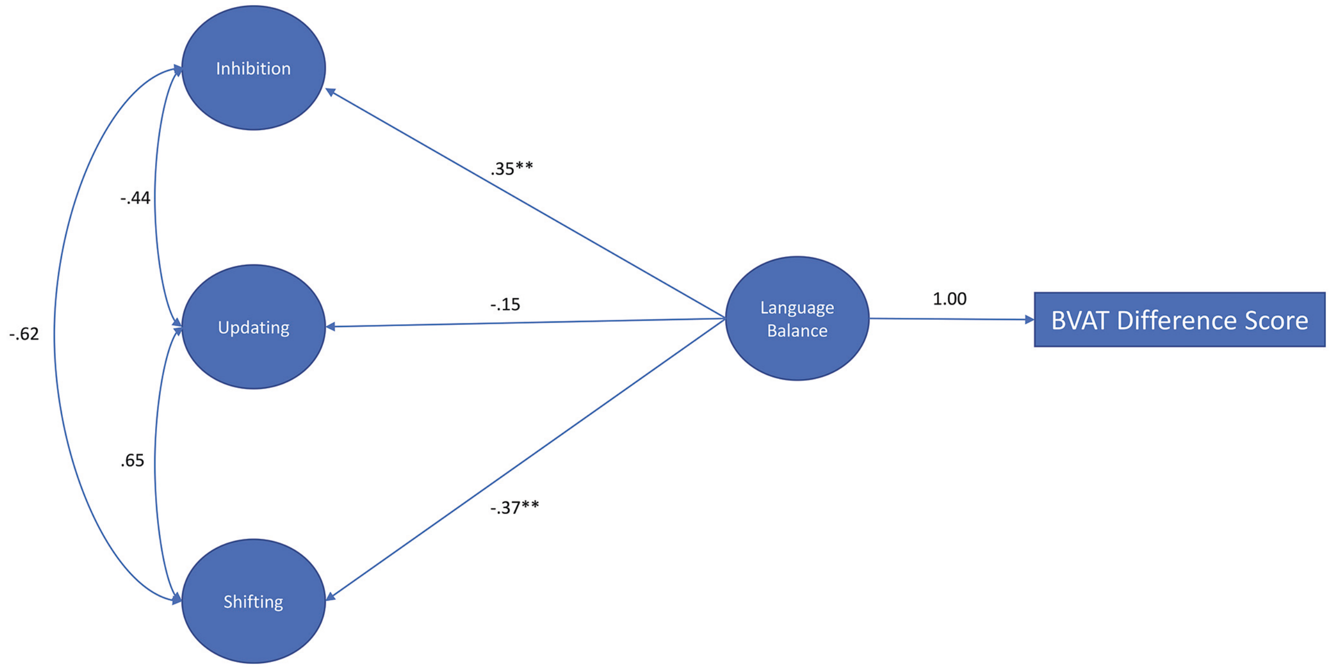
We know of one study that compared language groups on ERP activity during a nonverbal switch task (López Zunini et al., 2019). ERP results revealed larger switching and mixing N2 ERP amplitude in bilinguals along with lower switching and mixing costs. Accuracy rates were similar across groups. The authors concluded that because the N2 amplitude was not restricted to the switch condition, the evidence did not support a cognitive control advantage in bilinguals. Thus, their conclusions are in line with our findings which do not support a bilingual advantage on shifting abilities.

Our study also revealed that language balance predicted both neural inhibition and switch costs such that higher degree of balance was associated with greater N2 ERP amplitude and higher switch costs. These results are in line with ERP studies on language switching which link language proficiency, neural inhibition, and switch costs (Misra et al., 2012). That is, when bilinguals suppress their dominant language to respond in their other language, they exhibit stronger neural inhibition (as evidenced by greater amplitude of the N2 ERP component) and higher switch cost compared with when they suppress their nondominant language to respond in their dominant language. The added cost is believed to result from the additional inhibitory control necessary to suppress the stronger (dominant) language (Jackson et al., 2001).

Ultimately, it may be that the experience of managing two languages strengthens inhibition at the expense of shifting. In support, Friedman and Miyake's team (Friedman et al., 2011) showed that stronger Common EF (which is isomorphic with inhibition) is correlated with weaker shifting abilities. Friedman et al. (2016) found that although most of the variance in EF is stable, owing to genetic correlations across time, a small but significant change in Common EF and shifting specific abilities results from nonshared environmental influences. One example of a nonshared environmental influence is learning a second language. In the case of bilinguals, it may be that the added load to inhibitory processes to manage two languages strengthens inhibition at the expense of shifting.

Thus far, results of the inhibition and shifting tasks reveal congruency between performance and neural activity (Gajewski, Ferdinand, et al., 2018; Gonzalez-Rosa et al., 2013) and argue against a bilingual advantage. Indeed, the most balanced bilinguals showed the weakest performance and highest neural activity. Although these results

Figure 4
SEM Predicting Inhibition, Updating, and Shifting From Language Balance for Behavioral Data



Note. SEM = structural equations modeling. See the online article for the color version of this figure.
** $p < .01$.

support our contention that inhibition in bilinguals develops in tandem with second language proficiency, our findings argue against the premise that stronger inhibitory control yields better performance. Indeed, current study results suggest that stronger inhibition negatively impacts shifting ability.

Updating

Results of the *n*-back task revealed group performance differences favoring monolinguals in the absence of group differences in neural activity. We speculate that group differences in neural activity were not measurable because too few trials contributed to the ERP. Indeed, the total number of target trials was 54, 18 per *n*-back level. For comparison, there were 72 NoGo trials in the Go/NoGo task.

Unlike inhibition and shifting, updating behavior was unrelated to language balance. This finding is consistent with Friedman et al. (2016) who noted that unlike the Common EF and shifting specific abilities, nonshared experiences (such as the experience of being bilingual) do not significantly modify the stability of updating.

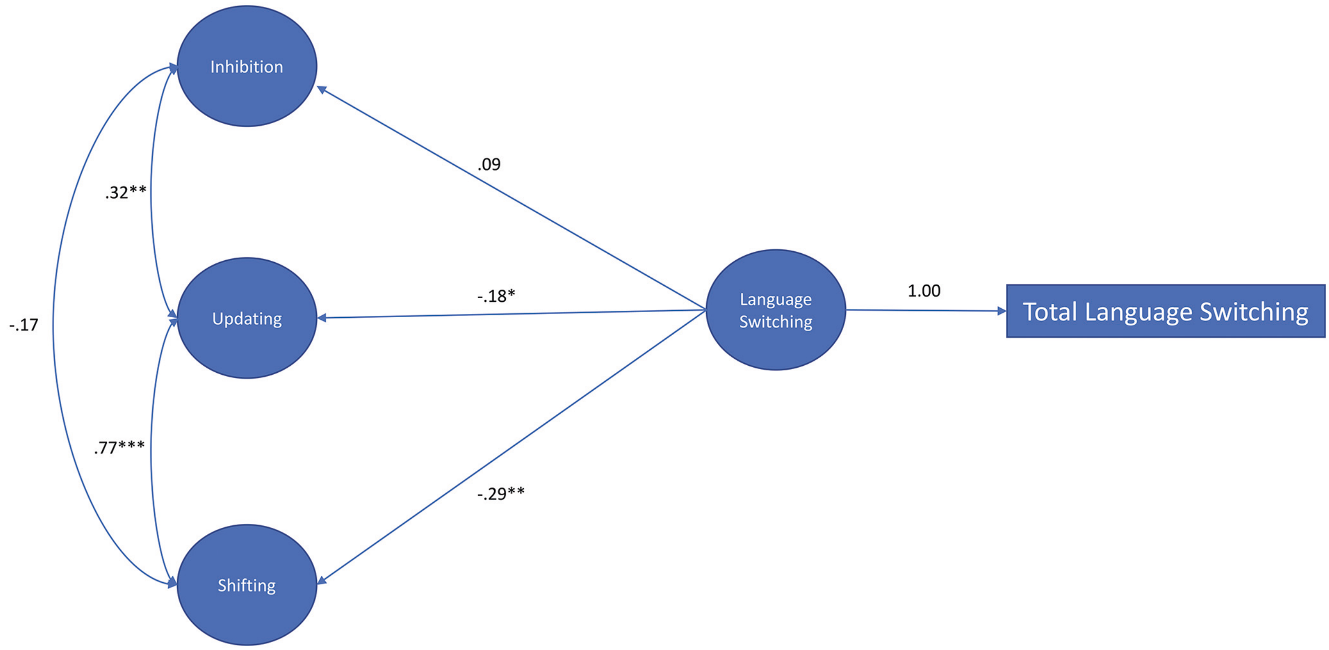
We know of one other ERP study that tested young adults (23 monolinguals and 21 bilinguals) on the *n*-back task (Morrison et al., 2019). Their study showed P3 ERP amplitude favoring bilinguals in the absence of group performance differences. Remarkably, their ERP findings are atypical in that the P3 amplitude in bilinguals increased with increasing memory load. This is contrary to expectation as P3 amplitude normally decreases with increasing memory load (Polich, 2007). The researchers interpreted increases in P3 amplitude as an indication of more cognitive resources. However, given the atypical pattern of neural activity and the lack of congruency between neural activity and performance, it is difficult to interpret and put into context the results of their study.

Results of meta-analyses showing larger effect sizes for executive control (vs. storage) component of WM have been attributed to the positive effects of bilingualism (Grundy & Timmer, 2017; Linck et al., 2014). These researchers concluded that WM is an important component of language processing and performance on measures of second language proficiency.

In the current study, we objectively measured language proficiency and generated a measure of language balance. However, language balance did not predict updating performance. Language balance, however, predicted updating neural activity, but the relationship was negative. Thus, higher linguistic balance was associated with fewer updating resources. Together, results of the current study do not support their contention.

One large-scale study tested the construct validity of the *n*-back task and found that performance in young adults shares variance mainly with EF (Gajewski, Hanisch, et al., 2018). They tested 533 participants and grouped them by age, young (20–40), middle-aged (41–60), and old (61–80). They found that the *n*-back in young adults was significantly correlated with Stroop interference and Trail Making Test Part B, a neuropsychological instrument of shifting abilities. This relationship of the *n*-back to interference suppression and shifting would argue against a bilingual advantage on WM given that most studies evaluating bilinguals on either interference suppression or shifting abilities do not reveal a bilingual advantage especially in young adults, our target population (Paap et al., 2017, 2020; Ware et al., 2020).

Figure 5
SEM Predicting Inhibition, Updating, and Shifting From Language Switching for ERP Data



Note. SEM = structural equations modeling; ERP = event-related brain potential. See the online article for the color version of this figure.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Summarily, our findings do not support the contention of a bilingual updating advantage. However, given the incongruency between neural activity and performance findings, these results must be interpreted with caution. In light of the reported relationship between *n*-back performance and other EF abilities (Gajewski, Hanisch, et al., 2018) and the lack of support for a bilingual advantage on these abilities, it may be that the experience of managing two languages does not strengthen memory updating. Future WM ERP studies that reveal congruent neural activity and performance will help answer this question.

Language Switching and Language Balance

Recently, the context in which bilinguals speak their two languages (single context, dual context, dense code switching) has been proposed as a factor that may impact EF (Green & Wei, 2014). In the current study, language switching frequency was strongly correlated with language balance, $r(254) = 00.720$, $p < .001$. This makes sense because switching between languages requires proficiency in both languages unless language switching occurs to retrieve words not known in the active language. In this case, switching occurs because of disparate proficiency. To disentangle the effects of language switching and language balance on EF, we conducted a series of regression analyses with language balance and language switching entered simultaneously in the model. The results showed that language balance, but not language switching, was a significant predictor of inhibition (Go/NoGo) and shifting task performance. Additionally, language balance predicted shifting neural activity. Specifically, we found that higher language balance was associated with poorer performance on both inhibition and shifting tasks and greater shifting neural activity. Thus, language balance was linked to greater neural expenditure and weaker performance. Language balance, however, did not predict neural activity on the Go/NoGo task. This was unexpected since prior work from our laboratory (Fernandez et al., 2013, 2014) revealed correlations between NoGo ERP amplitude and BVAT scores. However, in previous studies we used only one of the three subtests of the BVAT (Oral vocabulary) as our measure of proficiency, and we did not compute a language balance score. Moreover, unlike the previous two studies, 50% of the trials in the current Go/NoGo task included a time pressure. These differences may have obscured the relationship between language balance and the neural inhibition marker in the current study.

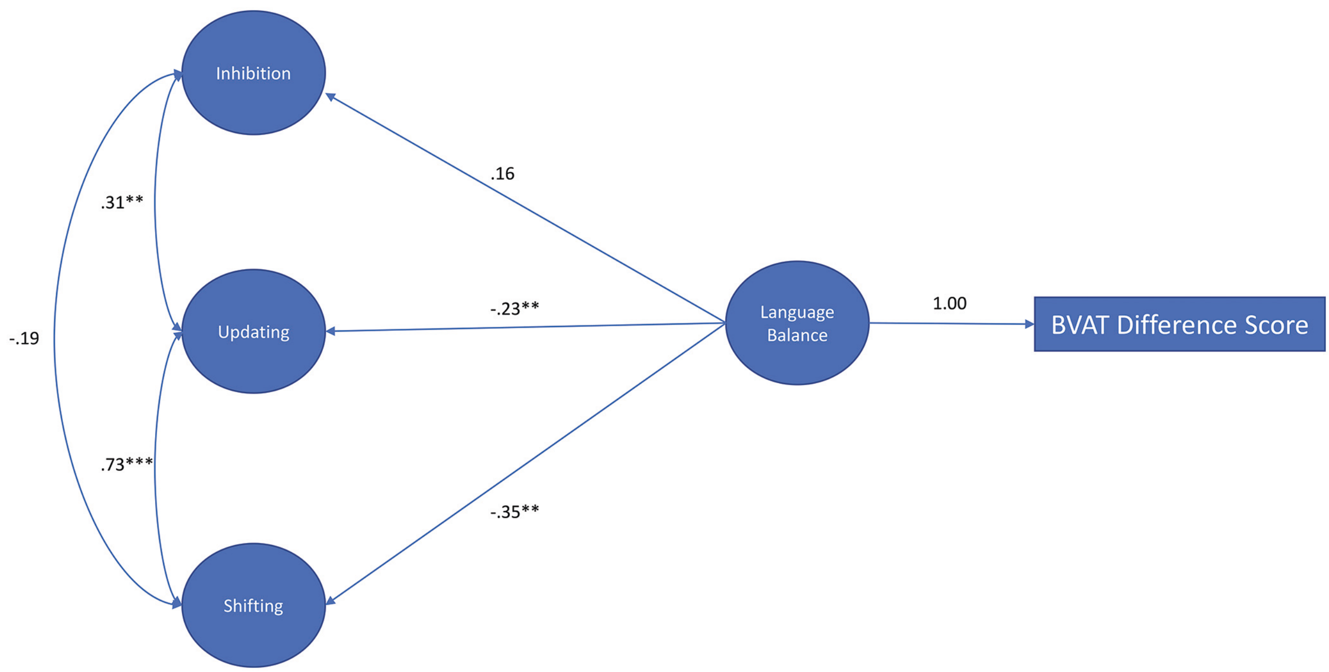
Hartanto and Yang (2020) investigated the relationship between the context in which bilinguals speak their two languages and EF. Passive bilinguals, those who reported no second language usage in their daily lives, were excluded from the study. Results revealed that dual language context predicted shifting abilities and dense code-switching predicted interference inhibition (measured by three versions of the flanker task).

In our study, 17 of 152 total bilinguals reported no language switching. All others reported switching between languages regularly. We did not have information to determine whether our bilinguals engaged in dense code switching. Language switching and language balance were strongly correlated, but language balance was a significant predictor and predicted both shifting and inhibition behaviors. However, in our study, the relationship was negative such that higher linguistic balance predicted poorer performance on shifting and response inhibition tasks.

It is difficult to reconcile differences in results between studies because the groups are different. They selected active bilinguals

Figure 6

SEM Predicting Inhibition, Updating, and Shifting From Language Balance for ERP data



Note. See the online article for the color version of this figure.
** $p < .01$. *** $p < .001$.

and excluded bilinguals who did not engage their two languages. Our language balance and language switching scores were obtained from a sample ranging from monolinguals to balanced bilinguals. Second language proficiency in their study was correlated with language context. The relationship was negative under single context, positive under dual context, and unrelated to dense code switching. In our study, language balance and language switching frequency were positively correlated.

Notably, most studies do not support the hypothesis that language context drives EF advantages in bilinguals (Jylkkä et al., 2017; Kałamała et al., 2020). In a recent study (Paap et al., 2021), participants were carefully screened to select groups of pure single context, dual context, dense code-switching, and monolinguals. Although a small number of participants met the strict criteria, their results failed to reveal a link between language context and EF task performance. Ultimately, our findings indicate that linguistic balance drives neural inhibition in bilinguals and that linguistic balance is negatively correlated with EF.

Structural Models of Executive Functioning

We used the three-factor (inhibition, shifting, updating) EF model proposed by Miyake et al. (2000) as the foundation for this project and expected that our language groups would yield similar models, because development of a second language may alter the mean score of each EF but would be unlikely to alter the structure of EF components. Surprisingly, the first model tested with the observed variables directly loading onto three EF components—inhibition, shifting, and updating—produced a poor fit to the data. Owing to the clear auditory and visual nature of the tasks, we tested a six-factor modality specific EF model. This functioning model provided a good fit to the data. Although an unexpected finding, it underscored the importance of considering the modality of the tasks used to measure these EF constructs.

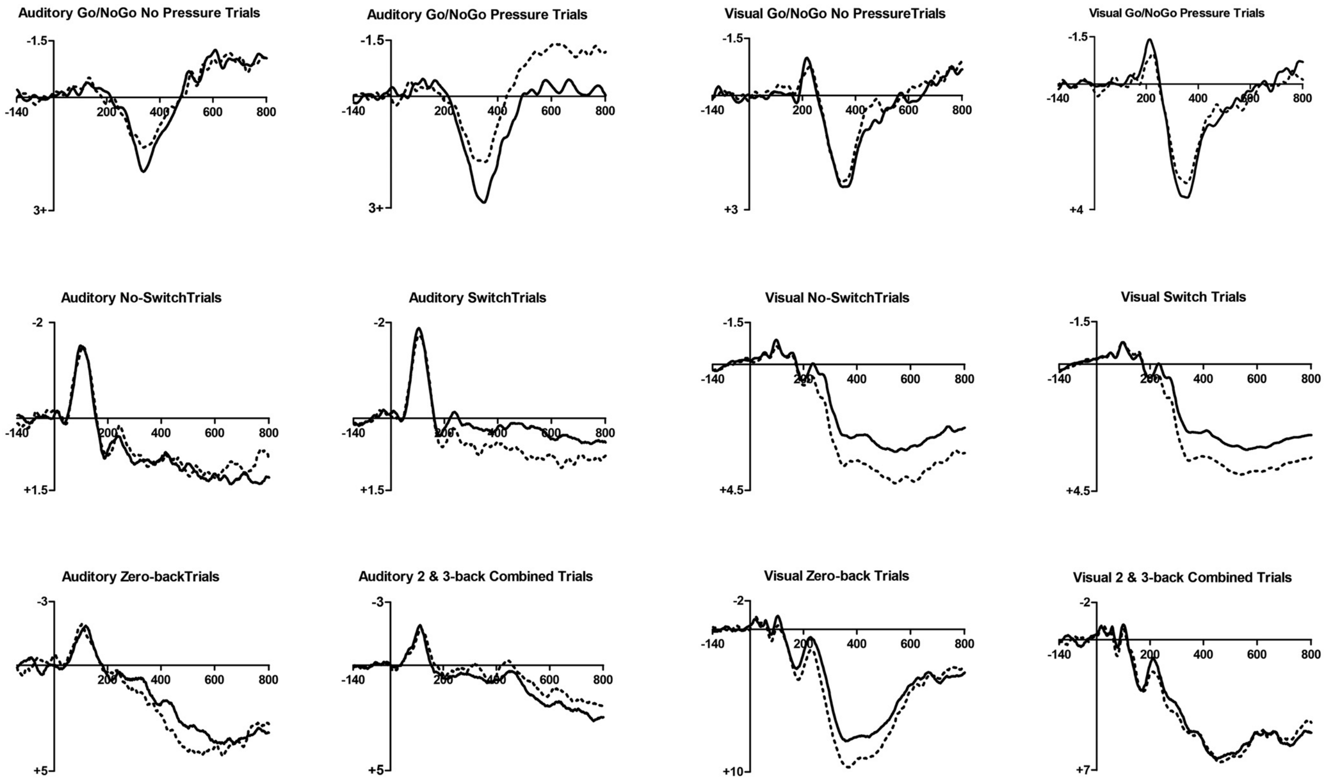
These tasks were designed to test differences between auditory and visual components of the three EF factors. As such, the results suggest that there was sufficient modality specific variance in the tasks that resulted in a poor fit for the initial three-factor model. However, the hierarchical three-factor model provided a better fit and measure of each of the three EF components once the sensory modality specific variance was included in the model. This finding on modality-specific variance supports the contention by Von Bastian and colleagues (von Bastian et al., 2016) that task specific (or modality specific) effects may explain inconsistent findings across studies which compare groups at the task level and argues in favor of structural equations modeling.

The hierarchical three-factor model provided the best fit to the data (behavioral and neural), and the model fit did not differ between the two language groups. This indicates that once modality specific variance is removed, the data fit Miyake and colleagues' (2000) model of EF. Of interest to the current study, the differences observed between the language groups in the six-factor modality specific model on inhibition and shifting were also observed in the hierarchical three-factor model. This provides further support for the idea that differences between monolinguals and bilinguals were attributable to inhibition and shifting rather than sensory modality specific variance.

One difference between the hierarchical three-factor model on behavioral data in the current study and the models by

Figure 7

ERP to Each EF Task by Test Modality and Group (Monolinguals Represented by Dashed Line, Bilinguals by Solid Line)



Miyake et al. (2000) is the direction of the paths between inhibition and both shifting and updating. In their work, the relationship between these constructs was positive. However, we observed a significant negative relationship between inhibition and both shifting and updating. One explanation is that the two studies used different dependent variables as measures of inhibition. We computed response bias, which combines hit rate on Go trials with false alarm rate on No/Go trials into one metric, as a proxy for response inhibition. Miyake's group computed hit rate on the Antisaccade task, false alarm rate on the Stop-signal task (a task similar to our Go/NoGo task) and RT difference on the Stroop task. It is plausible that these differences in dependent variables explain the difference in path correlations between latent variables in the two studies.

Limitations

Several limitations exist in the current study that are important to note. First, the sample size for latent variable modeling is sufficient when the entire sample is collapsed, but it is smaller than preferred for analyses by group. Nevertheless, the sample size is similar to those in other published works (Miyake et al., 2000). Second, the current study did not use a more comprehensive self-report measure of language context. A more comprehensive self-report measure in conjunction with objective language balance measures will help to clarify the role of switching on EF. A strength of this study is that our language groups were similar on demographic characteristics because others have suggested that demographic characteristics may be responsible for language group differences observed in prior work (Paap et al., 2015). However, because our sample is not typical of Spanish speaking bilinguals in the United States (see U.S. Census Bureau, 2020, for SES information) our results may not be representative of the population.

Future Directions

Given the relationship between inhibition and shifting revealed in this study, both at the behavioral and neural level, replicating these findings will contribute to our understanding of the link between bilingualism and EF. Additionally, using measures of interference inhibition, which was not assessed in the current study, will permit a broader understanding of the extent to which language control impacts inhibitory processes. This study revealed a strong relationship between language balance and both neural and behavioral indices of EF. Replicating these results will advance our understanding of neuroplasticity in bilinguals. This study demonstrated the unique contribution of test modality. Future work examining the impact of bilingualism may want to consider examining EF by other dimensions and using a latent model approach to extend the results of this study. Last, given the unanswered questions regarding the contribution to EF of the context in which bilinguals speak their two languages and the relationship to language balance, future studies would benefit from using objective measures of linguistic proficiency and comprehensive language switching questionnaires.

Summary

This study revealed that a higher degree of language balance predicted poorer performance on both inhibition and shifting tasks and greater shifting neural amplitude. Thus, language balance was linked to greater neural expenditure and weaker performance. Without objective measures of languages balance, this important finding would not have been revealed. Additionally, the current findings emphasize the importance of considering the sensory modality of the tasks used to study EF constructs. Although our final model did support a hierarchical three-factor model consistent with Miyake et al. (2000), the model was only a good fit of the data when six sensory modality-specific latent variables were included as first order factors.

References

Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., Cappa, S. F., & Costa, A. (2012). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cerebral Cortex*, *22*(9), 2076–2086. https://doi.org/10.1093/cercor/bhr287

Cespón, J., & Carreiras, M. (2020). Is there electrophysiological evidence for a bilingual advantage in neural processes related to executive functions? *Neuroscience and Biobehavioral Reviews*, *118*, 315–330. https://doi.org/10.1016/j.neubiorev.2020.07.030

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, *106*(1), 59–86. https://doi.org/10.1016/j.cognition.2006.12.013

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, *113*(2), 135–149. https://doi.org/10.1016/j.cognition.2009.08.001

Daffner, K. R., Chong, H., Sun, X., Tarbi, E. C., Riis, J. L., McGinnis, S. M., & Holcomb, P. J. (2011). Mechanisms underlying age- and performance-related differences in working memory. *Journal of Cognitive Neuroscience*, *23*(6), 1298–1314. https://doi.org/10.1162/jocn.2010.21540

Draheim, C., Hicks, K. L., & Engle, R. W. (2016). Combining reaction time and accuracy: The relationship between working memory capacity and task switching as a case example. *Perspectives on Psychological Science*, *11*(1), 133–155. https://doi.org/10.1177/1745691615596990

Falkenstein, M., Hoormann, J., & Hohnsbein, J. (1999). ERP components in Go/Nogo tasks and their relation to inhibition. *Acta Psychologica*, *101*(2–3), 267–291. https://doi.org/10.1016/S0001-6918(99)00008-6

Fernandez, M., Acosta, J., Douglass, K., Doshi, N., & Tartar, J. L. (2014). Speaking two languages enhances an auditory but not a visual neural marker of cognitive inhibition. *AIMS Neuroscience*, *1*(2), 145–157. https://doi.org/10.3934/Neuroscience.2014.2.145

Fernandez, M., Tartar, J. L., Padron, D., & Acosta, J. (2013). Neurophysiological marker of inhibition distinguishes language groups on a non-linguistic executive function test. *Brain and Cognition*, *83*(3), 330–336. https://doi.org/10.1016/j.bandc.2013.09.010

Fogarty, J., Barry, R., De Blasio, F., & Steiner, G. (2018). ERP components and behavior in the auditory equiprobable go/no-go task: Inhibition in young adults. *Psychophysiology*, *55*(7), e13065. https://doi.org/10.1111/psyp.13065

Friedman, N. P. (2016). Research on individual differences in executive function: Implications for the bilingual advantage hypothesis. *Linguistic Approaches to Bilingualism*, *6*(5), 535–548. https://doi.org/10.1075/lab.15041.fri

Friedman, N. P., Miyake, A., Altamirano, L. J., Corley, R. P., Young, S. E., Rhea, S. A., & Hewitt, J. K. (2016). Stability and change in executive function abilities from late adolescence to early adulthood: A longitudinal twin study. *Developmental Psychology*, *52*(2), 326–340. https://doi.org/10.1037/dev0000075

Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: A behavioral genetic analysis. *Developmental Psychology*, *47*(5), 1410–1430. https://doi.org/10.1037/a0023750

Gajewski, P. D., Ferdinand, N. K., Kray, J., & Falkenstein, M. (2018). Understanding sources of adult age differences in task switching: Evidence from behavioral and ERP studies. *Neuroscience and Biobehavioral Reviews*, *92*, 255–275. https://doi.org/10.1016/j.neubiorev.2018.05.029

Gajewski, P. D., Freude, G., & Falkenstein, M. (2017). Cognitive training sustainably improves executive functioning in middle-aged industry workers assesses by task switching: A randomized controlled ERP study. *Frontiers in Human Neuroscience*, *11*, 81. https://doi.org/10.3389/fnhum.2017.00081

Gajewski, P. D., Hanisch, E., Falkenstein, M., Thönes, S., & Wascher, E. (2018). What does the n-back task measure as we get older? Relations between working-memory measures and other cognitive functions across the lifespan. *Frontiers in Psychology*, *9*, 2208. https://doi.org/10.3389/fpsyg.2018.02208

George, E. M., & Coch, D. (2011). Music training and working memory: An ERP study. *Neuropsychologia*, *49*(5), 1083–1094. https://doi.org/10.1016/j.neuropsychologia.2011.02.001

Gonzalez-Rosa, J. J., Inuggi, A., Blasi, V., Cursi, M., Annovazzi, P., Comi, G., Falini, A., & Leocani, L. (2013). Response competition and response inhibition during different choice-discrimination tasks: Evidence from ERP measured inside MRI scanner. *International Journal of Psychophysiology*, *89*(1), 37–47. https://doi.org/10.1016/j.ijpsycho.2013.04.021

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(2), 67–81. https://doi.org/10.1017/S1366728998000133

Green, D. W., & Wei, L. (2014). A control process model of code-switching. *Language, Cognition and Neuroscience*, *29*(4), 499–511. https://doi.org/10.1080/23273798.2014.882515

Grundy, J. G., & Timmer, K. (2017). Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second Language Research*, *33*(3), 325–340. https://doi.org/10.1177/0267658316678286

Gullifer, J. W., & Titone, D. (2020). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*, *23*(2), 283–294. https://doi.org/10.1017/S1366728919000026

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

Jackson, G. M., Swainson, R., Cunnington, R., & Jackson, S. R. (2001). ERP correlates of executive control during repeated language switching. *Bilingualism: Language and Cognition*, *4*(2), 169–178. https://doi.org/10.1017/S1366728901000268

Jylkkä, J., Soveri, A., Wahlström, J., Lehtonen, M., Rodriguez-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., 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

Kausel, L., Zamorano, F., Billeke, P., Sutherland, M. E., Larrain-Valenzuela, J., Stecher, X., Schlaug, G., & Aboitiz, F. (2020). Neural dynamics of improved bimodal attention and working memory in musically trained children. *Frontiers in Neuroscience*, *14*, 554731. https://doi.org/10.3389/fnins.2020.554731

Kousaie, S., & Phillips, N. A. (2012). Conflict monitoring and resolution: Are two languages better than one? Evidence from reaction time and event-related brain potentials. *Brain Research*, *1446*, 71–90. https://doi.org/10.1016/j.brainres.2012.01.052

Kousaie, S., & Phillips, N. A. (2017). A behavioural and electrophysiological investigation of the effect of bilingualism on aging and cognitive control. *Neuropsychologia*, *94*, 23–35. https://doi.org/10.1016/j.neuropsychologia.2016.11.013

Krizman, J., Marian, V., Shook, A., Skoe, E., & Kraus, N. (2012). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(20), 7877–7881. https://doi.org/10.1073/pnas.1201575109

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

Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods*, *51*(1), 40–60. https://doi.org/10.3758/s13428-018-1076-x

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, *21*(4), 861–883. https://doi.org/10.3758/s13423-013-0565-2

López Zunini, R. A., Morrison, C., Kousaie, S., & Taler, V. (2019). Task switching and bilingualism in young and older adults: A behavioral and electrophysiological investigation. *Neuropsychologia*, *133*, 107186. https://doi.org/10.1016/j.neuropsychologia.2019.107186

Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed.). Bradford Books.

Luo, L., Craik, F. I., Moreno, S., & Bialystok, E. (2013). Bilingualism interacts with domain in a working memory task: Evidence from aging. *Psychology and Aging*, *28*(1), 28–34. https://doi.org/10.1037/a0030875

Misra, M., Guo, T., Bobb, S. C., & Kroll, J. F. (2012). When bilinguals choose a single word to speak: Electrophysiological evidence for inhibition of the native language. *Journal of Memory and Language*, *67*(1), 224–237. https://doi.org/10.1016/j.jml.2012.05.001

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

Morrison, C., Kamal, F., & Taler, V. (2019). The Influence of Bilingualism on working memory event-related potentials. *Bilingualism: Language and Cognition*, *22*(1), 191–199. https://doi.org/10.1017/S1366728918000391

Munoz-Sandoval, A. F., Cummins, J., Alvarado, C. G., & Ruef, M. L. (1998). *Bilingual verbal ability tests*. Riverside Publishing Company.

Nieuwenhuis, S., Yeung, N., van den Wildenberg, W., & Ridderinkhof, K. R. (2003). Electrophysiological correlates of anterior cingulate function in a go/no-go task: Effects of response conflict and trial type frequency. *Cognitive, Affective & Behavioral Neuroscience*, *3*(1), 17–26. https://doi.org/10.3758/CABN.3.1.17

Paap, K. R., Anders-Jefferson, R., Mikulinsky, R., Masuda, S., & Mason, L. (2019). On the encapsulation of bilingual language control. *Journal of Memory and Language*, *105*, 76–92. https://doi.org/10.1016/j.jml.2018.12.001

Paap, K. R., Anders-Jefferson, R., Zimiga, B., Mason, L., & Mikulinsky, R. (2020). Interference scores have inadequate concurrent and convergent validity: Should we stop using the flanker, Simon, and spatial Stroop tasks? *Cognitive Research: Principles and Implications*, *5*(1), 7. https://doi.org/10.1186/s41235-020-0207-y

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., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, *69*, 265–278. https://doi.org/10.1016/j.cortex.2015.04.014

Paap, K. R., 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

Paap, K. R., Sawi, O. M., Dalibar, C., Darrow, J., & Johnson, H. A. (2014). The brain mechanisms underlying the cognitive benefits of bilingualism may be extraordinarily difficult to discover. *AIMS Neuroscience*, *1*(3), 245–256. https://doi.org/10.3934/Neuroscience.2014.3.245

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Ramautar, J. R., Kok, A., & Ridderinkhof, K. R. (2006). Effects of stop-signal modality on the N2/P3 complex elicited in the stop-signal paradigm. *Biological Psychology*, *72*(1), 96–109. https://doi.org/10.1016/j.biopsycho.2005.08.001

Raven, J., Raven, J. C., & Court, J. H. (2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. Pearson Assessment.

Rey-Mermet, A., Gade, M., & Oberauer, K. (2018). Should we stop thinking about inhibition? Searching for individual and age differences in inhibition ability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(4), 501–526. https://doi.org/10.1037/xlm0000450

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*(2), 23–74.

Schneiders, J. A., Opitz, B., Tang, H., Deng, Y., Xie, C., Li, H., & Mecklinger, A. (2012). The impact of auditory working memory training on the fronto-parietal working memory network. *Frontiers in Human Neuroscience*, *6*, 173. https://doi.org/10.3389/fnhum.2012.00173

Shulley, L. J., & Shake, M. C. (2016). Investigating the relationship between bilingualism, cognitive control, and mind wandering. *Journal of Cognitive Psychology*, *28*(3), 257–274. https://doi.org/10.1080/20445911.2015.1128438

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*(1), 34–50. https://doi.org/10.1037/0096-3445.117.1.34

Thomas, S. J., Gonsalvez, C. J., & Johnstone, S. J. (2009). Sequence effects in the Go/NoGo task: Inhibition and facilitation. *International Journal of Psychophysiology*, *74*(3), 209–219. https://doi.org/10.1016/j.ijpsycho.2009.09.002

U.S. Census Bureau. (2020). Income and poverty in the United States: 2019. https://www.census.gov/library/publications/2020/demo/p60-270.html

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

van den Noort, M., Struys, E., Bosch, P., Jaswetz, L., Perriard, B., Yeo, S., Barisch, P., Vermeire, K., Lee, S. H., & Lim, S. (2019). Does the bilingual advantage in cognitive control exist and if so, what are its modulating factors? A systematic review. *Behavioral Sciences*, *9*(3), 27. https://doi.org/10.3390/bs9030027

von Bastian, C. C., Souza, A. S., & Gade, M. (2016). No evidence for bilingual cognitive advantages: A test of four hypotheses. *Journal of*

Experimental Psychology: General, *145*(2), 246–258. https://doi.org/10.1037/xge0000120

Ware, A. T., Kirkovski, M., & Lum, J. A. G. (2020). Meta-analysis reveals a bilingual advantage that is dependent on task and age. *Frontiers in Psychology*, *11*, 1458. https://doi.org/10.3389/fpsyg.2020.01458

Woumans, E., Ceuleers, E., Van der Linden, L., Szmalec, A., & Duyck, W. (2015). Verbal and nonverbal cognitive control in bilinguals and interpreters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(5), 1579–1586. https://doi.org/10.1037/xlm0000107

Woumans, E., Van Herck, S., & Struys, E. (2019). Shifting gear in the study of the bilingual advantage: Language switching examined as a possible moderator. *Behavioral Sciences*, *9*(8), 86. https://doi.org/10.3390/bs9080086

Received September 28, 2020
Revision received August 4, 2022
Accepted August 8, 2022 ■