Unstrooping the Stroop Effect

Can Bilinguals Reduce Stroop Interference Through Other-Language Mediation?

Ana C. Oliveira

University of Florida

Submitted under the supervision of Dr. Jorge Valdes Kroff and Dr. Lise Abrams to the University Honors Program at the University of Florida in partial fulfillment of the requirements for the degree of Bachelor of Sciences, summa cum laude in Psychology.

Abstract

Variations of the Stroop task continue to be employed as a robust, automatic measure of interference (MacLeod, 1991). 54 Spanish-English bilingual speakers and 54 English monolingual speakers participated in a button-press Stroop task administered in English (stimuli and response). The task presented participants with four conditions; baseline (BL), English labels (EL), Spanish labels (SL), and alternative labels (AL). Participants saw words in the computer screen in either *blue*, *green*, or *yellow* ink and had to press the corresponding option in a button-response box. The SL condition served as a suggestion to bilingual participants to use their second language to help them complete the task more efficiently. The Stroop interference effect was defined by comparing reaction time (RT) between incongruent and neutral trials for eligible stimuli. Results showed that bilinguals were overall faster and had a smaller Stroop interference effect compared to monolinguals but there was no significant difference between groups for the SL condition. Bilingual participants presenting faster overall RT compared to monolinguals could be evidence to an advantage in executive functioning.

Keywords: Stroop effect, interference, bilingual, conflict adaptation

Unstrooping the Stroop Effect

Can Bilinguals Reduce Stroop Interference Through Other-Language Mediation?

A growing body of evidence supports the claim that bilingual speakers present executive functioning advantages due to their experience with more than one language (Morales, Gómez-Ariza, & Bajo 2013). Among cognitive tasks that investigate executive functioning, the Stroop task measures inhibitory control by measuring interference. Participants must name the ink color despite the color word presented (*blue* presented in yellow) while their reaction time is recorded and compared with control stimuli such as control lists (color words in black ink), neutral words in different color inks, or congruent trials (*green* presented in green. Tzelgov, Henik. & Leiser, 1990). Despite different manipulations in the task, it generally serves as a robust measure of interference between word-reading, a task-irrelevant stimulus, and naming colors (Milham et al., 2001). Several studies have used the task for different purposes including the study of selective attention, conflict resolution, executive control in normal and special populations (Naylor, Stanley, & Wicha, 2012).

Most studies using the Stroop task to investigate bilingual populations have tested different groups of speakers within each of their languages (stimuli and response in the same language for separate conditions; Milham et al., 2001; Rosselli et al., 2002; Wang, Fan, Liu, & Cai, 2016) as well as between them (stimuli and response in different languages in the same condition; Tzelgov et al., 1990; Costa, Albareda & Santesteban, 2008; Preston and Lambert, 1969). Therefore, researchers have created and adapted the Stroop task in various ways to better understand how bilinguals deal with automaticity and to what extent the possible activation of their languages could correlate with the magnitude of the Stroop effect (see Francis, 1999 for a review). Bilinguals experience a within language Stroop superiority effect where the magnitude of the Stroop effect for within-language tasks is generally twice as much as between languages if there is no orthographic overlap for those languages (Naylor et al., 2012). Although there is no consensus about how bilinguals deal with the conflict of having more than one language in the brain, a mechanism of inhibition proposed by Green (1998) and Kroll et al. (2012) would allow bilinguals to use only one language despite a likely simultaneous activation of their languages which would explain the reduced Stroop interference between languages. To further explore this idea, we propose a new design for the task where bilingual speakers will perform the task in English only (stimuli and response) and receive the suggestion to use their second language to reduce the interference from word reading.

One of the most challenging aspects about working with bilingual groups is defining the level of proficiency and dominance for the languages they speak. Researchers usually use standardized grammar/vocabulary tests, self-report questionnaires and age of acquisition of a second language to determine proficiency and dominance (Portocarrero, Burright, & Donovick, 2007). Using those measures, analyses usually include any interaction between the magnitude of the interference effect across languages and bilingual proficiency in each language (Naylor et al., 2012). However, using grammar tests alone could be deceiving if you have a highly proficient bilingual that did not have formal education on their second language. Self-report could skew the level of proficiency, and age of acquisition could mask the real stage of acquisition in cases where speakers acquire a second language at a certain age but do not use that language in daily life (e.g. English-Spanish bilinguals that mostly speak Spanish with distant relatives). We used several measures to determine level of proficiency for the current task (see *Methods* section) to assess how language dominance would correlate with performance in the task.

In addition to the classical Stroop interference that measures cognitive inhibition in the response-level (eligible trials), we were interested in comparing that to a non-response level block (ineligible trials). Milham et al. (2001) showed that the anterior cingulate cortex (ACC) and the dorsolateral prefrontal cortex (PFC) which are involved with attentional control were activated differently depending on whether the level of conflict was at the response level or not. Furthermore, a response set effect suggests that response eligible words cause a larger Stroop effect due to greater conflict occurring at the response level rather than the stimulus evaluation phase (Naylor et al., 2012). A response set effect for monolinguals would be equivalent to a between-within language Stroop difference for bilinguals (Roelofs, 2010). Having a non-response level block will allow us to analyze if the proposed strategy is relevant at the response level.

The goal for this study is to further explore differences within the bilingual population and between bilingual and monolingual participants. Both groups were presented the same conditions where the manipulation is the use of labels as a proposed strategy where bilinguals could use their non-target language to map their responses and monolinguals could use words that prototypically represent colors to reduce the Stroop interference (e.g. peas for *green*). In a previous study with the same type of stimuli and similar conditions, researchers tested 24 Spanish-English bilinguals in a button-press response Stroop task. Preliminary accidental findings suggested that bilinguals could use the non-test language as an alternative response mapping strategy, i.e. bilinguals may recruit Spanish to map their button responses (i.e., associate responses to *azul* “blue”, *amarillo* “yellow”, *verde* “green”) to diminish the Stroop effect. Monolinguals showed stable Stroop interference for the later two sessions while bilinguals showed greater reduction of RT for incongruent and smaller difference between incongruent and neutral for later sessions. As a follow-up, we extended the set of stimuli and tested more participants in the same conditions (see *methods* section). Ultimately, this research will be another stepping stone for a better understanding about the interaction between bilingualism, cognitive control, and the purported bilingual advantage.

Research questions and hypotheses

Studies that used the Stroop task as a measure of conflict adaptation testing bilingual samples have explored many facets of how a second language in the brain relates to reaction time and interference. However, the study of bilingualism still raises many question as to which extent a second language changes the way speakers deal with conflict. We focused on three main questions:

1. Could bilinguals adopt an other-language strategy to reduce Stroop interference?
2. Is there a same-language analog for monolinguals?
3. What aspects of bilingualism correlate with the Stroop effect and the application of this strategy?

In agreement with previous findings, results should reveal overall slower RT for bilinguals compared to monolinguals (Rosselli et al. 2002). However, bilinguals should also have smaller Stroop interference effect compared to monolinguals and a greater reduction of that interference in later sessions compared to BL especially for the SL condition agreeing with preliminary findings.

Monolinguals could also present a reduction of interference for the AL condition but smaller in magnitude than bilinguals for the last two conditions. This prediction is based on previous findings that bilinguals have better conflict adaptation skills from the constant juggling of two languages in the brain (Kroll, Dussias, Bogulski & Kroff, 2012) and therefore present better suppression of task-irrelevant stimulus (Milham et al., 2001). Suarez et. al (2014) found that English fluency predicted better performance of English-Spanish bilinguals for incongruent trials in a Stroop task in Spanish. Similarly, we should find a correlation between performance in the Stroop task in English and fluency in Spanish.

Methods

Participants

This study was approved by the Institutional Review Board (IRB) of the University of Florida. Each participant was given a description of the experiment and then signed an informed consent form. A control sheet was used to document each participant's experiment number, tasks completed, and any other notes necessary. Participants were recruited in two ways; IRB-approved flyers were posted on boards around campus inviting Spanish-English bilingual speakers to participate in the study for a small monetary compensation whereas English monolingual speakers were recruited through the Undergraduate Research Pool, SONA. Researchers were able to recruit 54 bilinguals (39 females) with ages ranging from 18 to 30 years (M = 21.3 years, SD = 2.91) and 54 monolinguals (34 females) with ages ranging from 18 to 29 years (M = 19.93 years, SD = 2.05). Due to missing data, five participants were excluded from the bilingual group and four from the monolingual.

According to Francis (1999), using the term "bilingual" encompasses numerous definitions that are not used consistently in the literature. She defines bilingualism according to Grosjean (1992, p. 51): "Bilingualism is the regular use of two (or more) languages, and bilinguals are those people who need and use two (or more) languages in their everyday lives."

Aligned with this definition, we recruited Spanish-English bilinguals speakers who learned Spanish before or at the same time as English and speak both languages in their daily lives. English monolinguals are speakers who can only carry a conversation in English with limited knowledge of another language (up to school/college requirement level). To determine proficiency level, bilingual participants completed different language tests in both languages which will be further discussed in the *materials* section.

Materials

The different tasks in this project are either computer-based or filled out on paper by research assistants. The Stroop task is performed on a computer where participants see one word at a time on the screen. Words appear in either blue, yellow, or green ink and participants must select which ink color is being presented in a push button response box. The task design and stimuli were adapted from Milham et al. (2001). There are three types of stimuli; for eligible stimuli the color word is also a possible response option (e.g. *green* in blue ink) while ineligible stimuli consist of color words that are not a response option (e.g. *orange* in yellow ink). The last type of stimuli are neutral words (e.g. *farmer*) used for comparison to establish the amplitude of the Stroop effect. These three types of stimuli (See Table 1) are shown pseudorandomly and organized into four sessions with 192 trials each; first, participants must memorize the order for

Ineligible Stimuli		**Eligible Stimuli**	
Incongruent	Neutral	Incongruent	Neutral
red	*stage*	*green*	*deal*
orange	*farmer*	*yellow*	*plenty*
brown	*tax*	*blue*	*horse*

Table 1: Types of stimuli for Stroop task

what colors should be pressed in the button response box which will serve the baseline (BL) for the other sessions. In the second session, English labels (EL) for each color are placed right above the buttons. For the third and fourth sessions, labels for the colors in Spanish (SL) and alternative words (AL) that are prototypically associated with the colors (e.g. *ocean* for blue) are used and the sessions are counterbalanced between participants (see Figure 1). After the task is done, participants answer a post-experiment questionnaire (Appendix A) that assesses participants' understanding of experiment goals and their use of strategies including bilinguals use of their second language. The second computer-based task, the AX Continuous Performance

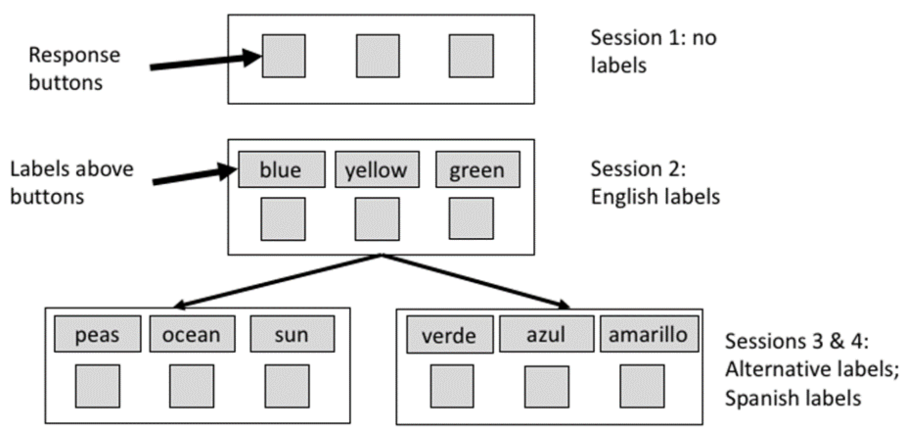


Figure 1: the four conditions for the Stroop task

Task (AX-CPT), measure reactive (inhibitory) and proactive (monitoring) cognitive control. Letters appear on the screen one at a time where participants must press either 'yes' or 'no' in a push button response box after each letter according to a few conditions. After the four first letters participants must press 'no' regardless of what letter appears in which order. For the fifth letter participants will press 'yes' only if the first letter in the sequence was a red 'A' (cue) and the last letter is a red 'X' (probe). In any other case, participants will press 'no' after the fifth letter as well. The AX trials will be the target response and compose 70% of trials. AY trials compose 10% of trials and require participants to use proactive control due to the cue that prepares them to a "yes" response. BX trials also appear 10% of the time and demand participants to engage in reactive control where they need to inhibit the "yes" response after the probe (see Figure 2).

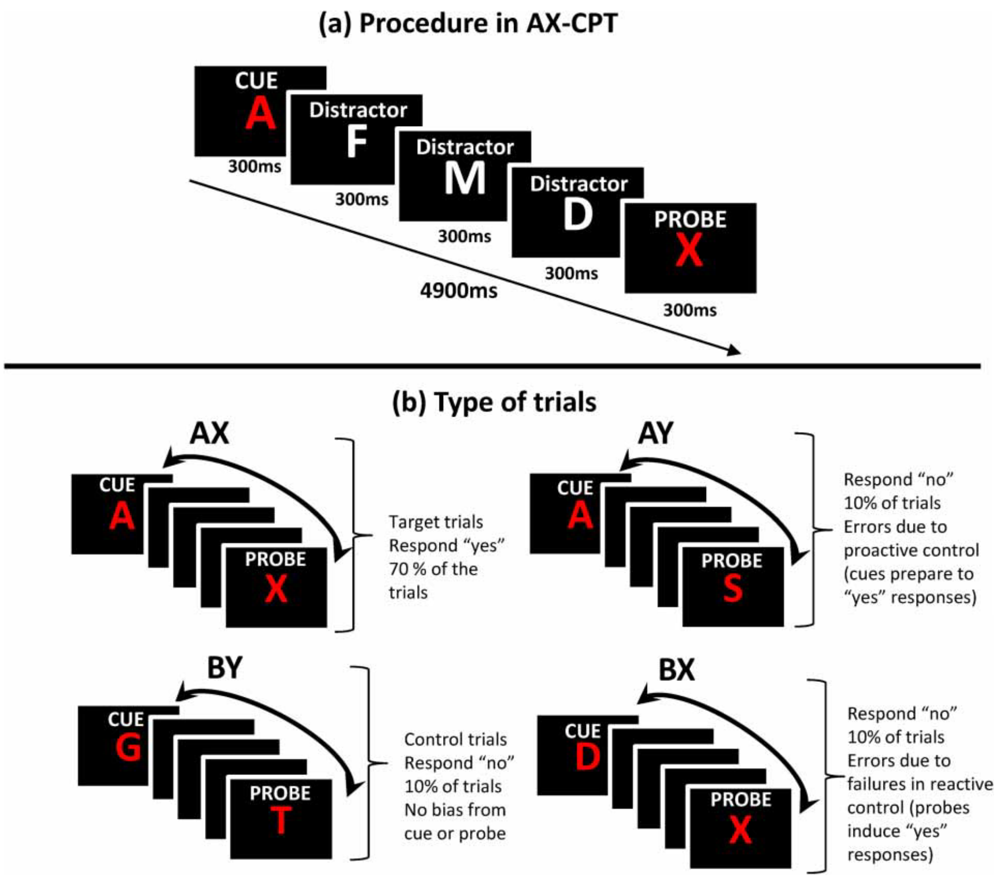


Figure 2: Procedure for the AX-CPT.

(a) The series of events in a typical target trial.

(b) The 4 different types of trial, the correct response for each of them, and the proportion of a given trial during the task.

From Morales et. al (2013)

The Backward Digit Span Task (Working Memory Test from Wechsler Memory Scale-Revised) is administered by research assistants where they read a string of numbers out loud and instruct participants to repeat it back to them in the reverse order. This task has been used by researchers as a measure of working memory capabilities (Ramsay & Reynolds 1995).

To determine language proficiency, two multiple choice standardized grammar tests are used; an adapted version of the *Diploma de Español como Lengua Extranjera* (DELE) is the Spanish grammar assessment and the *Michigan English Language Institute College Entrance Test* (MELICET) is the English grammar version. In addition to these grammar tests, research assistants administered a modified bilingual version of the Boston Naming Task (BNT; Kaplan et al., 1983) where participants see images on the screen and have to name them in English and Spanish in separate blocks with distinct images. This task attempts to quantify vocabulary size and the ratio found by dividing Spanish by English scores indicated if a participant was Spanish dominant (ratio ≥ 0.8) or English dominant (ratio < 0.8).

The Language History Questionnaire (LHQ. Appendix B) presents participants with questions about their language background including the number of languages they speak, the frequency in which they speak those languages, age of acquisition, language preference, and whether they code-switch. These questions will help researchers determine in addition to grammar tests how proficient/dominant participants are in each of their languages and it will serve to double-check monolingual participants' eligibility criteria.

Procedure

Testing occurred at the Bilingual Sentence Processing Lab at Dauer Hall and lasted approximately two hours for bilinguals and one hour and a half for monolinguals. Research assistants explained each section of the informed consent to participants and had them sign it. Participants went through each task on the lab computer except for the post-experiment questionnaire and digit span task. The software E-prime ran both the Stroop task and the AX-CPT where it recorded RT and accuracy onto the computer. We designed the LHQ, DELE, and MELICET through the Qualtrics website that records and compiles the answers. Finally, the

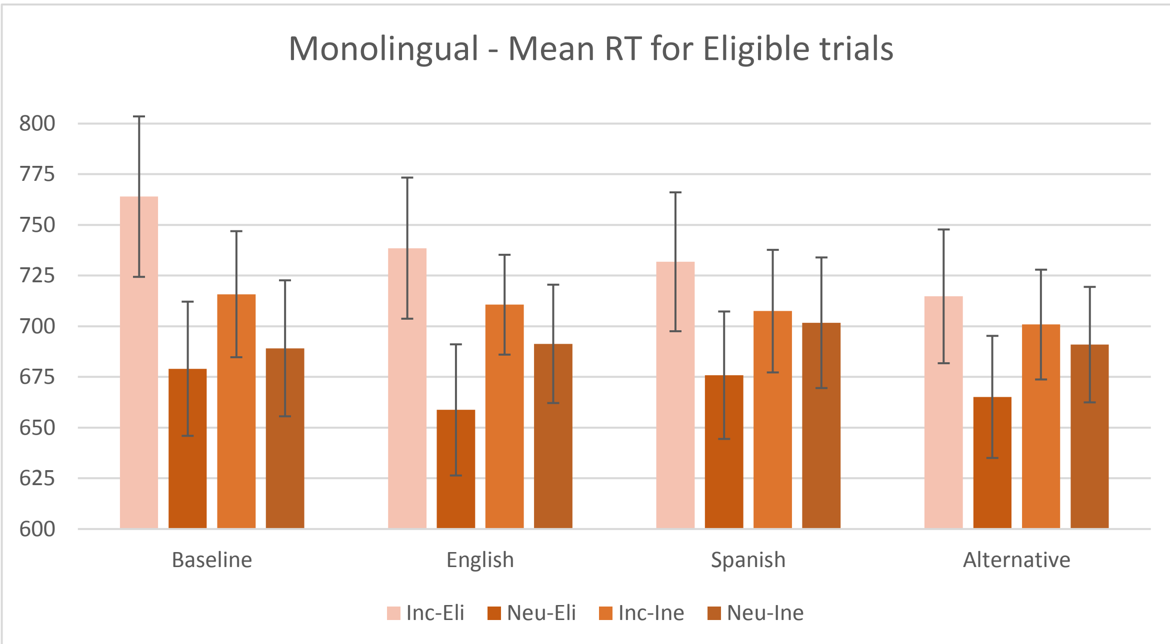
BNT (images presented on the computer screen), the digit span task, and the post-experiment questionnaire were collected by research assistants on paper.

Results

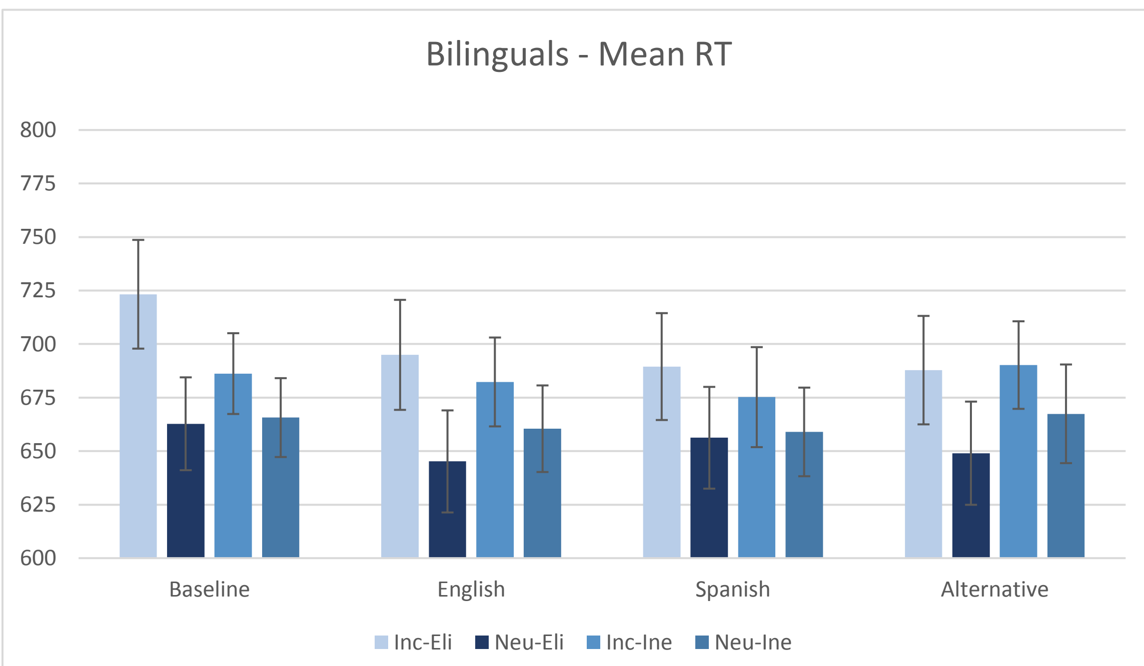
We used repeated measures analysis of variance (ANOVA) as the main statistical test. Post-hoc t-tests helped determine which variables were driving the interactions. Bivariate correlation tests were used to find associations between Stroop variables, language proficiency and cognitive measures. We averaged RT for each participant for every level and excluded any trials over or under 2.5 standard deviations from each individual mean. The error analysis revealed an approximate 2% error rate and did not show any significant differences between groups. Analysis 1 had 16 levels, 4 (Session) x 2 (Congruency) x 2 (Response Type) x 2 (Group). Analysis 2 consisted of 8 levels, 4 (Session) x 1(Incongruent-neutral) x 2 (Response Type) x 2 (Group). We ran correlations within the levels of analysis 2 with the digit span task and all proficiency tests (MELICET, DELE, BNT, and BNT ratio) for bilinguals to determine if proficiency had any correlation with RT for those levels. Analysis 3 had 6 levels, 3 (Session: AL-BL; SL-BL; EL-BL) x 2 (Congruency) x 1 (Eligible). We ran the same correlations from analysis 2 but with the levels of analysis 3.

Analysis 1

After running repeated measures ANOVA, results showed a main effect for congruency $F(1,106) = 90.48$, $MSE = 7019.84$, $p < .01$ where incongruent trials had higher mean RT than neutral mean RT indicating that both groups presented the Stroop interference effect. This difference was present for all sessions ($p < .01$) in both eligible and ineligible blocks as pairwise comparisons showed in post-hoc t-tests. The ANOVA also revealed an interaction between session, congruency, and response type $F(3,318) = 2.63$, $MSE = 1104.447$, $p < .05$ in which incongruent eligible trials differed significantly from BL to all other conditions ($p < .04$) and EL differed



Graph 1: Error bars represent ± 1 SD from the mean

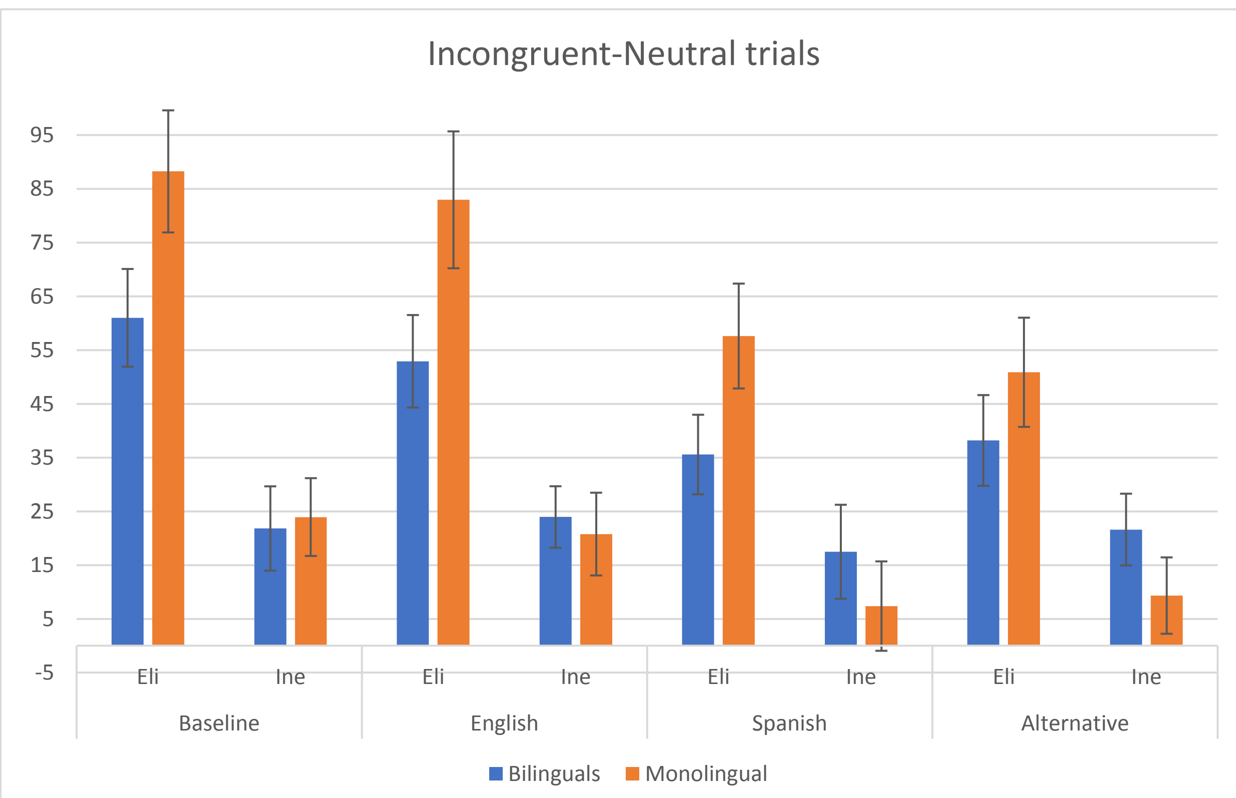


Graph 2: Error bars represent ± 1 SD from the mean

from AL significantly ($p < .05$). No other difference between sessions was significant ps>.05. This analysis also showed an interaction for congruency, response type, and group $F(1,106) = 7.81$, $MSE = 2880.09$, $p < .01$ where a post-hoc t-test revealed that the mean difference between incongruent and neutral trials for bilinguals in the eligible condition was 46.915 ($p < .01$) while for monolinguals was 69.916 ($p < .01$) showing a bigger Stroop effect for monolinguals overall. For an overview of mean RT for each group see Graph 1 and Graph 2. No other effects were significant, ps > .05.

Analysis 2

After subtracting mean neutral trial RT from incongruent trial RT for each session and response type, the repeated measures ANOVA indicated a main effect for session $F(3,318) =$



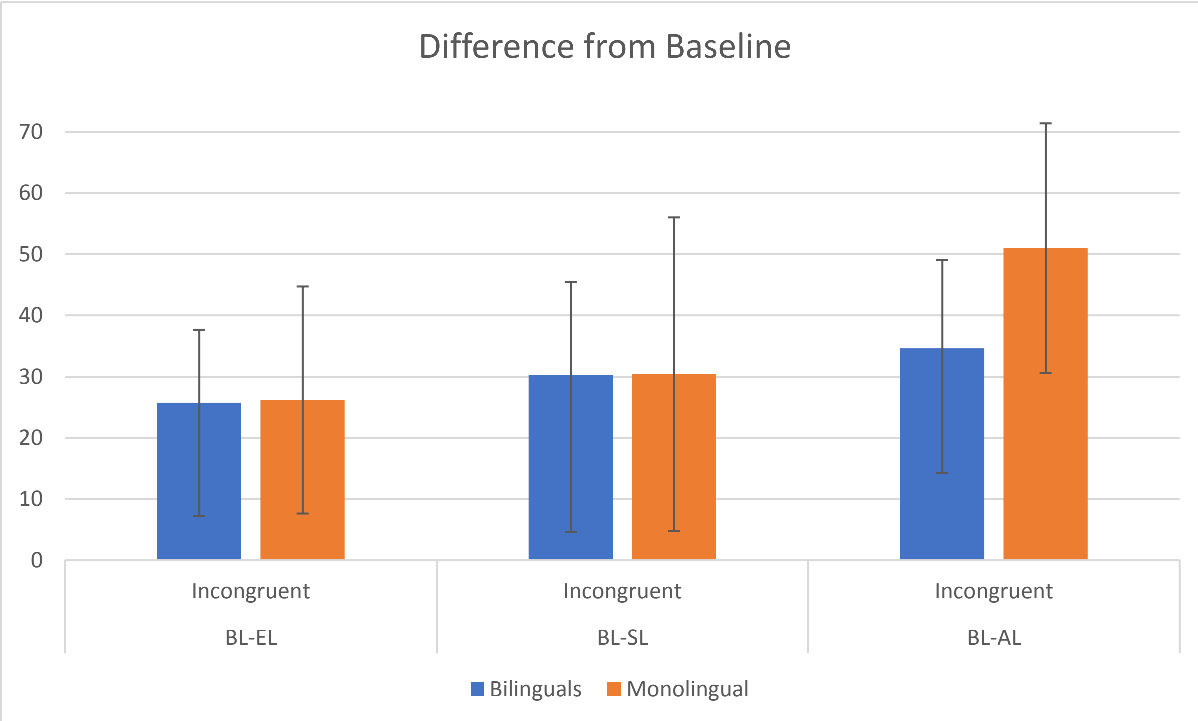
Graph 3: Error bars represent ± 1 SD from the mean

10.076, $MSE = 2159.736$, $p < .001$ and response type $F(1,106) = 60.422$, $MSE = 5760.177$, $p < .001$ where eligible trials have higher RT than ineligible. Analysis also showed an interaction between response type and group $F(1,106) = 7.808$, $MSE = 5760.177$, $p < .006$ where bilinguals presented smaller difference than monolinguals for eligible trials (See Graph 2). The ANOVA also displayed an interaction between session and response type $F(3, 318) = 2.63$, $MSE = 2208.893$, $p < .05$ where BL differed from SL ($p<.001$) and AL ($p<.001$), EL differed from SL ($p<.002$) and AL ($p<.001$), but SL showed no significant difference from AL ($p < .745$) for eligible stimuli. No other effects were significant, ps > .05. All correlations between Stroop variables and the digit span task and proficiency tasks were not significant ps > .05.

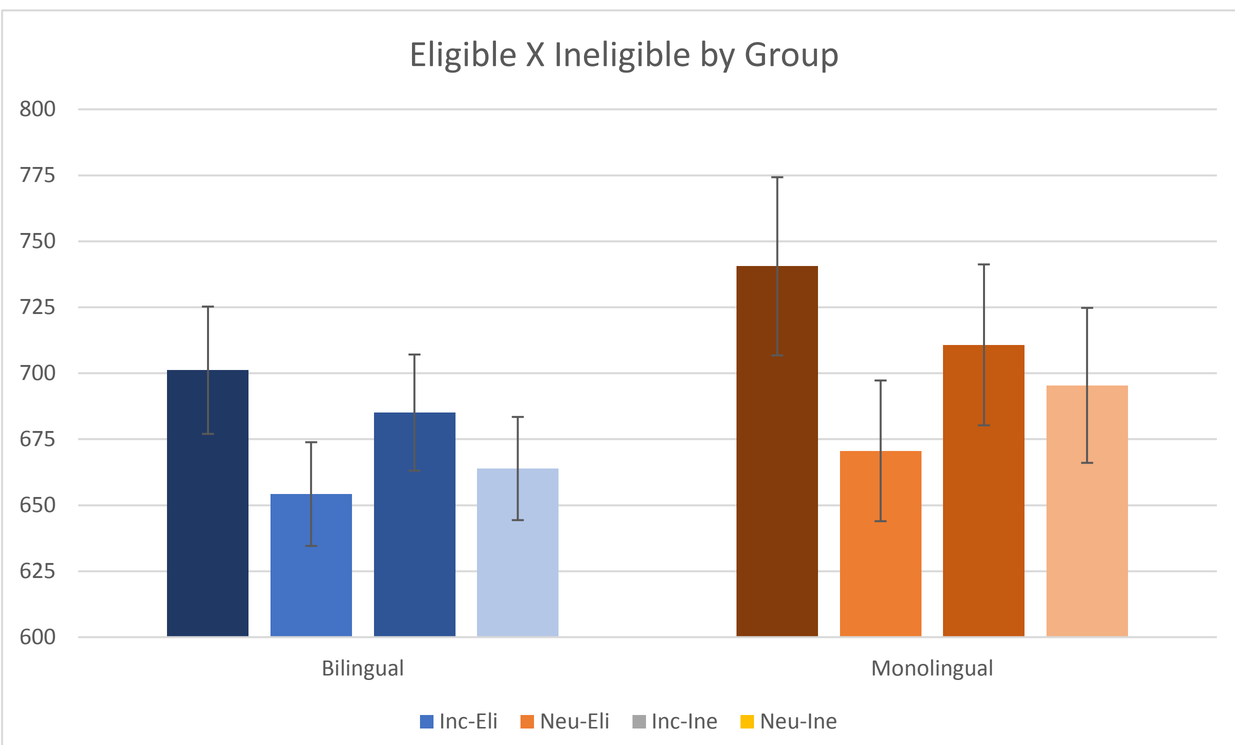
Analysis 3

The difference between each session and baseline was calculated for incongruent and neutral trials for the eligible condition only. Repeated measures ANOVA revealed a main effect for congruency $F(1,106) = 3389.317$, $MSE = 3389.317$, $p < .001$ where incongruent trials had higher RT than neutral. The analysis showed an interaction between session and congruency $F(2,212) = 8.186$, $MSE = 1107.749$, $p < .001$ where incongruent trials were only significantly different from neutral for SL minus BL ($p < .001$) and AL minus BL ($p < .001$) with mean differences of 28.03 and 30.09 respectively. No other effects were significant, ps > .05. Graph 3 shows the overall mean for incongruent trials for both groups.

Bivariate tests found a negative correlation between both SL-BL $r(108) = -.189$, $p<.05$ and AL-BL $r(108) = -.213$, $p<.03$ and the digit span test showing that participants who performed better in the digit span task presented smaller difference between these two conditions and BL. No other correlations were significant, ps > .05.



Graph 4: Error bars represent + 1 SD from the mean



Graph 5: Error bars represent + 1 SD from the mean

Discussion

The current study aimed to investigate whether bilinguals would reduce Stroop interference in later sessions by utilizing their non-target language as a strategy, how they compare to monolinguals, and whether monolinguals could find a same-language strategy to reduce interference. We found that the Stroop interference effect was present for both groups as participants had slower RT for incongruent eligible stimuli compared to neutral eligible stimuli across all sessions. Bilingual participants were overall faster than monolingual participants for both incongruent and neutral trials for eligible stimuli only although results lacked significance for any difference across sessions between the two groups. Monolinguals showed an overall larger Stroop interference effect than bilinguals. For both groups combined, BL had higher RT than the other three conditions while EL had higher RT than AL for incongruent eligible stimuli. The difference between incongruent and neutral RT differed significantly among sessions except between SL and AL. The difference between each session and BL for incongruent and neutral trials was significantly different from each other for SL and AL conditions only.

According to previous findings, bilinguals should be overall slower than monolinguals (Rosselli et al, 2002; Prior, 2012); however, they showed faster RT compared to monolinguals and sustained that pattern throughout the study which agrees with findings for different attentional tasks (Bialystok & Craik, 2010; Bialystok, Craik, & Luk, 2008). Bilinguals have long been the object of research for executive functioning and different studies have demonstrated that having two languages in the brain can result in better monitoring and inhibitory control (Prior, 2012, Morales et. al, 2013). Additionally, in the present task bilinguals could be using Spanish as a strategy to reduce RT for both incongruent and neutral trials without the need for suggestion. Language proficiency measures did not predict reduced Stroop interference but more analyses between proficiency and RT for incongruent trials could reveal the correlation between Spanish dominance and faster RT based on results from Suarez et. al (2014) and Tse & Altarriba (2012). Moreover, the difference in motivation for completing the study could also explain this unexpected result as bilinguals received a small monetary compensation for their participation and have limited opportunities to participate in studies compared to monolinguals that were completing the study for class credit.

Another important limitation for any language study is the method of assessing proficiency. Having different proficiency measures could serve as a better way to determining whether a participant is balanced or dominant in one of their languages. However, current tests present their own limitations; for instance, the BNT task used here is outdated and grammar tests might not be the most effective way to detect language dominance, although results for the BNT in Spanish showed correlation with scores for DELE and the BNT in English was correlated with scores for the MELICET. Self-rating through the LHQ also presented challenges and might not be a clear representation of proficiency. Moreover, bilinguals could be overtly introduced to the strategy of using their non-target language instead of using labels as a suggestion. The Stroop effect is an automatic robust measure of interference that is not easily reduced; therefore, even if the strategy was introduced directly it would only be effective if bilinguals can use the non-target language to map their responses.

The lack of interaction between group and session for the Stroop effect was also unexpected as the main hypothesis predicted that the SL condition would show reduced Stroop interference for bilinguals. A possible explanation is the lack of power of 54 bilingual participants for the desirable effect. Opposing the idea that bilinguals used the non-target language strategy early on, proficiency measures did not predict their performance on the Stroop task. The BNT ratio revealed that only 26% of bilingual participants were Spanish dominant and neither that nor their language preference correlated with a reduced Stroop effect. Age of acquisition and time in the US have not been analyzed with Stroop data. Other analyses between proficiency measures and Stroop data are needed and will be explored in the future.

Another possible confounding variable is participants using squinting or looking away in an effort to avoid the conflict with word reading. Preston and Lambert (1969) asked their participants to avoid squinting or looking from the corner of their eyes which was reported in the post-experiment questionnaire as a strategy some participants used in the present study. This strategy could help participants perform better but it could also hinder their performance by causing more fatigue; therefore, all participants should be instructed to avoid squinting and looking from the corner of the eye to achieve uniformity across participants in future studies.

Future Directions

For now, results have shown that bilinguals are faster and had less Stroop interference than monolinguals, but further analyses are still needed to draw stronger conclusions. The AX-CPT data still needs to be organized and analyzed. We hypothesize that better inhibitory control in the AX-CPT task will predict better performance on the Stroop task.

Another way to look at the progression of cognitive control would be to analyze the change within each session. The first third of trials could be used as baseline for the latter third of trials. This analysis could reveal how participants progressed within a same session and what that progression looked like as a group.

From the many proficiency measures collected, there might be better and more effective ways to rate bilingual proficiency and these differences could reveal important interactions with

Stroop data. We need to look at age of acquisition and find how the different measures agree with each other. We will run more correlation tests between proficiency measures and decide on how to use them to further analyze differences among bilinguals and contrast that to monolinguals.

The data still contain some participants that could be acting as outliers for the group and raising the mean RT. Preliminary analyses showed that removing participants that had mean RT over 2.5 SD from group mean for more than one session would reduced the overall mean by 10-30ms. Careful analysis of outliers could change later interactions and correlations for the group.

Acknowledgements

My sincere gratitude to Dr. Jorge Valdes Kroff who put me in charge of this project from its initial stages and guided me every step of the way allowing me to have autonomy and gain valuable research experience. And Dr. Lise Abrams who worked closely with us since the beginning of this project and shared her lab resources with us. Dr. Abrams' guidance was decisive for the long hours of statistical analyses and making sense of the data. I am also grateful to every research assistant that helped us test participants throughout the 3-semester testing period.

References

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, memory, and cognition*, *34*(4), 859.

Bialystok, E., & Craik, F. I. (2010). Cognitive and linguistic processing in the bilingual mind. *Current directions in psychological science*, *19*(1), 19-23.

Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends in cognitive sciences*, 16(2), 106-113.

Costa, A., Albareda, B., & Santesteban, M. (2008). Assessing the presence of lexical competition across languages: Evidence from the Stroop task. *Bilingualism: Language and Cognition*, *11*(1), 121-131.

Dalrymple-Alford, E. C. (1968). Interlingual interference in a color-naming task. *Psychonomic Science*, 10(6), 215-216.

Francis, W. S. (1999). Cognitive integration of language and memory in bilinguals: semantic representation. *Psychol. Bull*. 125, 193–222.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. Bilingualism: *Language and cognition*, 1(2), 67-81.

Grosjean, F. (1992). Another view of bilingualism. In R. Harris (Ed.), Cognitive processing in bilinguals (pp. 51-62). Amsterdam: Elsevier.

Heidlmayr, K., Moutier, S., Hemforth, B., Courtin, C., Tanzmeister, R., & Isel, F. (2014). Successive bilingualism and executive functions: The effect of second language use on inhibitory control in a behavioural Stroop Colour Word task. *Bilingualism: Language and Cognition*, *17*(3), 630-645.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). Boston Naming Test (2nd ed.). Boston: Lea & Febiger.

Kiyak, H. A. (1982). Interlingual interference in naming color words. *Journal of Cross-Cultural Psychology*, 13(1), 125-135.

Kroll, J. F., Dussias, P. E., Bogulski, C. A., & Kroff, J. R. V. (2012). Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In *Psychology of learning and motivation* (Vol. 56, pp. 229-262). Academic Press.

Lesh, T. A., Westphal, A. J., Niendam, T. A., Yoon, J. H., Minzenberg, M. J., Ragland, J. D., ... & Carter, C. S. (2013). Proactive and reactive cognitive control and dorsolateral prefrontal cortex dysfunction in first episode schizophrenia. *NeuroImage: Clinical*, 2, 590-599.

Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., & Kramer, A. F. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Cognitive Brain Research*, *12*(3), 467-473.

Morales, J., Gómez-Ariza, C. J., & Bajo, M. T. (2013). Dual mechanisms of cognitive control in bilinguals and monolinguals. *Journal of Cognitive Psychology*, 25(5), 531-546.

Naylor, L. J., Stanley, E. M., & Wicha, N. Y. Y. (2012). Cognitive and electrophysiological correlates of the bilingual Stroop effect. *Frontiers in Psychology*, 3, 1-18.

Portocarrero, J. S., Burright, R. G., & Donovick, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology*, *22*(3), 415-422.

Preston, M. S., & Lambert, W. E. (1969). Interlingual interference in a bilingual version of the Stroop color-word task. *Journal of Verbal Learning and Verbal Behavior*, 8(2), 295-301.

Prior, A. (2012). Too much of a good thing: Stronger bilingual inhibition leads to larger lag-2 task repetition costs. *Cognition*, *125*(1), 1-12.

Ramsay, M. C., & Reynolds, C. R. (1995). Separate digits tests: A brief history, a literature review, and a reexamination of the factor structure of the Test of Memory and Learning (TOMAL). *Neuropsychology Review*, *5*(3), 151-171.

Roelofs, A. (2010). Attention and facilitation: Converging information versus inadvertent reading in Stroop task performance. *Journal of Experimental Psychology*: Learning, Memory, and Cognition, 36(2), 411.

Rosselli, M. Ardila, A. Santisi, M., Arecco, M., Alvatierra, J., Conde, A., Lenis, B (2002). Stroop effect in Spanish–English bilinguals. *Journal of the International Neuropsychological Society*, **8**, 819–827.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662.

Suarez, P. A., Gollan, T. H., Heaton, R., Grant, I., Cherner, M., & Group, H. (2014). Second-Language Fluency Predicts Native Language Stroop Effects: Evidence from Spanish–

English Bilinguals. *Journal of the International Neuropsychological Society*: JINS, 20(3), 342–348.

Tse, C. S., & Altarriba, J. (2012). The effects of first-and second-language proficiency on conflict resolution and goal maintenance in bilinguals: Evidence from reaction time distributional analyses in a Stroop task. *Bilingualism: Language and Cognition*, 15(3), 663-676.

Tzelgov, J., Henik, A., & Leiser, D. (1990). Controlling Stroop interference: Evidence from a bilingual task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(5), 760.

Wang, R., Fan, X., Liu, C., & Cai, Z. G. (2016). Cognitive control and word recognition speed influence the Stroop effect in bilinguals. *International Journal of Psychology*, *51*(2), 93-101.

Appendix A – Post Experiment Questionnaire

Please try to answer the following questions as honestly as you can. Provide any extra information that you think is relevant and be as specific as possible. If you do not understand a question, please ask.

1. What do you think was the purpose of this experiment?

2. Were all of the questions of the same difficulty, or did you find that some of the questions were easier or harder to answer than other questions?

   a. If some were easier/harder, what made them easier/harder?

   b. For those that were harder, did you use any strategies to make the task easier?

3. Did you use any overall strategies to help you complete the task? If so, what strategies did you use?

4. (Spanish speakers) Which language(s) did you use to help you complete the task?

Appendix B – Language History Questionnaire

UF UNIVERSITY of FLORIDA

In what country were you born?

How long did you live in your country (in years)?

How long have you lived in the US (in years)?

Have you lived in another country besides your birth country (if applicable) or the US for longer than 6 months?

- Yes
- No

>>

UF UNIVERSITY of FLORIDA

Welcome! In this survey you will provide information on your experience as a bilingual/multilingual individual. Please fill out your responses as accurately as possible. The survey should last about 8 minutes.

Name

Sex

- Male
- Female

Age

Were you born in the US?

- Yes
- No

List the country and length of time (e.g France, 3 years). If more than one country, list each new country on a separate line.

At what age did you first begin to learn Spanish (**select 0** if from birth or indicate in years, e.g. 3)?

At what age did you being to learn English (**select 0** if from birth or indicate in years, e.g. 3)?

If you can hold a conversation in other languages besides English or Spanish, please list those languages here.

In general, which language do you **prefer** to use?

- Spanish
- English
- Both equally
- Other

In general, which language do you feel that you **SPEAK** better?

- Spanish
- English
- Both equally
- Other

In general, which language do you feel that you **UNDERSTAND** better?

- Spanish
- English
- Both equally
- Other

In general, which language do you feel that you **READ** better?

- Spanish
- English
- Both equally
- Other

In general, which language do you feel that you **WRITE** better?

- Spanish
- English
- Both equally
- Other

What language or languages do you speak at home?

- Spanish
- English
- Both
- Other

In what language(s) do you speak to your mother? (leave blank if not applicable)

- Spanish
- English
- Both
- Other

In what language(s) do you speak to your father? (leave blank if not applicable)

- Spanish
- English
- Both
- Other

In what language(s) do you speak to your sibling(s)? (leave blank if not applicable)

- Spanish
- English
- Both
- Other

In what language(s) do you speak to your closest friend?

- Spanish
- English
- Both
- Other

On a scale from one to 10 (1 = rudimentary proficiency, 10 = complete fluency), please indicate your level of proficiency in **ENGLISH** in speaking, listening, writing, and reading. Use the slider to leave your answer.

Least Proficient								Most Proficient	
1	2	3	4	5	6	7	8	9	10

Speaking

Listening

Writing

Reading

On a scale from one to 10 (1 = rudimentary proficiency, 10 = complete fluency), please indicate your level of proficiency in **SPANISH** in speaking, listening, writing, and reading. Use the slider to leave your answer.

Least Proficient								Most Proficient	
1	2	3	4	5	6	7	8	9	10

Speaking

Listening

Writing

Reading

On average, what percentage of the time are you **exposed to English, Spanish, or other languages** (your answers must total 100%)?

English	
Spanish	
Other	
Total	0

When choosing a language to speak with a person who is equally fluent in all of your languages, what percentage of the time would you choose to speak each language? (your answers must total 100%)

English	
Spanish	
Other	
Total	0

Codeswitching means using both Spanish and English in the same sentence when talking to someone else (e.g. *Pero no encontraron el car over there?*). Do you ever codeswitch?

- Yes
- No
- Not sure

How often do you codeswitch when you are speaking to another bilingual friend or family member?

Never	Rarely	Sometimes	Often	Always
1	2	3	4	5

How frequently do your bilingual friends and family members codeswitch with you during a conversation?

Never	Rarely	Sometimes	Most of the Time	Always
1	2	3	4	5

How frequently do your bilingual friends or family members interact with you through social media written with codeswitching (e.g. e-mail, twitter, instagram, etc.)?

Never	Rarely	Sometimes	Most of the Time	Always
1	2	3	4	5

Do you think that codeswitching is an important part of who you are?

- Yes
- No
- Not Sure