The Role of Executive Control in Bilingual Lexical Access

BY
JOANNA C. BOVEE
B.A., Lake Forest College, 2006
M.A., University of Illinois at Chicago, 2009

DISSERTATION

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Psychology
in the Graduate College of the
University of Illinois at Chicago, 2014

Chicago, Illinois

Defense Committee:
Gary E. Raney, Chair and Advisor, Psychology
Susan Goldman, Psychology and Learning Sciences
Viorica Marian, Northwestern University
Kara Morgan-Short, Psychology and Hispanic Studies
James Pellegrino, Psychology and Learning Sciences

ACKNOWLEDGMENTS

I would like to thank my dissertation committee – Gary Raney, Susan Goldman, Viorica Marian, Kara Morgan-Short, and Jim Pellegrino – for sharing their expertise and providing helpful feedback and suggestions. An extra round of thanks goes to Gary Raney, my advisor who has been extremely encouraging throughout my time at UIC.

I would also like to thank my undergraduate research assistants who helped my collect my data: Alessandra Arreola, Alvaro Cruz, and Gabby Suarez. Thanks to Dr. Vittorio Picco for the laughs that kept my spirits up and to Roberta Catherine for always watching over me.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF ABBREVATIONS

BL	Bilingual
CST	Color-shape task
L1	First language
L2	Second language
L3	Third language
LEAP-Q	Language Experience and Proficiency Questionnaire
PNT	Picture-naming task
PPVT	Peabody Picture Vocabulary Test
RHM	Revised Hierarchical Model
RT	Reaction time
RT-CR	Reaction time for correct trials only
SAS	Supervisory attentional system
SLL	Second language learner
TVIP	Test de Vocabulario de Imagenes Peabody
VF	Verbal fluency
VSNB	Visual-spatial n-back

SUMMARY

The purpose of the current study was to investigate the role of executive control components during bilingual lexical retrieval (i.e., word retrieval). Executive control components of interest included inhibition, updating of working memory, and shifting. The goals were to identify (a) which components play a role during lexical retrieval and (b) whether increased executive control ability is related to higher proficiency in a second language. To do this, I tested highly proficient English-Spanish bilinguals (BLs) and native English speakers who were current Spanish language learners (SLLs) on two switch tasks: one non-verbal (shape and color identification) and one verbal (picture-naming in English and Spanish). Additionally, participants completed a battery of tests that measured individual differences in executive control. More proficient bilinguals were predicted to exhibit greater executive control on the verbal switch task. In particular, they were expected to show more advanced use of the updating component than the less proficient bilinguals. Better executive control as measured by the individual differences battery was expected to be associated with more efficient processing during the verbal and nonverbal switch tasks.

Results showed very few differences between participants groups. SLLs and BLs had similar switching and mixing costs on the non-verbal and verbal switching tasks. Both participant groups had similar working memory and shifting skills. The bilinguals did show enhanced inhibitory skills compared to SLL. Measures of executive control did not correlate well with the two switching tasks, and therefore were not predictive of performance on the switching tasks. These data suggest that being bilingual does not always afford enhanced executive control. Using two languages frequently could lead to enhanced control of language processing, but this does not necessarily extend to the non-verbal domain.

I. INTRODUCTION

If you are English-Spanish bilingual and see smoke coming out of an apartment building, would you yell the English word “fire” or the Spanish equivalent “fuego”? When talking with a friend about the soccer match you watched over the weekend, do you call it a “game” or a “partido”? In each case words from both languages are activated in the mental lexicon, but only one can be produced at a time. As a result, every time bilinguals speak, they experience competition between their languages. Despite the competition, selecting words from the appropriate language may appear effortless. One potential reason the process is performed so easily is that bilinguals are thought to have enhanced executive control processes, such as the ability to control their attention when retrieving words from memory.

The purpose of the current study was to investigate the role of executive control components during bilingual lexical retrieval (i.e., word retrieval). Executive control components of interest include inhibition, updating of working memory, and shifting. The goals were to identify (a) which components play a role during lexical retrieval and (b) whether increased executive control ability is related to higher proficiency in a second language.

The current research has the potential to enhance previous research that shows bilinguals have cognitive advantages on a number of nonverbal tasks (e.g., Bialystok, 2008; Costa, Hernandez, & Sebastian-Galles, 2008; Emmorey, Luk, Pyers, & Bialystok 2008). Since the late 1990s, a substantial number of studies have focused on determining when or what type of tasks bilinguals demonstrate enhanced executive control, but less established is how bilinguals develop this advantage. The general hypothesis is that compared to knowing and using only one language, knowing and using two languages recruits more executive control relative to a single language when engaged in language processing (Bialystok, Craik, Klein, & Viswanathan, 2004;

Kroll & Dijkstra, 2002). For each semantic concept bilinguals want to verbally communicate, they must choose between two words before selecting the word for production. Executive control is needed to make the correct selection. In other words, bilinguals have an additional step in lexical access than monolinguals (Kroll & Dijsktra, 2002; Kroll & Stewart; 1994).

If bilinguals are exposed to and use their two languages enough, the resulting increased executive control in the verbal domain can spill over to the nonverbal skills as well (Bialystok, 2008). It is important to understand not only where the benefits occur in the nonverbal domain but also how they originate in the verbal domain, which is a topic that typically receives less attention. Additionally, over half the world's population speaks two or more languages. Much of the cognitive research thus far is based on a monolingual model of cognitive function. If monolinguals and bilinguals process information differently, the current view is incomplete. The current study will help to create a more complete understanding of general human cognition by exploring whether bilinguals develop enhanced control of their executive processes as a result of learning two languages.

In order to study how bilinguals might develop enhanced executive control, I had a group of English-Spanish bilinguals complete a variety of executive control tasks as well as a lexical retrieval task in which they had to resolve interference between their two languages. Participants were of high or beginning-to-moderate proficiency levels in Spanish. If executive control in general improves due to using two languages, the executive control ability should vary as a function of proficiency level. Based on previous research (Kroll & Stewart, 1994), bilinguals (BLs) who are proficient in both languages and can use both to the same extent should be more likely to have to constantly resolve the interference between their two languages. Emerging bilinguals, also known as second language learners (SLLs), will initially experience less interference as their L2 (second language)knowledge is limited and therefore has fewer opportunities to interfere with L1 (first language).

To preview the organization of this paper, I first provide a brief summary of executive control components and a model of cognitive control in bilingual lexical access. Next, I review some of the relevant research on bilingual lexical access. I will then identify where further research is needed and subsequently present my hypothesis, predictions, and methodology. Next, I state the results of the analyses and subsequently provide my conclusions and a discussion on the findings and further research.

Components of Executive Control

As described below, executive control has been divided into multiple components. Theories differ as to whether the components are distinct from each other (e.g., Miyake, Friedman, Emerson, Witzki, Howerter, & Wager, 2000) or whether there is a central system that controls subordinate systems (e.g., Baddeley, 1996; Norman & Shallice, 1986). Most important for my purpose is to understand how each component might be used in bilingual language processing.

The executive control system has been proposed as a composition of three basic mechanisms: shifting between tasks or mental sets, updating and monitoring of working memory representations, and inhibition of dominant or prepotent responses (Miyake, Friedman, Emerson, Witzki, Howerter, & Wager, 2000; Dempster, 1993; Harnishfeger, 1995; Nigg, 2000). I describe each of these below.

Shifting. The first proposed component of the central executive system is shifting, which has the role of disengaging attention from one task and engaging it in another task (Callejas, Lupiàñez, Funes, & Tudela, 2005; Monsell, 1996; Miyake et al., 2000). In a task switching paradigm, such as a Stroop-like task (Monsell, 2003) or switching between two languages, shifting is simply a requirement. Without the shifting mechanism, a person would become fixated on a single task or language and not be able to switch when necessary. It is possible that bilinguals are able to engage and disengage more efficiently than monolinguals, but because past research has not emphasized this component, I do not focus heavily on it in the current study.

Updating. Updating includes the revision and monitoring of working memory representations, such as task goals and memory traces, (Miyake et al. 2000; Callejas et al., 2005). When new information is received by the executive control system, this mechanism acts to monitor and code incoming information as relevant or irrelevant to the task at hand. It then allows applicable information to enter into the working memory processing space (Miyake et al., 2000; Koch, Prinz, & Allport, 2005; Rubin & Meiran, 2005; Smith & Jonides, 1999). This mechanism is important for keeping distractor items from reaching working memory where they could interfere with the relevant information. Additionally, the updating component evaluates the current contents of working memory and decides whether information already receiving attention is still needed to complete the current task or goal (Morris & Jones, 1990). If the information is no longer needed, it is not refreshed and quickly decays from working memory.

Despite its importance in executive control, updating is not always successful, used, or needed in cognitive processing. Sometimes interfering information may enter working memory and interference among items occurs. How or when this failure or lack of updating occurs is unclear. For example, if the updating component were always 100% correct there would be no need for additional mechanisms to deal with competing information (e.g., inhibitory control discussed in the paragraph below). Updating is relevant to bilingual speech production in that bilinguals need to keep track of which is the to-be-used language for the task at hand and be able to focus attention on that language or set of words from that language. For the task at hand, updating would be needed to identify the goal language in which to name an item, but does not always go as far as to identify the target word in the goal language. When updating of working memory is not in control, inhibitory control is likely to assume control.

Inhibition. The third proposed component of the central executive system is inhibition, which is the active suppression of irrelevant information (as determined by the goal in working memory) that then allows attention to remain on the current task (Friedman & Miyake, 2004; Harnishfeger, 1995). Inhibition is a mechanism that can be used when competition occurs among items in working memory and acts to select a target item by reducing the activation, or accessibility, of a non-target item. When word choices from both languages are active in a bilingual person's working memory, inhibitory control could be used to make the word from the nontarget language less active, which would facilitate selecting the word from the target language for retrieval because the word in the target language would have greater activation.

Executive Control in Bilingual Lexical Access

A common assumption in bilingual word selection research is that both of a bilingual's languages are activated in parallel during language processing. Several studies support the conclusion that when identifying a conceptual representation, word choices from both languages are made available and thus allow for the possibility of competition between lexical items during selection (e.g., Colomé, 2001; Costa, Miozzo, & Caramazza, 1999; Hermans, Bongaerts, De Bot, & Schreuder, 1998).

If there is parallel activation, how do bilinguals choose words from the target language? One account of bilingual lexical access that is heavily cited is the Inhibitory Control (IC) model (Green, 1998). Green proposes that the regulation of the bilingual lexico-semantic system is much like the regulation of any other action; specifically, the executive control system has multiple levels of control that play a role in choosing the target word. The IC model is based on Norman and Shallice's (1986) model of executive control, which consists of different levels of control for routine and non-routine behavior. Routine behavior is directly controlled by what is called a contention scheduling system that recruits schemas, which contain rules and goals for the task at hand, to guide behavior. Non-routine behavior is guided directly by the supervisory attentional system (SAS). Green suggests that bilingual lexical access control can occur at two levels. The first is the language-level and occurs through the use of schemas in which the target language is identified. The second is at the lexico-semantic level (word level) in which words from both languages (translation equivalents) compete for selection, but only one can be chosen based on the target language. The lexico-semantic level requires more direct control from the SAS. Depending on whether regulation occurs at the language level or word level, different components of executive control might be used.

As described by Green (1998), the first possible form of executive control in bilingual lexical access comes at the language level when a language schema is selected. A language schema can be thought of as a set of production rules used to achieve the goal of producing words in a specific language. The language schema only activates production rules for the target language, and therefore is referred to as language-specific control. In this manner, language selection is controlled because the nontarget language schema is ignored, and, subsequently, the nontarget language and its lexical items, although activated to a small degree, receive little to no attention. Once the chosen language schema has been activated, the selection mechanism then evaluates lexical node competitors within the target language and chooses the best match for the conceptual representation. The language schema will remain active in working memory until a) the goal is achieved, b) the goal changes, or c) another language schema is activated. Executive control of this type is in line with an updating/monitoring account of control as it is based on the concepts of complex goals controlling behavior (Miyake, et al. 2000; Nigg, 2000) and is more likely to be found in highly proficient bilinguals for whom speaking in both languages is an relatively automatic, routine process.

Green (1998) also proposes that control of lexical access can occur at the word-selection level. This occurs especially when the action is still considered non-routine, as it might be for emerging bilinguals. When a bilingual person identifies the semantic concept he or she wants to communicate, possible word choices are activated in both languages. Because more than one possible word choice could accurately represent the concept, the executive control system must evaluate the language tags associated with each word item and subsequently suppress the competing item(s) from the nontarget language so that the target item's activation is relatively higher than the non-target(s). This is selection through inhibition. Control through inhibition is considered language non-specific in that both languages compete for selection (Costa & Santesteban, 2004; Costa & Caramazza, 1999).

How does Green's proposed model of the regulation of the bilingual lexico-semantic system fit in with more general theories of cognitive control? Clearly, it fits well with Norman and Shallice's model (1986) as it is the foundation of the Inhibitory Control model. Green's language task schema selection fits well with Miyake et al.'s (2000) updating component and lexical selection fits well with Miyake et al.'s inhibition component. Both of these components of executive control might be used in lexico-semantic access and perhaps we could see control exerted using either of these control processes at different stages of the selection process.

To add to the complexity, individual differences outside the domain of executive control, such as language proficiency, could moderate the way in which words are retrieved and the processes used to access words. Language proficiency changes how a word is associated to its conceptual representation, and thus could affect the way in which the word is retrieved. Kroll and Stewart's (1994) Revised Hierarchical Model (RHM) of bilingual language production can be used to explain the evolution of second language (L2) word production from low proficiency to high proficiency in the L2. According to the RHM, when a second language learner acquires new words in the L2, the words will be conceptually mediated by the first language (L1). In other words, to produce a word in L2, the speaker first selects the L1 lexical item, which is linked directly to the conceptual representation for the target concept, and then translates it to the L2. As proficiency increases, the associative link between L2 and the conceptual level is strengthened and eventually translation through L1 is not necessary.

For non-proficient bilinguals, the L1 mediation would eliminate the need to select the L2 target word in a language-specific manner because the L2 word is not competing for access. The L1 word would need to be activated before the corresponding L2 word could be activated. This dual-language activation will result in competition between the two words, which, according to Green's IC model (1998), will be resolved through inhibition. Thus, lexical retrieval would be considered language-nonspecific as target word candidates come from both languages.

For highly proficient balanced bilinguals, direct lexical retrieval is possible for both languages. This opens up the opportunity for control to be at the language-level as L2 is no longer dependent on L1. In the beginning stages of lexical retrieval, proficient bilinguals could create a goal to access the target language and activate the appropriate language schema. Updating of working memory would then be essential as a monitor to ensure the current language-schema matches the current word-selection goal.

Identifying Executive Control Processes at Work

How do we identify when a particular component of executive control is being used? To separate inhibition from updating of goals in working memory, simple paradigms that involve switching between two tasks can be used. A task-switching paradigm is a common method used to study how the executive system accomplishes disengaging from one task and engaging in a different task (see Appendix H for a list of key terms and definitions related to switch tasks). Because two different tasks are constantly being used, task switching is a suitable method for looking at how the executive control system handles competing information. For example, in a Stroop-like paradigm, participants switch between a color naming task and a word-naming task. Regardless of the task used, there will always be a cost associated with switching between tasks, which results in slower reaction times and lower accuracy for trials following the switch. For example, when two successive trials require color naming (task repetition), the second trial will be faster than when a color naming trial is preceded by a word naming trial (task switching).

Monsell (2003) attributes the additional time needed to switch tasks to both persisting inhibition of the previous task-set's activation and the time to reconfigure the goal statement. Each of these can be linked to the executive control components necessary to make the switch. At the beginning of a trial, the updating component brings into working memory the target, the task goal, and the task-set, which includes the possible answers. If there is competing information, either persisting from a previous trial or caused by an aspect of the current trial, inhibition will be used to suppress the irrelevant information. Once the response is selected and made, a new trial begins. If the updating component identifies that there is a new goal, the shifting component is used to disengage from the prior task and direct attention to the production rules for the new task. If the new trial's task is a repetition of the previous trial, no shifting between task sets will be needed and no costs will be added.

The variables of interest for the task switching paradigm are switching costs and mixing costs. To measure these costs, reaction times are needed from two types of task blocks. The first is a pure-block task in which participants need to select and make the correct response for the task, but there is no variation in the task type. For example, participants might be shown shapes, such as circles or squares, and only need to identify the shape. Reaction time for these trials (single-task trials) acts as a baseline measure for simple response time. The second block is a mixed-block task in which participants will have to perform one of two tasks during each trial. Typically, the participants receive a cue for which task to perform immediately before the stimulus onset or simultaneously with the stimulus onset. For example, participants might be shown shapes, such as circles or squares that are displayed in blue or red, and are cued before each trial to identify the shape or the color. Two different trial types can occur during mixed-blocks. The first is a repetition trial which has the participants repeat the task completed on the preceding trial (e.g., trial n-1 = name shape, trial 1 = name shape). The second trial type is a switch trial, which requires the participants to switch to the other task set. That is, the task set does not match the immediately preceding one (e.g., trial n-1 = name color, tria1 = name shape).

By examining reaction time differences between trial types, it is possible to distinguish the source of the increased response time. Switching costs are represented by the difference between repetition trials and switch trials in the mixed-block task. Mixing costs represent the difference between repetition trials in mixed-blocks and single-task trials in the pure blocks. Switching costs are purported to be an indicator of inhibition whereas as mixing costs are thought to reflect the updating of working memory (Koch, et al., 2005; Prior & MacWhinney, 2010; Rubin & Meiran, 2005). The logic is as follows: switching costs will reflect the transient-interference from the response sets on a previous trial. Inhibition will be used to suppress the irrelevant information so that that target information can be selected. Mixing costs, however, will reflect the costs associated with the possibility of switching tasks. In the pure block, the goal is always the same (e.g., name the shape). In the mixed block, the participants do not know what the task will be on the upcoming trial until the trial begins and a cue is presented. This "unknowing" recruits executive control in the form of updating of goals. This additional recruitment and use of executive control will increase reaction time.

In summary, a task-switching paradigm with a pure block and a mixed block makes it possible to separate switching costs from mixing costs. Switching costs are associated with inhibitory control and mixing costs are associated with the updating of goals in working memory.

Task-Switching in Bilinguals

The research described thus far has referred to domain nonspecific, general task switching. The current research used a verbal version of task-switching, language-switching, to explore the role of executive control processes during lexical retrieval. Before describing the study, I will describe a few recent studies that provide the basis for my study. My study will replicate and extend upon findings from previous research that used task switching to examine executive control in bilinguals.

As previously mentioned, both inhibition and updating components of executive control have been proposed as the key mechanisms that allow bilinguals to resolve competition between their two languages (Bialystok, 2010; Costa & Santesteban, 2004). In terms of language production, the use of inhibition can be inferred when the cost to switch to the dominant language from the less dominant language (typically switching from L2 to L1) is greater than when switching to the less dominant language from the more dominant language (typically switching from L1 to L2). This is called an asymmetrical switch cost (Calabria, Hernández, Branzi, & Costa, 2011; Costa & Santesteban, 2004; Meuter & Allport, 1999) and occurs because greater inhibition is needed to inhibit the dominant language as it is generally more active than the less dominant language. Returning to the dominant language requires more effort to "reactivate" it for selection. If bilinguals are highly proficient in both of their languages, there should be symmetrical switch costs between languages because the resting activation levels of each language should be comparable.

Costa and colleagues (e.g., Costa & Santesteban, 2004; Costa, Santesteban, & Ivanova, 2006) have conducted several studies with bilinguals completing verbal switch tasks. In their research, the stimuli were pictures and for each trial the participants must name the object represented by the picture (e.g., a table). Target language was visually cued and switched randomly across trials. Costa and colleagues found that less proficient bilinguals had asymmetrical switch costs, whereas highly proficient bilinguals did not. They interpreted this as evidence for inhibitory control in less proficient bilinguals and evidence for goal monitoring in highly proficient bilinguals. No pure blocks were included their studies, therefore they could not measure mixing costs as a more direct measure of goal monitoring. As a result, my own view is that their conclusion that highly proficient bilinguals used goal monitoring is weak. The lack of asymmetric switch costs alone does not strongly support the lack of inhibition because it is possible that inhibition was used to switch between both languages but the amount of inhibition was similar in each switch direction.

Prior and MacWhinney (2010) were among the first researchers to directly examine *both* the role of inhibition and updating in bilinguals and monolinguals. They used a nonverbal task-switching paradigm that required participants to either identify the shape (circle vs. triangle) or the color (red vs. green) of visual stimuli. The bilinguals in Prior and MacWhinney's study demonstrated a significant advantage when switching between tasks, which are thought to be linked to transient proactive interference and attenuated by enhanced inhibitory ability. In contrast, bilinguals showed no advantage for mixing costs, which are thought to be linked to updating of working memory. Prior and MacWhinney also evaluated some individual differences measures, such as working memory (O-Span) and English vocabulary knowledge (Peabody Picture Vocabulary Test [PPVT]), but there were no differences between bilinguals and monolinguals on these measures. A possible explanation for the lack of mixing cost differences could be the exclusive use of a non-verbal task. Because Prior and MacWhinney compared monolinguals to bilinguals, they could not use a language switch task because monolinguals do not have a second language. One might expect bilinguals to show an advantage on a task that is more related to where the cognitive advantage is proposed to have originated, which is language use.

Soveri, Rodriguez-Fornells, and Laine (2011) examined individual differences in executive control in switch tasks using Swedish-German bilinguals. They collected several measures of executive control ability and self-report information about how often their participants switched between their two languages when speaking to other Swedish-German bilinguals outside of the experiment. They found that bilinguals who reporting switching between languages the most often were more likely to show lower mixing costs on a switching task than those who switched between languages less frequently. This result demonstrates a relationship between how often one switches between languages and mixing costs, but no directionality (causality) can be inferred from this study. Had Soveri and colleagues included adequate language proficiency measures and used bilinguals of varying proficiency, they could have examined whether the degree of bilingualism is associated with the size of mixing costs, and consequently, updating of working memory.

Prior and Gollan (2011) compared performance across two groups of bilinguals (Spanish-English and Chinese-English) and monolinguals on a nonverbal switch task. The Spanish-English bilinguals were more proficient in English than the Chinese-English participants. The researchers collected several individual differences measures, including language background information, verbal fluency, the Shipley vocabulary test, and the matrices subtest of the k-bit intelligence test. On the nonverbal switch task, participants had to either identify the shape (circle vs. triangle) or color (red vs. green) of a visual stimulus item. Spanish-English bilinguals showed a small advantage in reduced switching costs as measured by reaction time compared to Chinese-English bilinguals and monolinguals, but showed no advantage for mixing costs as there was no significant difference in this measure across the groups. Results were similar for the verbal task, which required participants to name digits, with the Spanish-English bilinguals again exhibiting slightly smaller switch costs and comparable mixing costs relative to the Chinese-English bilinguals (the monolinguals could not complete this task). There were, however, asymmetrical mixing cost differences as a function of language dominance, with the more dominate language incurring more mixing-costs. Although Prior and Gollan used a verbal switch task (naming a digit), it could be thought of as recognition task rather than a recall task based on the argument that number identification is more of an automatic process compared to naming pictures of objects. If this is correct, the task would be relatively easier to complete compared to a picture-naming task and therefore require less cognitive control.

In summary, results from prior research are inconclusive and the main question remains unanswered. What executive control mechanism or mechanisms guide(s) lexical access in bilinguals? However, this recent research provides a foundation for a study with improved methodology and more a focused purpose.

The purpose of the current research was to extend prior research by investigating switching and mixing costs in language and non-language switch tasks using bilinguals of different proficiency levels. To better understand the contribution and hypotheses, I need to summarize the intended methodology. Half the bilinguals will be highly proficient in both languages and half will be English-dominant with limited Spanish proficiency. Both bilingual groups will complete a nonverbal and a verbal switch-task and several executive control tests. I will measure switching and mixing costs in both tasks and analyze the patterns in the data while controlling various individual difference measures. The outcome will provide a more precise test of the role of updating and inhibition in language selection in bilinguals and a more controlled evaluation of whether individual differences in executive functions influence switching and mixing costs.

The current study adds to previous research on bilingual switch costs in several aspects. To begin, it is the first study that compares switching and mixing cost across bilinguals of *different* proficiency levels. High-proficiency bilinguals are proficient in both English and Spanish and low proficiency bilinguals are highly proficient in English, but still learning Spanish. Participants first completed a shape/color identification switch task (similar to that of Prior & MacWhinney, 2010, and Prior & Gollan, 2011) during which switching costs and mixing costs were measured. Additionally, it is the first study to measure both switching *and* mixing costs on a bilingual picture-naming task. In this task, participants were shown a picture of an object, such as a key, along with a language cue, and had to name the object as quickly as possible in the cued language. Naming times were collected for pure blocks that only tested naming time in one language and in mixed blocks in which the target language switched across trials so that objects had to be named in either Spanish or English. Furthermore, language proficiency based on vocabulary knowledge was measured using methods that allow for easy and direct comparison across languages. Lastly, an extensive battery of executive control tasks was given so that I could conduct a detailed analysis of the relationship between each executive control component (shifting, working memory, and inhibitory control) and performance on the language-switch task.

Results from this study could help to clarify the conflicting research regarding executive control mechanisms during bilingual lexical access. Whether the key mechanism that resolves competition between a bilinguals two languages is inhibitory control, updating of working memory, or a combination of the two, it could lead us one step further to understanding in what ways bilingualism is associated with enhanced executive control. Additionally, it will add to our overall knowledge of how cognition works in not just monolinguals, but to a more comprehensive view that includes bilinguals.

Hypotheses

Switching (Non-Verbal) Task. Two predictions were made for the color-shape task.

Hypothesis 1A: switching costs. I predicted that results should be similar to those found in prior research as this task replicates research by Prior and MacWhinney (2010) and Prior and Gollan (2011). That is, more proficient bilinguals will have reduced switching costs compared to less proficient bilinguals. Additionally, there should be no difference as a function of task – neither proficient or less-proficient bilinguals will show asymmetrical switching costs between task sets. This will be shown by the following pattern of results: the difference in mean reaction time between non-repetition (switch) and non-switch (repeat) trials in the mixed block will be similar for both low- and high-proficiency bilinguals. Additionally, there will be no difference in reaction time between these two trial types as a function of task type (color vs. shape identification).

Hypothesis 1B: mixing costs. If there are any differences in mixing costs between the two groups, the more proficient bilinguals should show smaller mixing costs than the less proficient bilinguals. This will be shown by the following pattern of results: the difference in mean reaction time between all trials in the pure block (single-task trials) and the repetition trials in the mixed block (repeat trials) will be smaller for the proficient bilinguals than the less proficient bilinguals.

Picture-naming (Verbal) Task. Two predictions were made for the picture-naming task.

Hypothesis 2A: switching costs. Both low- and high-proficiency bilinguals will exhibit switching costs. For the low- proficiency bilinguals, these costs will be asymmetrical such that switching to L1 is more costly than switching to L2 and requires a greater amount of inhibitory control. For the high-proficiency bilinguals, there will be no difference in switch costs size as a function of language as both languages should require the same amount of inhibitory control. This will be shown by the following pattern of results: a switch cost will be evidenced by longer mean reaction times for switch trials than repeat trials in the mixed block. For the less proficient bilinguals, switch costs will be larger when the switch is from Spanish to English than vice versa. For the proficient bilinguals, the direction of the switch will not matter.

Hypothesis 2B: mixing costs. High-proficiency bilinguals will have smaller mixing costs than low-proficiency bilinguals. This will be shown by the following pattern of results: the difference in mean reaction time between all single-task trials in the pure block and the repeat trials in the mixed block will be smaller for the high-proficiency bilinguals than the low-proficiency bilinguals.

Individual differences in executive control. Assuming my measures of individual differences in shifting, working memory, and inhibitory control are valid, better performance on these individual differences tasks should be associated with smaller mixing costs (working memory) and smaller switch costs (inhibitory control) on the nonverbal switch task. Switching ability should account for a significant portion of the remaining variance in performance on the switch task.

Hypothesis 3: switching costs. If general inhibitory control is related to the inhibitory control used in language processing, and my measures of inhibitory control are valid, performance on tests of inhibitory control should highly correlate with both the verbal and non-verbal switching tasks. That is, higher performance on the individual difference measures should be associated with smaller switch costs on the verbal task. If individual differences in inhibitory control are not associated with inhibitory control in the verbal task, then it is unlikely that the advantages bilinguals have shown on inhibition tasks in past research solely comes from the use of inhibition in bilingual lexical access.

Hypothesis 4: mixing costs. If the updating of working memory is related to how goals are identified and actualized is related to bilingual lexical access, then measures of individual differences should be highly correlated with mixing costs on the verbal switch task. That is, smaller mixing costs should be associated with better updating ability. If individual differences in updating are not related to updating in the verbal task, then it is unlikely that any advantages the bilinguals might show in goal monitoring is unrelated to the use of goal monitoring in bilingual lexical access.

Hypothesis 5: If advantages in executive control are related to bilingualism, differences in the working memory and the inhibitory control tasks should be associated with language proficiency. That is, high-proficiency bilinguals should perform better on the working memory and inhibitory control tasks than low-proficiency bilinguals.

II. Methods

Participants

Data from 42 participants were collected. All participants were 18 or older (*M* = 20.4, *SE* = 3.1). Twenty-one participants were native English speakers who were also beginning-to-intermediate Spanish Language Learners who were currently or had been recently (within the last six months) enrolled in a Spanish language class (males = 6; females = 15). This group had at least two years of high school Spanish or two semesters of college Spanish. The other 21 participants were Spanish-English bilinguals who had been exposed to both languages before the age of six and had continued to use both languages up to the time of the study (males = 3, females = 18).

Language-related Individual Differences Measures

Language history questionnaire. All participants completed the Language Experience and Proficiency Questionnaire (LEAP-Q; Marion, Blumenfeld, & Kaushanskya, 2007). This questionnaire consists of questions relating to language acquisition, proficiency, exposure, environment, etc. (see Appendix A). While information was gathered on all of the languages the participants may speak, the two languages of interest for the questionnaire were English and Spanish. The LEAP-Q is a good self-report measure of language background and proficiency, as it has been normed (including on English-Spanish bilinguals) and reliably predicts relationships between self-report measures and behavioral measures of language proficiency.

Lexical knowledge. English and Spanish proficiency measures focused on participants' word knowledge. Two tasks, each with an English and a Spanish portion, were completed at the end of the session. I focused on lexical knowledge because my verbal task was a lexical task.

Syntactic knowledge was not as critical as it would be in a task that required sentence-level processing.

Productive vocabulary – semantic (category) and phonemic (letter) fluency. The participants' productive vocabularies were measured using a verbal fluency test. The purpose of this test was to provide an individual difference measure of word production in both English and Spanish. The expressive vocabulary test consisted of two parts: naming category exemplars (semantic fluency) and naming words that start with a particular letter (phonemic fluency). Restrictions for responses included no repeated words or close variants of a word (e.g., train and trains), no proper nouns (e.g., a country or a person's name), and no numbers. The instructions and test letters and categories were presented using the E-prime software. Participants had 60 seconds to name aloud as many exemplars as possible. Timing was kept via the software and participant responses were recorded both in written form by the experimenter and auditory form using a voice recorder.

Categories and letters were chosen from Gollan, Montoya, and Werner's (2002) study on verbal fluency in Spanish-English bilinguals. Gollan et al. tested participants on 10 possible letters and 12 possible categories. For the purposes of the current study, I excluded two of their categories, Countries and Countries in Europe, as often the English and Spanish equivalents are cognates. This left 10 possible categories, four of which were used in the current study. The four test letters and test categories for the current study were chosen based on response rates reported by Gollan et al. for naming only in Spanish and naming only in English. In each language set, the items with highest number of correctly produced responses were compared. The first highest four overlapping letters and categories were chosen for this study. The letters tested were P, C, R, and D. The categories tested were Animals, Clothing, Colors and Fruits.

Each participant named exemplars for half (2) of the letters and half (2) of the categories in English and the other half in Spanish. Language order, naming type (letter vs. category), and target letter/category were counter-balanced across participants.

Receptive Vocabulary – picture identification task. The participants' receptive vocabularies were measured using standardized picture identification tests. The purpose of these tests was to provide an individual difference measure of vocabulary in both English and Spanish.

The Peabody Picture Vocabulary Test (PPVT-IV; Dunn & Dunn, 2007) and the Test de Vocabulario de Imagenes Peabody (TVIP; Dunn, Padilla, Lugo, & Dunn, 1986) were chosen because of their use in previous bilingual studies (Luo, Luk, & Bialystok 2009; Marian et al. 2007) and because of their reliability as standardized measures.

In both language versions of the untimed tests, participants viewed a set of four pictures/drawings and identified the one that best represents the word given to them by the experimenter (see Appendix B for example stimuli). The total raw score for each language was calculated using the guidelines specified by the publisher.

Non-verbal Individual Differences Measure

General intelligence – Raven's Progressive Matrices. General Intelligence was assessed using an untimed abbreviated version of Raven's Advanced Progressive Matrices (Raven, Raven, & Court, 1998). Participants were shown an image of a pattern with a piece cut out of it and were given eight possible pieces that could complete the pattern going both across and down (see appendix C for example stimuli). To indicate their responses, participants entered the number on the keyboard (1-8) that corresponded to their answer choice. Participants first completed two practice items with the experiment (Set 1: Items 1 and 2) before completing 18 experimental trials on their own (Set 2: even numbered items). The full set was not used due to time contraints. Previous research has shown that performance on an abbreviated set strongly correlates well with performance on all 60 items (Bilker, Hansen, Brensinger, Richard, Gur, & Gur, 2012).

Experimental Tasks

Nonverbal: color-shape identification. This task was based on the nonverbal switching task used by Prior and MacWhinney (2010) and Prior and Gollan (2011). For this task participants had to identify either an item's shape (circle vs. triangle) or color (red vs. blue) via button presses (see Appendix D for stimuli).

Simple blocks. Participants began the task by identifying items in a series of simple blocks. The blocks were considered simple as trials within a block only had one goal: shape identification *or* color identification. Half the blocks required the participants to identify the shape and the other half required the participants to identify color.

Each simple block consisted of 48 trials in random order. Order of simple block type was counterbalanced across participants. Stimuli were presented using E-Prime and responses were made using response box. The button on the left had a picture of a red triangle and the button on the right had a picture of a blue triangle. The same two buttons were used to indicate responses in each task set.

Each trial started with a black fixation cross presented in the center of the white screen for 350 ms, followed by a blank screen for 150 ms. The stimulus item and a cue appeared at the same time in the center of the screen and remained on the screen for 4000 ms or until the participants responded. The cue was used in the simple blocks to keep stimuli presentation consistent with that of the mixed block. The cue was either a horizontal sequence of three black squares for shape identification or a rainbow for color identification.

Mixed blocks. Participants identified the stimulus item as specified by a cue located above the item. The order of the task sets was pseudo-randomized such that participants would not know what the target task for a trial was until the stimulus and cue were presented simultaneously. There were never more than four consecutive trials of the same type and there was always an equal number of switch and non-switch trials. The procedure for each trial was the same as in the simple blocks, but the identification task switched pseudo-randomly. There were 48 trials per block and 3 mixed blocks in total (144 total mixed trials).

Presentation order. The presentation of blocks was a sandwich design such that participants first completed two simple-blocks, one of each task type; 3 mixed block trials; and then the two simple-block trials with the order reversed (e.g., Block 1: color, Block 2: shape, Block 3: mixed, Block 4: mixed, Block 5: mixed, Block 6: Shape, and Block 7: color). This sandwich design was based on Prior and MacWhinney (2010) and used to minimize order and practice effects.

Verbal: picture-naming task. For this task, 24 pictures were selected from the International Picture Naming Project database (Bates et al., 2003). This database consists of 512 images and provides norms for image naming times, image characteristics, and word characteristics across seven languages, including English and Spanish (see Appendix E for stimuli). The following selection criteria were used to select items from the object stimuli pool.

1. All items were nouns.
2. Limited phonological similarity of name with other items in the database (i.e., no intra-linguistic or cross-linguistic homophones, rhyming words, cognates, false cognates).

3. All items were high frequency words and were matched across language on word frequency. The average log transformed frequency from the Celex Lexical database (Baayen, Piepenbrock, & Gulikers, 1995) was 5.04 in English and 5.07 in Spanish.
4. Items were matched as closely as possible across languages on number of syllables. On average, English words had 1.33 syllables and Spanish words had 1.96. There was never more than a difference of one syllable between the object names in each language.

Pre-test. Before starting the task, participants were shown the images via a PowerPoint presentation. Participants first viewed the items with their English names written below the image and then viewed the Spanish set or vice versa. After studying each set of images and names for up to two minutes, the participants were given a quiz in which only the pictures were presented. The participants wrote down the name of item below the item. If participants could not recall a name, the experimenter said it aloud and had the participants repeat it. Once the participants completed the study and test for one language, they repeated the process in the other language.

Simple Blocks. Participants began the task by naming items aloud in a series of simple blocks. The blocks are considered simple as only a picture was shown and the target language remained the same within a block. For example, participants were instructed to name all the pictures in the block in Spanish (as cued by a green frame around the image) or all in English as (cued by a blue frame around the image).

Each simple block consisted of 48 trials with each item presented twice, in pseudo-random order. That is, participants named all items once before any items were presented the second time. Order of simple block type was counterbalanced across participants. Before the presentation of the first trial in the block, an instruction screen informed the participants of the target language.

Each trial started with a fixation cross presented in the center of the screen for 1000 ms. The stimulus item and cue then appeared simultaneously in the center of the screen and remained on the screen for 2000 ms. A blank screen was then presented for 500 ms before the beginning of the next trial.

Mixed Blocks. Participants then identified the stimulus item as specified by the cue located above the item. The procedure for each trial was the same as in the simple blocks, but the target language switched pseudo-randomly. There were never more than four consecutive trials of the same time, and there was an equal number of switch and non-switch trials. There were 48 trials in a mixed block and 3 mixed blocks in all (144 mixed trials total).

Presentation Order. Like the color-shape task, the presentation of blocks was a sandwich design such that participants first completed two simple-blocks, one of each language, 3 mixed block trials, and then the two simple-block trials with the order reversed (e.g., Block 1: English, Block 2: Spanish, Block 3: mixed, Block 4: mixed, Block 5: mixed, Block 6: Spanish, Block 7: English).

Non-Verbal Executive Control Tasks

A series of six executive control tasks were completed by the participants. Each task was designed to tap into a specific executive processing ability, was nonverbal in nature, and has been used in previous bilingual research. Nonverbal tasks were chosen to avoid any differences in performance due to language proficiency between subjects. Tasks were presented via E-Prime 2.0 and responses were made on the E-Primes serial response box (SR-box).

Updating of working memory.

Auto-Symmetry Span. The Automated Symmetry Span task was the first measure of working memory ability and was administered using standard procedures (Unsworth, Redick, Heitz, Broadway, & Engle, 2009; see Appendix F for sample stimuli). Participants had to remember and recall sequences of squares (storage component) within a 4 X 4 matrix while completing a symmetry judgment task (processing component). The symmetry judgment portion required participants to look at an 8 X 8 matrix with a mix of white and black squares. They had to judge whether the image was symmetrical across the vertical access. The pattern was symmetrical half of the time. After they indicated their responses via a mouse click, they were immediately shown the 4 X 4 grid with one red square for 650 ms and had to remember the location of the red square. After seeing a series (2-5 items in length) of the symmetry judgment images and to-be-recalled red squares, participants were given a For the recall portion, participants had to click on the cells in an empty 4 X 4 grid for the recall portion. They had to click on the appropriate square to indicate the sequence of the red squares. Participants completed a training portion before beginning the experimental trials to ensure they understood the task. There were a total of 12 sets of trials with three of each set size (3, 4, 5, and 6) and participants were not aware of the set size until they were asked to recall the sequence of the red squares.

Visuospatial N-back task. To test participants' ability to update working memory, a visuospatial n-back test was completed. The task required the participants to remember (store) the location of the stimulus from one trial and compare it to the location of the stimulus from a previous trial (processing). The task was a replication from a recent bilingual study by Soveri et al. (2011) and originally adapted from Carlson, Martinkauppi, Rämä, Salli, Korvenoja, & Aronen

(1998). For this task, participants saw a black square presented in one of eight possible locations centered around a black fixation cross in the middle of a white screen (see Appendix G for example stimuli). The task was to indicate whether the currently displayed square was in the same location as the previous trial (N-1 trial) or the trial before the previous trial (N-2). For 1-back trials, participants were to press the left button (labeled "same") whenever the square occurred in the same location as the previous one (target trial). If the square appeared in any other location participants were to press the right button (labeled "different"). In the 2-back task, participants were to press the "same" button whenever the square is in the same location as the one presented two trials back. If in any other location, participants were to press the "different" button.

The 160-trial task was broken into 8 blocks consisting of 20 trials each. Within a block, trials were either 1-back or 2-back, but not both. The participants were told before beginning each block what was the comparison. Within a block, there were 6 target trials (in which participant should respond yes; the square was in the same location for both the n trial and the n-back trial) and 14 nontarget trials (in which participant should respond no, the square was not in the same location for both the n trial and the n-back trial). Order of the trials was pseudo-randomized and no response was made on the first trial of each block for n-1 blocks (because there was no comparison trial) and no response was mode on the first two trials of the n-2 blocks).

Each block started with a screen displaying the number 1 or the number 2 in the middle of the screen for 5000 ms to indicate the appropriate response for the sequence. Next, a fixation point appeared in the middle of the screen for 1500 ms followed by a square that appeared in one of 8 places (upper-left, upper-center, upper-right, middle-left, middle-right, bottom-left, bottom-center, and bottom-right) on the screen for 1500 ms. The participants had to respond during this time. After this time period, the next trial started regardless of whether the participants had responded or not. Each sequence was followed by a 15-second break. Following Soveri et al., the block order of the sequences were 1-back, 2-back, 2-back, 1-back, 2-back, 1-back, 1-back, 2-back for a total of 8 blocks. Participants received a 30-second break after the first four blocks have been completed.

Key data for this test were reaction time (RT) and accuracy. The dependent measure was the N-back effect, which is the difference in RT or accuracy between N-2 trials and N-1 trials. These differences are thought to represent the cost of updating working memory as task demands increase.

Shifting Tasks.

Number-Letter Task. To measure the participants' ability to efficiently switch between tasks, a number-letter task was used. This task, based off of similar tasks by Soveri et al. (2011) and Miyake et al. (2000) required participants to switch between tasks and response sets based on the location of the stimuli. Participants viewed a table consisting of one column and two rows. On each trial, a number and letter pair (e.g., 4 I) appeared either in the top row or the bottom row. A number or letter could appear in the first or second position of the number-letter pair (). see Appendix H for example stimuli).

When the stimulus pair appeared in the top row, participants had to decide whether the number is odd (press the left key) or even (press the right key). When the stimulus pair appeared in the bottom row, the participants had to decide whether the letter was a vowel (press the left key) or a consonant (press the right key). Number and letter response sets each consisted of 8 unique targets. The number-set consisted of four even number digits (2, 4, 6, and 8) and four odd number digits (3, 5, 7, and 9). The letter-set consisted of four vowels (A, E, I, and U) and four consonants (D, L, R, and S). Letters were chosen based on the most common letters across English and Spanish languages. The vowel "O" was more common than "U," but was not selected due to its resemblance to the number zero (0

Participants completed a single-task block of 80 trials followed by a mixed-task block of 80 trials with a 15 sec break in between. In the single-task block, participants were first presented with either 40 trials all located in the top row followed by 40 trials all presented in the bottom row (Version A) or vice versa (Version B). Thus, in the single-task block participants did not have to decide whether to respond to the number or the letter, but only what kind of number or letter. Which sub-block (i.e., top row or bottom row) first appears was assigned randomly. After the break, the participants completed the mixed-task block with 80 more trials in which stimuli randomly appeared in the top or the bottom row. Participants had to decide whether to respond to the number or the letter and then decide to which set the number or letter belonged.

A trial consisted of a fixation cross that appeared in the center of the screen for 300 ms, and was followed by the appearance of the two squares. The number-letter pair was in either the top square (top row) or bottom square (bottom row). The participants saw the stimuli for 1500 ms and could respond at any point during this interval. Next a blank screen appeared for 150 ms followed by the fixation cross that signaled the beginning of the next trial.

The key data were reaction time and accuracy. The dependent measures were the differences between responses in mixed set of trials (mixing costs) and the simple set of trials (switching costs). These differences represent the processing cost associated with switching between tasks.

Global-Local task. The second switching task was the global-local task modified from one used by Miyake et al (2000). In the global-local task, participants viewed an image, in which the lines of the ‘‘global’’ figure (e.g., a number: digits 1-4) were composed of much smaller, ‘‘local’’ figures (e.g., digits 1-4). Target task (naming the global vs. the local figure) was cued by the color of the figure. If the figure was blue, participants had to identify the global shape and if the figure was green, participants had to identify the local shape. Although Miyake et al.’s participants responded verbally, responses were made via the response box with buttons corresponding to 1, 2, 3, and 4 to keep the task as non-verbal as possible. Both reaction time and accuracy was recorded. The shifting cost was calculated by taking the difference in average reaction times on the switch trials and on non-switch trials (see Appendix I for example stimuli).

Participants engaged in 12 practice trials before starting the 120 experiment trials. The trials were pseudo-randomized such that half the trials required switching between task sets (e.g., trial 1: global, trial 2: local) and the other involved repeating task sets (e.g., trial 1: global, trial 2: global). A trial consisted of a fixation cross appearing in the center of the screen for 300 ms, followed by the appearance of the geometric figure. The participants saw the stimuli for up to 2000 ms and could respond at any point during this interval. Next a blank screen appeared for 150 ms followed by the fixation cross that signals the beginning of the next trial.

Inhibitory control.

Flanker task. Participants completed a flanker test to measure their ability to use inhibitory control to overcome interference of competing information. The Flanker task, originally developed by Eriksen and Eriksen (1974) is a common executive control task used in bilingual research (Costa, Hernández, Costa-Faidella, Sebastián-Gallés, & Núria, 2009; Emmorey et al., 2008; Gollan, Sandoval, & Salmon, 2011; Soveri, et al., 2011) and cognitive research in general. The task required the participants to view a horizontal row of 5 arrows in the center of the computer screen and decide which direction the center arrow was pointing (left vs. right). The flanking arrows on either side of the center arrow were all pointing in the same direction as the target (a congruent trial; e.g., → → → → →) or were all pointing the opposite direction as the target (an incongruent trial; e.g., → → ← → →). If the target arrow was pointing left, participants pressed the left button on the response pad; if the target arrow was pointing right, participants pressed the right button on the response pad. See Appendix J for example stimuli.

After completing a practice set of 10 trials with accuracy feedback, participants completed 100 experimental trials. A trial consisted of a black fixation cross (+) for 1000 ms, followed by a screen with the five arrows that appeared for 500 ms, which was then followed by a blank screen for 500 ms. Participants could respond while the arrows were displayed or during the blank screen. After the blank screen, the next trial started regardless of whether participants responded or not. No feedback was given.

Fifty congruent and 50 incongruent trials were presented in random order. Within the congruent trials, 25 of the trials had arrows that all pointed right and the other 25 trials had arrows that all pointed left. Within the incongruent trials, 25 of the trials consisted of 4 left-pointing flanker arrows and a right-pointing target arrow and 25 consisted of 4 right-pointing flanker arrows and a left-pointing target arrow (see Appendix A for an overview of all possible combinations).

Key data for this test were reaction time (RT) and accuracy. The dependent measures were the Flanker effects, which is the difference in RT or accuracy between congruent and incongruent trials. This difference is thought to represent the processing cost of inhibiting the conflicting flanker arrows (Soveri et al., 2011).

Simon Task. The second measure of inhibitory control was the Simon Task in which red and blue squares were presented on either the left or right side of a white screen. Using the response pad, participants identified the color of the item. In order to cause interference between responses, participants had to press a key on the left if the square was blue and a key on the right if the square was red. In order to respond accurately on incongruent trials, participants had to ignore the irrelevant information of spatial location and respond purely based on color (see Appendix K for stimuli).

Participants completed 100 experimental trials. A trial consisted of a black fixation cross presented in the center of the screen (+) for 500 ms, followed by a screen with red or blue square for 1000 ms, which was then followed by a blank screen for 350 ms. Participants could respond either while the stimulus item is displayed or during the blank screen. After the blank screen, the next trial started regardless of whether participants had responded or not. No feedback was given. Fifty congruent and 50 incongruent trials were presented in random order. Within the congruent trials, there were 25 trials in which the blue square appeared on the left side of the screen and 25 trials in which the red square appeared on the right side of the screen. Within the incongruent trials, 25 trials consisted of blue squares on the right side of the screen and 25 trials consisted of red squares on the left side of the screen. The key dependent measure was the Simon effect, which is the difference in reaction time between congruent and incongruent trials for trials in which participants responded correctly.

General Procedure

Participants first consented to participate and then the experimenter gave a brief overview of the types of tasks the participants would complete during the three hour session. Participants all completed the same tasks, but task order was semi-counter-balanced across two orders of tasks (Version A and Version B).

The first task was the picture-naming task, which included the presentation of the training slides, the pre-test, and the experimental task itself. All verbal responses were recorded digitally for later accuracy scoring and the naming onset latency was recorded via the response box. Half the participants completed the English portion first (Version A) and the Spanish portion second and the other half completed the Spanish portion first (Version B), followed by the English portion. Everyone then completed the color-shape task. Version A participants started with the simple-color block and Version B started with the simple-shape block.

The participants then completed the six executive control tasks. The participants in Version A completed the tasks in the following order: auto-symmetry span, Simon, global-local, visual-spatial n-back, number-letter, and Flanker. The other half completed the tasks in the alternative order: visual-spatial n-back, number-letter, Flanker, auto-symmetry span, Simon, and global-local. In both versions, participants were offered a five minute break after the third executive control task.

The next set of tasks was the vocabulary measures. If participants had completed the English portion of the picture-naming task first (Version A), they first completed the English verbal fluency test and then the PPVT. To avoid interference from testing the two languages back-to-back, participants completed the Raven's APM before completing the Spanish verbal fluency task and the TVIP. If the participants had started with a Spanish block (Version B) in the picture-naming task, they first completed the Spanish verbal fluency test and the TVIP. Before completing the equivalent measures in English, they completed the Raven's. Verbal fluency was always completed before the PPVT/TVIP in the respective language so that there would not be priming from the receptive measure on the productive measure.

Lastly, participants filled out the LEAP-Q. They were then debriefed and assigned credit or paid for their participation.

III. RESULTS

Background Measures

LEAP-Q. A large amount of descriptive data were collected using the LEAP-Q. Here I present a summary of participants' responses related to age of acquisition and age at which fluency was reached for various verbal skills, exposure to a language, and self-rated proficiency. Responses to the LEAP-Q are summarized in Table 1.

There was no significant difference between second language learners (SLL) and bilinguals (BL) in age and years of education.

Participants in the SLL group reported acquiring English three years earlier than BL participants and speaking fluently at a marginally earlier age than BL participants (p = .06). There was no difference in age for when SLL and BL participants reported beginning to read in English or the age at which they began to read fluently in English. As expected, SLL participants reported acquiring Spanish, learning to read in Spanish, and becoming fluent in speaking and reading Spanish at a later age than did BL participants[1].

The SLL group reported they were exposed to Spanish less frequently than BL participants and they preferred to speak in Spanish less than BL participants. However, both groups reported a similar preference for amount of time they choose to read in Spanish. This likely reflects the fact that most students attended English speaking schools.

For self-rated proficiency in English, SLL and BL participants reported similarly high proficiency levels for speaking, understanding, and reading, which led to similar English composite scores. The SLLs reported lower proficiency levels for speaking, understanding, and reading in Spanish than did BLs, which led to different Spanish composite scores.

[1] Only 8 of the 20 SLL participants reported becoming fluent speakers and readers in Spanish.

Receptive vocabulary. Receptive vocabulary was measured using Version A of the Peabody Picture Vocabulary Test (PPVT) and Version A of its Spanish equivalent, the Test de Vocabulario en Imagenes Peabody (TVIP). Raw scores were calculated for each test following the guidelines provided with the tests and used for analyses (see Table 2 for means). PPVT scores can range between 0 and 228. TVIP raw scores can range between 0 and 125. Although it is possible to convert raw scores to standardized scores for purposes of direct comparison between tests, this was not done as the majority of participants reached ceiling level on the TVIP. Therefore, scores for the TVIP would not be a valid comparison.

Separate one-way ANOVAs were conducted for both English and Spanish versions to test for group differences in raw scores. For the PPVT, SLLs (M = 209.0, SE = 1.7) performed significantly better (18 points) than did BL participants (M = 191.3, SE = 3.3), $F(1,40) = 23.50$, MSE = 140.21, $p < .001$, but these scores place both groups at an “adult” level in English receptive vocabulary. For the TVIP, BL participants (M = 107.1, SE = 2.5) performed slightly better (2 points) than SLLs (M = 105.0, SE = 3.8), but this difference did not approach significance, $F(1,40) < 1.0$, MSE = 217.25, *ns*. These scores place both groups at an “adult” level in Spanish receptive vocabulary. Five of the SLLs and 3 of the BLs groups scored within 5% of the maximum. This reflects a ceiling effect for the Spanish test.

Productive vocabulary. Productive vocabulary was measured using the verbal fluency task and is reported as the average number of items produced for each language (see Table 2 for means). A 2(Language: English, Spanish) X 2(Participant Group: SLL, BL) repeated-measures ANOVA was conducted with average number of items as the dependent measure. Collapsed across SLLs and BLs, the average number of items produced was 11.1 (SE = .4). There was a main effect of language, such that participants produced 5 more items on average in English than in Spanish, $F(1,40) = 50.38$, $MSE = 6.05$, $p < .001$. There was no main effect of participant group, $F(1,40) < 1.0$, $MSE = 10.95$, *ns*, indicating SLL and BL participants produced the same number of items. The main effect is qualified by a significant Language X Participant Group interaction, $F(1,40) = 14.22$, $MSE = 6.05$, $p < .001$. Follow-up tests revealed that SLL participants generated approximately 95% more items in English than they did in Spanish, but BL participants produced only 20% more items in English than in Spanish, $F(1,40) = 44.66$, $MSE = 8.00$, $p < .001$. This interaction reflects the fact that the SLL group was much more proficient in English than Spanish.

General intelligence – Raven's Progressive Matrices. Accuracy was defined as the proportion of trials answered correctly. All participants performed above chance accuracy ($M = .53$, $SE = .03$), which was 12.5%. Participants in the SLL group were 8% more accurate than participants in the BL group, but this difference was not significant, $F(1,40) = 2.70$, $MSE = .03$, $p = .11$.

Background measures summary. The analyses of background measures indicate that SLL and BL participants were matched on age, years of education, Raven's, and most self-report measures for English. The SLL group reported an earlier age of English acquisition, but the average age of acquisition for both groups was less than 4 years, indicating that both groups were early learners of English. The SLL group had slightly better vocabulary scores in English than did BLs, but both groups were proficient. The BL participants were exposed to Spanish earlier, more often, and speak it more frequently. As expected, BL participants rated themselves higher in Spanish proficiency than did SLLs. These self-ratings matched the findings of the productive vocabulary measure, but not the receptive vocabulary measure, which appears to be limited by a ceiling effect. In essence, these measures indicate that the two groups were similarly proficient with English, but the BL group was much more proficient with Spanish.

Experimental Tasks (Color-Shape Identification and Picture-Naming Tasks)

Data cleaning. For each participant, the final data set was created by excluding (a) the first trial in a block (because first trials could not be switch or repeat trials), (b) trials with reaction times under 200 ms (which reflect artificially fast responses), and (c) trials that had a reaction time of ±3 standard deviations from a participant's mean reaction time for the task. The same procedure was used for each task.

Analyses. For each variable, I ran separate ANOVAs for each of the three dependent measures: accuracy, reaction time for all trials, and reaction time for correct trials only. Accuracy was defined as the average proportion of correct trials. Reaction time for all trials (RT-All) is the average reaction time for all trials on which participants responded. Reaction time for correct trials only (RT-CR) is the average reaction time for the trials on which the participants responded correctly. There was very little difference between the results for reaction time based on all trials and correct trials only and the patterns of means was identical. In a few cases, effects that were marginally significant for RT-All became statistically significant for RT-CR. Therefore, I only report results for accuracy and reaction time for correct trials only (RT-CR).

For each measure I present analyses for switching and mixing costs as my primary predictions were based on these measures. As a reminder, switching cost is the difference between switch trials and repetition trials – both of which are in mixed blocks. Mixing cost is the difference between repetition trials in the mixed block and the average response time for trials in simple blocks (all trials in simple blocks are repetition trials).

Color-Shape task. Accuracy was defined as the average proportion of trials for which participants correctly indicated the color or shape as designated by the cue. Trials in which participants did not respond or in which reaction time was not recorded (due to response box error) were not included in the analyses. Reaction time was defined as the time from when the stimulus was displayed until the time the participants pressed a button on the response box. Reaction times for incorrect trials were excluded from the reaction time analyses.

Original hypotheses. Analyses of the picture-naming task were conducted to test Hypotheses 1A and 1B. Hypothesis 1A predicted that BLs would have reduced switching costs compared to SLLs due to enhanced inhibitory control. Hypothesis 2B predicted that if there was a group difference in mixing costs, BLs would have smaller mixing costs than SLLs. This would be due to a superior ability to update goals in working memory.

Switching costs. Switching costs in accuracy and reaction time were analyzed using 2(Trial Type: switch, repetition) X 2(Participant Group: SLL, BL) repeated-measures analyses of variance (ANOVA). For accuracy, there was no main effect of trial type [$F(1,40) = 1.24$, $MSE = .002$, $p = .273$], no main effect of participant group [$F(1,40) = 1.73$, $MSE=80219.26$, *ns*], and no interaction ($F < 1$).For reaction time, there was a significant main effect of trial type $F(1,40) = 83.21$, $MSE = 5690.86$, $p < .001$. Participants were 150 ms faster on repetition trials than on switch trials – that is, participants showed a switching cost. There was no main effect of participant group [$F(1,40) < 1.0$, $MSE=.005$, *ns*] and no two-way interaction [$F < 1$].

Results do not support the original hypothesis (1A) that BLs would have reduced switching costs compared to SLLs (see Table 3 for means).

Mixing Costs. Mixing costs in accuracy and reaction time were analyzed using 2(Trial Type: repetition, simple) X 2(Participant Group: SLL, BL) repeated-measures analyses of variance (ANOVA). There were a significant main effect of trial type for accuracy $F(1,40) = 5.93$, $MSE = .002$, $p < .05$– that is, participants showed a mixing cost. Participants were 3% more accurate on simple trials compared to repetition trials. There was no main effect of participant group or two-way interaction, $F < 1$.

There were a significant main effect of trial type for reaction time, $F(1,40) = 370.39$, $MSE = 13661.81$, $p < .001$, – again, participants showed a mixing cost. Participants were 491 ms faster on simple trials compared to repetition trials. There was a main effect of participant group on reaction time ($F[1,40] = 4.54$, $MSE = 36177.04$, $p < .05$. Overall, SLLs were 88 ms faster than bilinguals. The two-way interaction was not significant, $F < 1$.

Results do not support the original hypothesis (1B) in that BLs did not have reduced mixing costs compared to SLLS (see Table 3 for means).

Summary of group differences. There were no differences in accuracy as a function of participant group. SLL participants were faster on the repetition and simple trials combined compared to BL participants (main effect of group). Although both SLL and BL groups produced switching mixing costs, there was no difference in the size of these costs across language groups. Hypotheses 1A and 1B were not supported by either the accuracy or reaction time data.

Picture-Naming task. Accuracy was defined as the average proportion of trials on which participants correctly named the picture in the cued language. Trials in which the participants did not respond or in which reaction time was not recorded by the response box were not included in the analyses. Reaction time was defined as the time between the presentation of the stimulus and the participants' naming response. Reaction times for incorrect trials were excluded from the reaction time analyses.

Original hypotheses. Analyses of the picture-naming task were conducted to test Hypotheses 2A and 2B. Hypothesis 2A predicted that both SLL and BL participants would exhibit switching costs. For SLLs, costs for switching into English would be larger than switching into Spanish. This would indicate the need for SLL participants to inhibit their L1 (English) more than their L2 (Spanish), as L1 is their dominant response. For BLs, switch costs were predicted to be similar for each language as neither language is highly dominant. Hypothesis 2B predicted that BLs would have smaller mixing costs than SLLs in both languages. This would reflect their enhanced ability in updating goals in working memory.

Switching costs. Switching costs in accuracy and RT-CR were analyzed using 2 (Language: English, Spanish) X 2(Trial Type: switch, repetition) X 2(Participant Group: SLL, BL) repeated-measures analyses of variance (ANOVA). For accuracy, there were no significant main effects or interactions, $Fs \leq 1$.

For reaction time, there was a significant of main effect of language, $F(1,40) = 8.15$, $MSE = 4894.44$, $p < .01$. That is, for trials in the mixed blocks, participants were 31 ms faster when they named items in Spanish than in English. There was a main effect of trial type, $F(1,40) = 57.35$, $MSE = 1940.64$, $p < .001$. Overall, participants were 51 ms faster on repetition trials than on switch trials – that is, participants showed a switching cost in reaction time. There was not a significant of main effect of participant group, $F(1,40) = 1.65$, $MSE = 81472.88$, *ns*. There was a marginally significant Language X Participant Group interaction for reaction time, $F(1,40) = 3.65$, $MSE = 4894.44$, $p = .063$. Contrast tests revealed that for trials in the mixed blocks, SLLs were 51 ms slower in English than in Spanish ($F[1,40] = 5.67$, $MSE = 4894.44$, $p < .05$), whereas bilinguals had similar reaction times in both languages, $F < 1$. There was a significant Language X Trial Type interaction for reaction time, $F(1,40) = 11.07$, $MSE = 1703.37$, $p < .01$. Contrast tests revealed that reaction time was faster for repetition trials than for switch trials in both languages, but that this difference was 43 ms larger in Spanish ($F[1,40] = 57.14$, $MSE = 1940.64$, $p < .001$) than in English, ($F[1,40] = 9.93$, $MSE = 1940.64$, $p < .01$). None of the other interactions were significant.

Results partially support Hypothesis 2A that both SLLs and BLs would have switching costs, but that SLLs would have larger switching costs for English than for Spanish. My prediction was accurate in that both participant groups had reaction time switching costs. However, for each group, switching costs in English were the same as in Spanish (see Tables 4A and 4B for means).

Mixing costs. Mixing costs in accuracy and reaction time were analyzed using 2 (Language: English, Spanish) X 2(Trial Type: switch, repetition) X 2(Participant Group: SLL, BL) repeated-measures analyses of variance (ANOVA). For accuracy, there were no significant main effects, Fs < 1. There was a marginally significant Language X Trial Type interaction, $F(1,40) = 2.99$, $MSE = .001$, $p = .09$. Participants did not show mixing costs in English, but did for Spanish, though neither mixing cost was significantly difference from zero, (Fs < 1). None of the other interactions were significant.

For reaction time, the main effects of language and participant group were not significant ($Fs < 1$), but there was a main effect of trial type, $F(1,40) = 119.26$, $MSE = 2915.60$, $p < .001$. Overall, participants were 91 ms faster on simple trials than on repetition trials – that is, participants showed a mixing cost in reaction time. Interestingly, the 3-way interaction between language, trial type, and participant group was significant, $F(1,40) = 4.51$, $MSE = 1456.94$, $p < .05$.

The interaction was followed up at the levels of task language. For English trials, the simple effect of trial type was significant, such that participants were 132 ms faster on simple trials than on repetition trials, $F(1,40) = 125.73$, $MSE = 2915.60$, $p < .001$. That is, participants experienced mixing costs when naming items in English. Neither the simple effect of participant group nor the interaction between trial type and participant group were significant, $F < 1$, *ns*. For Spanish trials only, the simple effect of trial type was again significant, such that participants were 50 ms faster on simple trials than on repetition trials, $F(1,40) = 17.90$, $MSE = 2915.60$, $p < .001$. That is, participants experienced mixing costs when naming items in Spanish. There was no simple effect of participant group ($F < 1$, *ns*), but the interaction between trial type and participant group was significant, $F(1,40) = 4.90$, $MSE = 2915.60$, $p < .05$. Pairwise comparisons revealed that for SLLs, there was no significant difference between simple and repetition trials – that is, they did not demonstrate any mixing costs in Spanish, $F < 1$, *ns*. In contrast, BLs were significantly faster on simple trials than on repetition trials by 76 ms – that is, they had mixing costs in Spanish, $F(1,40) = 20.78$, $MSE = 2915.60$, $p < .001$.

In summary, SLLs and BLs had similar mixing costs in English, and SLLs had smaller mixing costs in Spanish compared to BLs. Results did not support my hypothesis (2B) that BLs would have reduced mixing costs compared to SLLs (see Tables 4A and 4B for means).

Summary of group differences. Both SLLs and BLs had switching and mixing costs as measured by reaction time. SLLs were slower overall on the English trials in the mixed block (switch and repetition trials) than on the Spanish trials in the mixed block, whereas BLs' reaction times did not vary as a function of language. Hypothesis 2A was not supported by accuracy data and only partially supported by reaction time data (i.e., for SLL, reaction time switching costs were *not* larger when switching in English compared to switching into Spanish). Neither participant group differed in the size of switching tasks a function of task language. This was predicted for BLs but not SLLs.

Hypothesis 2B predicted that BLs would have smaller mixing costs in both languages compared to SLLs and that mixing costs would not differ as a function of language. This hypothesis was partially supported by the reaction time data, in that switching costs did not differ as a function of language for bilinguals. However, bilinguals did not have smaller mixing costs compared to second language learners as measured by reaction time and accuracy. Mixing costs were either the same (in English) or larger (in Spanish).

Relationship between the Color-Shape and Picture-Naming Tasks

Although this was not a research question that I initially presented in the proposal, it is theoretically interesting to explore the relationship between the two switching tasks. Of interest is whether there is a difference in the relationship between the two switching tasks when a person is highly proficient in two languages versus when a person is still learning the second language. In other words, do BLs and SLLs approach the tasks in the same ways?

To test the moderating hypothesis, I used reaction time from the color-shape task as my focal predictor and participant group as my moderating predictor. If both participant groups approach the two tasks in the same way, performance on the color-shape task will explain a similar amount of variance in performance on the picture-naming task for SLLs and BLs. If the two groups are processing the stimuli and responding to the two tasks in different manners (i.e., participant group moderates performance on the picture-naming task), then the color-shape task will explain a different amount of variance in performance on the picture-naming task for SLLs and BLs.

To test the model, I first centered the color-shape RT-CR means for each trial type and then computed the Color-Shape X Participant Group interaction terms. I then ran a linear regression by entering color-shape RT, participant group, and the interaction term to predict picture-naming RT-CR. Six separate linear regression analyses were conducted for the reaction times for each trial type crossed with each language in the picture-naming task. For example, reaction time from color-shape switch trials in English was used to predict RT for picture naming switch trials in English, reaction time from color-shape switch trials in Spanish was used to predict RT for picture naming switch trials in Spanish, and so on. If participant group moderated the relationship between the two RT-CR measures, the interaction term was significant. When the interaction term was significant, I followed it up by performing tests of simple slopes. This looked at the relationship between the two RT-CR measures for each participant group separately and indicated whether the each of these simple slopes was significantly different from zero.

Results indicated that performance on the color-shape task was a significant predictor of performance on the picture-naming task for all of the six models (*p*s < .05). The general pattern suggests that the faster participants were on a given type of color-shape trial (e.g., switch trials), the faster the participants were on the corresponding picture naming trials (i.e., switch trials). In contrast, results indicated that participant group was a significant predictor of performance on the picture-naming task for only one of the six conditions (English simple trials: $B = -69.78$, $t[39] = -2.15$, $p < .001$) and a marginally significant predictor in another (Spanish simple trials: $B = -62.18$, $t[39] = -1.84$, $p = .074$). This suggests that language background alone was generally not a good predictor of performance on the picture-naming task. However, this finding was qualified by the interaction between color-shape task RT and participant group. The interaction was significant for English simple trials (B = -.75, t[39] = -2.46, $p < .05$) and marginally significant for three other trial types (Spanish simple trials: B = -.547, t[39] = -1.72, p = .09; English repetition trials: B = -.35, t[39] = -1.79, p = .08; and Spanish switch trials: B = -.37, t[39] = -1.67, p = .10). This suggests that the explanatory power of the color-shape task performance depended on language background (see Figures1A-1F for simple slopes graphs).

Simple slopes for the association between color-shape reaction time and picture naming reaction time for SLLs and BLs revealed that the relationship was generally stronger for SLLs than for BLs. For both English simple trials, color-shape RT was a significant predictor of picture-naming RT for both SLLs (b = 1.24, t[39] = 5.30, $p < .001$) and BLs (b = .50, t[39] = 2.56, $p < .05$), but was stronger for SLLs. The same was true for Spanish simple trials: (SLLs: b = 1.02, t[39] = 5.17, $p < .001$; BLs: b = .48, t[39] = 1.86, p = .07). For English repetition trials, color-shape RT was a significant predictor of picture-naming RT for SLLs (b = .51, t[39] = 3.80, $p < .001$), but not BLs (b = .17, t[39] = 1.199, p = .24). The same was true for Spanish switch trials: (SLLs: b = .56, t[39] = 3.86, $p < .001$; BLs: b = .19, t[39] = 1.14, p = .26, respectively).

In summary, participant group had a moderating effect in the relationship between color-shape task performance and picture-naming task performance for four of the trial types (English simple, Spanish Simple, English repetition, and Spanish switch). In all cases, there was a stronger relationship between the two reaction time measures for SLLs than for BLs. In two of the four cases (English repetition and Spanish switch), the two tasks were not significantly related. These results are interesting because they indicate differences in cognitive processes used by the SLLs and the BLs in the switch tasks. More specifically, how SLLs perform on the nonverbal task is more closely associated with their verbal task performance than it is for BLs. Second language learners seem to treat the two tasks more similarly than do bilinguals – that is, the addition of a verbal component to the task does not change the way in which the participants switch between the two task sets (color and shape, English and Spanish). In contrast, bilinguals appear to switch between the two tasks sets in different manners. Could this be a reflection of an expertise effect in switching between languages? Bilinguals have much more practice at switching between languages than they do between color and shape identification. It seems as though their switching skills are highly specialized to verbal tasks and not nonverbal tasks, as evidenced by the smaller associations, as compared to bilinguals) between the RT-CR measures for the color-shape and picture-naming tasks. If this is the case, I would not expect to see much carryover of these skills to any of the non-verbal executive control tasks. This would be the opposite of the common explanation of why bilinguals have advanced executive control. If bilinguals are relying heavily on verbal processing for switching between languages, this would not transfer to the nonverbal domain and subsequently have an impact on general executive control.

Executive Control Tasks

Working Memory Tasks (Auto-Symmetry Span and Visual-Spatial N-Back).

Auto-Symmetry Span. Span score was defined as the sum of all trials in which all items were recalled in the correct serial order. Total score was defined as the sum of items recalled in the correct serial position, regardless of whether the entire trial was correctly recalled. Two participants were excluded (one from each participant group) because they made more than six errors in the symmetry judgment portion of the task. This is a standard procedure when using the task.

A one-way ANOVA was conducted to compare the span scores for the two participant groups. There was no significant difference between SLL (M = 18.55, SE = 1.96) and BL (M = 18.15, SE = 1.88), $F(1,38) < 1.0$, MSE = 73.72, *ns*.

Additionally, a one-way ANOVA was conducted to compare the total span score for the two groups. There was no significant difference between SLL (M = 28.57, SE = 1.74) and BL (M = 27.05, SE = 1.82), $F(1,38) < 1.0$, MSE = 66.70, *ns*.

Summary of group differences. The SLL and BL participants performed equivalently on the Symmetry Span task.

Visual Spatial N-Back. Accuracy was defined as the average proportion of trials for which the participants responded correctly by indicating whether the square appeared in the same location as the previous trial (one back) or two trials previous (two back). Reaction time was defined as the time from when the stimulus was displayed until the time the participants pressed a button on the response box. Reaction times for incorrect trials were excluded from these analyses. Data were not collected for one participant in the SLL group due to a computer malfunction.

A 2(Trial Type: N-1, N-2) X 2(Participant Group: SLL, BL) repeated-measures ANOVA was conducted with accuracy as the dependent measure (see Figure 2 for means). There was a significant main effect of trial type, such that participants were 3% more accurate on N-1 trials than on N-2 trials $F(1,39) = 10.79$, MSE = .002, $p < .01$. There was no significant main effect of participant group, $F(1,39) = 1.24$, MSE = .008, $p > .05$; nor was there an interaction between trial type and group, $F(1,39) < 1$, MSE = .002, *ns.*

A 2(Trial Type: N-1, N-2) X 2(Group: SLL, BL) repeated-measures ANOVA was conducted with reaction time for correct trials only as the dependent measure (see Figure 2B for means). There was a significant main effect of trial type, such that participants responded 84 ms faster on N-1 trials than on N-2 trials, $F(1,39) = 23.55$, $MSE = 6104.09$, $p < .001$. There was a marginally significant main effect of participant group, such that SLLs responded 81 ms faster than BLs, $F(1,39) = 1.24$, $MSE = .008$, $p = .05$. The interaction between trial type and participant group was not significant, $F(1,39) = 1.46$, $MSE = 6104.09$, $p > .05$.

Summary of group differences. The SLL and BL participants had similar accuracy rates. The SLL participants were faster on correct trials than were BL participants.

Results from the two working memory tasks did not match predictions. I had predicted that BLs would perform better on working memory tasks than would SLLs due to the expectation of enhanced executive control. Instead, there were no group differences in working memory other than response time.

Shifting tasks (number-letter and global-local tasks).

Number-Letter Task. Accuracy was defined as the average proportion of trials in which the participants responded correctly by indicating whether the number was odd or even or whether the letter was a vowel or a consonant. Reaction time was defined as the time from when the stimulus was displayed until the time the participants pressed a button on the response box. Reaction times for incorrect trials were excluded. Data were not collected from three participants in the SLL group due to computer malfunctions.

Accuracy. A 3(Trial Type: simple, repetition, switch) X 2(Participant Group: SLL, BL) repeated-measures analysis of variance (ANOVA) was run (see Figure 4A for means). There was a significant main effect of trial type, such that participants were more accurate on simple trials than on repetition and switch trials (6.9% and 9.4%, respectively) and 2.5% more accurate on repetition trials than on switch trials, $F(2,74) = 37.27$, $MSE = .002$, $p < .001$. There was no main effect of participant group, $F(1,37) < 1$, MSE =.005, *ns*, and no interaction between trial type and participant group, $F(2,74) < 1$, MSE = .002, *ns*.

Reaction Time. A 3(Trial Type: simple, repetition, switch) X 2(Participant Group: SLL, BL) repeated-measures analysis of variance (ANOVA) was run (see Figure 4B for means). There was a significant main effect for trial type, $F(2,74) = 320.50$, $MSE = 3215.66$, $p < .001$. Means comparison tests indicated that average reaction time was faster on simple trials than repetition or switch trials (114 ms and 322 ms, respectively), $p < .001$. Repetition trials were also 207 ms faster than switch trials, $p < .001$. There was a no main effect of participant group, $F(1,37) < 1$, $MSE = 18831.41$, *ns*. The interaction between trial type and participant group was significant, $F(2,74) = 9.16$, $MSE = 3215.66$, $p < .001$. Contrast tests revealed that both groups performed equivalently on repetition and switch trials, Fs < 1, *ns*, but that SLL participants were 64 ms faster than BL participants on simple trials, $F(1,74) = 4.53$, $MSE = 18831.41$, $p < .05$. This means that the SLL participants had significantly larger switching costs (M = 151 ms, SE = 12) than did BL participants (M = 84 ms, SE = 20).

Summary of group differences. The SLL and BL groups had similar levels of accuracy in identifying the number or letter. For simple trials, reaction times for SLL participants were faster than BL participants. However, bilinguals had smaller mixing costs than SLL participants.

Global-Local. Accuracy was defined as the average proportion of trials in which the participants responded correctly by indicating the number shown as specified by the cue. Reaction time was defined as the time from when the stimulus was displayed until the time the participants pressed a button on the response box.

Accuracy. A 2(Focus: global, local) X 2(Congruency: match, mismatch) X 2 (Participant Group: SLL, BL) repeated-measures ANOVA was conducted (see Figure 3A for means). There was a significant main effect of focus, such that participants were more 37% more accurate on global trials than on local trials, $F(1,40) = 268.73$, $MSE = .022$, $p < .001$. There was also a significant main effect of congruency, such that participants were 51% more accurate on trials when the numbers matched than on trials when they did not, $F(1,40) = 4916.16$, $MSE = .002$, $p < .001$. There was no main effect of participant group: SLL and BL had identical accuracy rates ($M = .72$, $SE = .007$), $F(1,40) < 1$, $MSE = .004$, *ns*.

The Focus X Congruency interaction was significant, $F\ (1,40) = 326.17$, $MSE = .019$, $p < .001$. As seen in Figure 3a, contrast tests revealed that for matched trials, there was no difference in accuracy as a function of focus (Global: $M = .97$, $SE = .008$; Local: $M = .98$, $SE = .007$), $F\ (1,40) < 1$, $MSE = .022$, *ns*. In contrast, for mismatched trials there was a significant difference in accuracy, such that accuracy remained relatively high for global trials ($M = .84$, $SE = .024$), but was extremely poor for local trials ($M = .09$, $SE = .021$), $F\ (1,40) = 298.16$, $MSE = .022$, $p < .001$.[2]

Reaction Time. The reaction time analysis was run based on data from all trials in which participants responded, rather than correct trials only because there were not enough data for the local-mismatched trials to provide reliable results due to the very high error rate in that condition (as described in the accuracy analysis above).

A 2(Focus: global, local) X 2(Congruency: match, mismatch) X 2 (Participant Group: SLL, BL) repeated-measures ANOVA was conducted (see Figure 3B for means). There was a significant main effect of focus, such that participants were 31 ms faster on global trials than on local trials, $F(1,40) = 9.18$, $MSE = 4311.69$, $p < .01$. There was a significant main effect of congruency, such that participants were 123 ms faster on matched than on mismatched trials,

---

[2] Six participants had a 0% accuracy rate for local-mismatched trials. I excluded them and reran the accuracy analysis. The patterns of data remained the same and the accuracy rate for local-mismatched trials remained low at 11%.

$F(1,40) = 126.84$, $MSE = 4983.48$, $p < .001$. The Focus X Congruency interaction was significant, $F(1,40) = 326.17$, $MSE = 4311.69$, $p < .001$. Follow-up analyses showed that for matched trials, participants were 48 ms faster on the global trials than on the local trials, $F(1,40) = 11.65$, $MSE = 4311.69$, $p < .01$. In contrast, for incongruent trials, reaction time was equivalent for local and global trials, $F(1,40) = 1.39$, $MSE = 4311.69$, *ns*.

Summary of group differences. There were no differences between the SLL and BL participants in accuracy or reaction time.

I had not predicted differences in shifting, as none had been previously reported in the literature. There were no group differences in shifting, which supports previous research.

Inhibition (Flanker Task and Simon Task).

Flanker Task. Accuracy was defined as the average proportion of trials in which participants responded correctly by indicating the whether the arrow pointed left or right. Reaction time was defined as the time from when the stimulus was displayed until the time the participants pressed a button on the response box. Reaction times for incorrect trials were excluded from this analysis.

Accuracy. A 2(Trial Type: Congruent, Incongruent) X 2(Participant Group: SLL, BL) repeated-measures ANOVA was conducted (see Figure 6A for means). There was a significant main effect of trial type, such that participants were 4.5% more accurate on congruent trials than on incongruent trials, $F(1,40) = 63.58$, $MSE = .001$, $p < .001$. There was a marginally significant main effect of participant group, such that bilinguals ($M = .98$, $SE = .005$) were 1.2% more accurate than SLLs ($M = .97$, $SE = .005$), $F(1,40) = 3.5$, $MSE = .001$, $p = .069$. This was qualified by the Trial Type X Participant Group interaction, $F(1,40) = 4.14$, $MSE = .001$, $p < .05$. For congruent trials, there was no difference in accuracy as a function of participant group, *F*(1,40) < 1, *MSE* = .001, *ns*. In contrast, for incongruent trials, bilinguals (*M* = .96, *SE* = .006) 2% were more accurate than SLLs (*M* = .94, *SE* = .005), *F*(1,40) = 6.0, *MSE* = .001, *p* < .05. This meant that the flanker effect was significantly larger for SLLs (*M* = .06, *SE* = .04) than for BLs (*M* = .03, *SE* = .04).

Reaction Time. A 2(Trial Type: Congruent, Incongruent) X 2(Participant Group: SLL, BL) repeated-measures ANOVA was conducted (see Figure 6B for means). There was a significant main effect of trial type, such that participants were 63 ms faster on congruent trials than on incongruent trials, *F*(1,40) = 513.89, *MSE* = 161.21, *p* < .001. There was a significant main effect of participant group, such that SLLs (*M* = 416 ms, *SE* = 9) were 27 ms faster than BLs, *F*(1,40) = 4.81, *MSE* = 3146.81, *p* < .05. There was a marginal Trial Type X Participant Group interaction, *F*(1,40) = 3.64, *MSE* = 161.21, *p* = .06. Contrast tests showed a marginal difference in reaction time on incongruent trials as a function of participant group, such that SLLs were 22 ms faster than BLs, *F*(1,40) = 3.10, *MSE* = 3146.81, *p* = .09. The pattern was similar for congruent trials, with SLLs responding 32 ms faster than BLs, *F*(1,40) = 6.27, *MSE* = 3146.81, *p* < .05. This resulted in a marginally larger Flanker effect for SLLs (*M* = 68 ms, *SE* = 4) than for BLs (*M* = 57, *SE* = 4).

Simon Task. Accuracy was defined as the average proportion of trials in which the participants responded correctly by indicating whether the square was red or blue. Reaction time was defined as the time from when the stimulus was displayed until the time the participants pressed a button on the response box. Reaction times for incorrect trials were excluded from analyses.

Accuracy. A 2(Trial Type: Congruent, Incongruent) X 2(Participant Group: SLL, BL) repeated-measures ANOVA was conducted (see Figure 5A for means). There was a significant main effect of trial type, such that participants were 3% more accurate on congruent trials than on incongruent trials, $F(1,40) = 8.93$, $MSE = .002$, $p < .01$. There was not a significant main effect of participant group, $F(1,40) < 1$, $MSE = .002$, *ns.* Additionally, there was no interaction between trial type and participant group, $F(1,39) < 1$, $MSE = .002$, *ns.*

Reaction Time. A 2(Trial Type: Congruent, Incongruent) X 2(Participant Group: SLL, BL) repeated-measures ANOVA was conducted with RT-CR as the dependent measure (see Figure 5B for means). There was a significant main effect of trial type, such that participants were 29 ms faster on congruent trials than on incongruent trials, $F(1,40) = 41.32$, $MSE = 418.39$, $p < .001$. There was a significant main effect of participant group, such that SLL participants were 43 ms faster than BL participants, $F(1,40) = 6.93$, $MSE = 5475.14$, $p < .05$. There was no interaction between trial type and participant group, $F(1,40) = 1.22$, $MSE = 418.39$, $p > .05$.

Summary of group differences. The SLL and BL participants had equivalent accuracy rates on the Simon task. For reaction time, SLL participants responded faster than BL participants.

Summary of group differences. For accuracy, SLL participants were (marginally) less accurate overall and demonstrated a larger Flanker effect. For reaction time, SLL participants were faster overall and again demonstrated a (marginally) larger Flanker effect.

For the inhibitory control tasks, results of the Flanker task partially support the original hypotheses that BLs would demonstrate enhanced inhibitory control. Although there was no evidence of this in the reaction time data, there was evidence for this conclusion in the accuracy data. Specifically, the Flanker effect was larger (worse performance) on incongruent trials than on congruent trials. Both groups showed this pattern, but importantly, the decrease in performance for incongruent trials was less for BLs than SLLs.

Relationship between the Executive Control Tasks

As a manipulation check, the relationship in performance on the two tasks for each executive control ability was explored using bivariate correlations. The tasks used to measure each executive control ability were selected based on their use in previous research on executive control. If pairs of tasks measure the same cognitive processes, then correlations between the two tasks should be moderate-to-strong. All correlations are reported in Table 5. Below I summarize only the notable relationships.

Working memory. Bivariate correlations were run between symmetry span scores and four measures from the visual-spatial n-back task: overall accuracy, the difference in accuracy between N-1 and N-2 trials, average reaction time for correct responses only (RT-CR) and the difference in RT-CR between N-1 and N-2 trials. All correlations were weak (r = -.14 - .18) and not significant, ps > .05.

Shifting. Bivariate correlations were run for four dependent measures from the Number-Letter task and the global local task: overall accuracy and RT; and the number-letter switch cost and the global-local difference as measured by accuracy and RT. Most notably, there were significant, positive correlations between overall accuracy on both tasks (r = .40) and between global-local RT differences and overall RT on the number letter task (r = .60). This indicates that the two tasks overlap in what they measure.

Inhibitory control. Bivariate correlations were run for four measures from both the Flanker task and the Simon task: overall accuracy; the difference in accuracy between incongruent and congruent trials; average RT-CR; and the difference in reaction time between incongruent and congruent trials (i.e., the Flanker effect and the Simon effect). There were significant, positive correlations between overall accuracy on both tasks (r = .40) and overall RT-

CR on both tasks (r = .58). The correlation between the Simon effect and Flanker effect for accuracy was marginally significant, (r = .29), but the same comparison for RT-CR was not, (r = .16).

Summary. There were no significant relationships between the two measures of working memory, suggesting these two tasks do not tap the same processes. Thus, one or both of these tasks is not an ideal measure of working memory for the current populations. There were a few positive relationships between the two shifting tasks for accuracy and RT-CR, suggesting that these two tasks account for some common variance in shifting performance. The Simon and Flanker tasks demonstrated strong positive relationships on the expected measures for accuracy and RT-CR, suggesting that these two tasks account for similar variance in inhibitory control.

Relationship between Executive Control and Switching Task Performance

A set of correlations using data from the individual difference measures was conducted to see which of the executive control measures correlated with a) switching costs on the verbal task, b) mixing costs on the verbal task, and c) proficiency levels in English and Spanish. Proficiency levels were analyzed in three ways: (1) the ratio of English to Spanish exposure as reported on the LEAP-Q, (2) the difference between the composite scores for self-rated proficiency in English and Spanish, and (3) the ratio of average items produced on the verbal fluency task between English and Spanish. Below I address my original predictions with RT-CR measures from the picture-naming task. See Table 6 for the complete set of correlations

Hypothesis 3: I predicted that there would be a negative association between switch costs and inhibitory control. That is, better performance on the inhibitory control measures (Flanker and Simon Tasks) would be associated with smaller switching costs.

Results do not support Prediction 3. As show in the top of Table 6, the relationships between RT-CR for the Flanker effect, the Simon Effect, and English and Spanish switching costs were not significant. The correlations were small and ranged from $r = -0.17$ to $r = 0.16$. Note that these values are roughly centered around zero.

Hypothesis 4: I predicted that there would be a negative association between mixing costs and updating of working memory. That is, better performance on the updating measures (Spatial N-Back and S-SPAN tasks) would be associated with smaller mixing costs.

Results do not support Prediction 4. As shown in the bottom of Table 6, the relationships between S-Span score, the difference in reaction times for N-2 and N-1 trials on the Visual Spatial N-Back task, and English and Spanish mixing costs for reaction times were not significant. The correlations ranged from $r = -0.17$ to $r = 0.01$. Although there appears to be trend for the correlations to be negative, the correlations were very small.

Hypothesis 5: I predicted that there would be a positive correlation between Spanish proficiency and enhanced inhibitory control and updating ability.

As seen in Table 6, Results weakly support Prediction 5. The only relationship that approached significance ($r = .29$, $p = .07$) was between the Flanker effect for reaction time and the ratio of English to Spanish exposure (fifth column). This moderate correlation indicates that the more English (and the less Spanish) a person was exposed to, the larger the Flanker effect. This matches the prediction that higher Spanish exposure would be associated with enhanced inhibitory control. However, the remaining correlations were not significant.

IV. General Discussion

In this section, I will first summarize the results and then interpret them in terms of support for my hypotheses. Next I will discuss limitations and strengths of the current study. Lastly, I will address conclusions and future research directions.

Summary of Results

Summary of Color-Shape and Picture-Naming Tasks. For the Color-shape task there were no switching or mixing costs as measured by accuracy. Reaction time data revealed that there was a small switching cost (150 ms) associated with the color-shape task, but switching costs did not vary as a function of participant group. There also was a large mixing cost (490 ms). Lastly, SLL participants were 88 ms faster than BLs on the non-switch trials (repetition and simple), which was the only significant difference between groups.

For the Picture-Naming task, there were no differences related to switching costs when accuracy data were examined. For reaction time, there were significant differences between participants were faster on repetition trials and on Spanish language trials. However, these differences were qualified by the Trial Type X Language interaction. There were switching costs for both languages, but the costs were larger in Spanish (73 ms) than in English (30 ms). Additionally, there was an interaction between language and participant group. Whereas SLLs were significantly 51 ms faster when naming in Spanish than in English, BLs named items in Spanish at approximately the same speed as in English. Accuracy data revealed that mixing costs differed as a function of task language, with a larger mixing cost in English (.01) than in Spanish (-.01), although the mixing cost for each language group was not significantly difference from zero. Reaction time data showed that there was a small mixing cost (91 ms) associated with the Picture-Naming task in general (i.e., collapsed across task language and participant group).

Furthermore, mixing costs were 80 ms larger in English than in Spanish. There were no differences in overall reaction time as a function of participant group.

Reaction time on the Color-Shape task was then used to predict reaction time on the Picture-naming task as a function of participant group. When examining reaction times of individual trial types (simple, repetition, and switch) separately for both languages (English and Spanish), participant group acted as a moderating variable. For four of the six trial types (English Simple, Spanish Simple, English Repetition, and Spanish Switch), performance on the color-shape task was a significant predictor of the speed of picture-naming task performance for SLLs but *not* for BLs. That is, for SLLs the relationship between the two tasks was positive such that as color-shape reaction time increased, so did picture naming time. This relationship was not present for the bilinguals in that their slopes were not significantly different from zero.

Summary of Executive Control Tasks. For each executive control component, performance was compared between the two tasks used to measure each component. If the two tasks were tapping into similar processes, there should be a significant correlation between measures for the tasks.

The two working memory tasks were not correlated, with correlations hovering close to $r = 0.0$. The auto-symmetry span task is reliable measure, as performance on it has been shown to correlate well with performance on O-Span and R-Span tasks (Unsworth et al., 2009. To my knowledge, the visual-spatial n-back task has only been used in one study (Soveri, et al. 2011). It is possible that this task tested different aspects of working memory than the auto-symmetry span task, is not a valid measure of working memory, or both.

The second set of tasks was shifting tasks (global-local and number-letter). Accuracy and reaction time performance on these two tasks was moderately to strongly correlated (positive correlations), with a stronger correlation between two of the reaction time measures (Number-letter RT and Global-Local RT difference; $r = .60$) than with the overall accuracy on the two tasks ($r = .40$)

The third set of tasks was inhibitory control tasks (Flanker and Simon). Accuracy and reaction time performance on these two tasks were moderately to strongly correlated, with a moderate relationship between the two overall accuracy measures ($r = .40$) and a strong relationship between the two RT measures ($r = .58$).

In short, the shifting and inhibition tasks had concurrent reliability, but the working memory tasks did not.

Summary of Group differences in Executive Control. Participants did not differ across groups on the measures of working memory in terms of accuracy. For the visual-spatial n-back task, SLL participants were faster on correct trials than were BL participants. This is the second case in which SLL participants performed faster overall on a task (the first being in the Color-Shape task).

Participants performed similarly on the global-local shifting task. Accuracy and reaction time was similar across groups for the global-local task. However, accuracy was very low for all participants on one trial type (9% accuracy rate for local-mismatch trials). This outcome makes the results unreliable as participants were not performing the task correctly. It is possible that they were almost always responding to the Global-Number. This strategy would lead to high accuracy rates in both of the global trials (matched or mismatched) and in the local-matched trial as the global and local number are the same, but low accuracy in the local-mismatch as the global number never matches the local number. In other words, 75% of the time, participants could respond to the global number and answer correctly. The rest of the time, they had a 25%

change of guessing correctly. Participants performed even lower than that. I looked at participants' responses for incorrect local-mismatched trials and compared them to what the global number was for the trial. Only 1.1% of the time did the participants respond with the global figure's digit. It is unclear as to what strategy the participants were actually using. In retrospect, the task may have been too difficult.

For the Number-Letter shifting task, accuracy was the same across groups but there were significant differences in reaction time between SLL and BL participants. Again, SLLs had faster overall reaction times compared to bilinguals. More importantly, bilinguals had smaller mixing costs than SLL participants. This is evidence for enhanced executive control as bilinguals showed less of a slow down with the addition of switch trials.

The two inhibitory tasks, Flanker and Simon, had conflicting results. Accuracy on the Simon task was equivalent across groups. In contrast, accuracy on the Flanker task differed by participant group such that SLLs had a larger Flanker effect (by 1.3%) than did bilinguals. This is possible evidence for enhanced inhibitory control in the bilingual group – the opposite of what was seen in the Simon task. For reaction time, SLLs were faster than BLs on both tasks (63 ms faster on the Flanker and 43 ms on the Simon). This particular difference does not hold much meaning as the SLLs tended to be faster than BLs on several tasks.

Summary of Picture-Naming Task, Executive Control Tasks, and Proficiency. Relationships between performance on the picture-naming task (switching and mixing costs), the executive control tasks, and proficiency measures (English to Spanish exposure ratio, difference in self-rated proficiency composite scores, and the English to Spanish verbal fluency ratio) were explored. The only relationship that approached significance occurred between the Flanker effect

RT-CR and the ratio of English-to-Spanish exposure (*r* = .29, *p* = .066). There were no other significant relationships.

Support of Hypotheses

As described below, only a few of the results supported my hypothesis. The group differences were not as striking as originally predicted. However, due to a small sample size, it is too early to draw strong conclusions. The potential of the study is exciting, but the statistical power is disappointing. See Table 7 for a summary of predicted and actuals results.

Switching (non-verbal) task.

Hypothesis 1A: Switching costs. I predicted that more proficient bilinguals would have reduced switching costs compared to less proficient bilinguals. This would have been shown by the following pattern of results: the difference in mean reaction time between switch and repeat trials in the mixed block would be smaller for BLs compared to SLLs. Instead, results showed that the two groups had similar switching costs. That is, bilinguals did not demonstrate an advantage in inhibitory control on the non-verbal switching task.

Hypothesis 1B: Mixing costs. I predicted that if there were any differences in mixing costs between the two groups, the more proficient bilinguals should show smaller mixing costs than the less proficient bilinguals. This would be shown by the following pattern of results: the difference in mean reaction time between all trials in the simple block (single-task trials) and the repetition trials in the mixed block would be smaller for the proficient bilinguals than the less proficient bilinguals. Results showed that the two groups had similar mixing costs. Bilinguals did not demonstrate an advantage in updating goals in working memory on the non-verbal switching task compared to SLLs

Picture-Naming (Verbal) Task.

Hypothesis 2A: Switching costs. I predicted that both SLL and BL participants would exhibit switching costs. For the SLLs, these costs were expected to be asymmetrical such that switching to L1 would be more costly than switching to L2 due to a greater amount of inhibitory control required when switching into their dominant, native language. For the BLs, there would be no difference in switch costs size as a function of language as both languages should require the same amount of inhibitory control. This would have been shown by the following pattern of results: a switch cost would be evidenced by the longer RTs for switch trials than for repeat trials in the mixed block. For the less proficient bilinguals, switch costs would be larger when the switch is from Spanish to English (switch into L1) than vice versa. For the proficient bilinguals, the direction of the switch would not matter. Reaction time data did not support this hypothesis. There were no significant differences in switching costs as a function of participant group. The only significant difference was that SLLs were slower on English trials than on Spanish trials, whereas bilinguals did not differ in naming time as a function of language. This effect was primarily driven by repetition trials, such that RT was similar on switch trials for SLLs (English 931 ms = Spanish 903 ms), but different as a function of language on repetition trials (English 913 ms > Spanish 839 ms). This could be due to carry over effects of inhibition from English switch trials to English repetition trials and indication that SLLs have a more difficult time switching to English than to Spanish, but this is not how differential amounts of inhibition used is typically demonstrated.

Hypothesis 2B: Mixing costs. I predicted that high-proficiency bilinguals would have smaller mixing costs than low-proficiency bilinguals. This would be shown by the following pattern of results: the difference in reaction time between repeat trials in the mixed block and all single-task trials in the simple block would be smaller for the high-proficiency bilinguals (BLs) than the low-proficiency bilinguals (SLLs).

Reaction time data partially supported this hypothesis. The interaction between language, trial type, and participant group was significant, but it did not match my prediction. Both participant groups had similar mixing costs in English and BLs had larger mixing costs in Spanish (SLLs did not have any mixing costs). This is evidence for group differences in cognitive processing as a function of language, but if the data are reliable it indicates and advantage for the SLL. This is the opposite of what should occur based on prior research and theory. It is unclear as to why this pattern of results occurred or what the implications of this finding might be.

Individual differences in executive control.

The next three hypotheses were based on the assumption that my measures of individual differences in shifting, working memory, and inhibitory control were valid. I predicted that individual differences in these tasks would correlate switching and mixing costs.

Hypothesis 3: Switching costs. If general inhibitory control was related to the inhibitory control used in language processing, performance on two tests of inhibitory control (Flanker and Simon) would highly correlate with both the verbal and non-verbal switching tasks. That is, better performance on the Flanker and Simon tasks would be associated with smaller switching costs on the verbal task. If individual differences in inhibitory control were not associated with inhibitory control in the verbal task, then it is unlikely that the advantages bilinguals have shown on inhibition tasks in past research (e.g., Bialystok et al., 2004; Costa et al., 2008) solely comes from the use of inhibition in bilingual lexical access. Results revealed no significant correlations and the predictions were not supported. The lack of significant differences suggests that executive control may not be not as closely related to language switching as previously found – at least for participants in the current study. Independent cognitive mechanisms may be used for the language tasks that are not used in executive control tasks.

Hypothesis 4: Mixing costs. If updating of working memory is related to how goals are identified and actualized is related to bilingual lexical access, then the two measures of working memory (auto-symmetry and visual-spatial n-back) should highly correlate with mixing costs on the verbal switch task. That is, smaller mixing costs should be associated with better updating ability. Results revealed no significant correlations and the predictions were not supported. Again, the lack of significant differences suggests that participants use at least some independent mechanisms in the language switching task and the executive control tasks.

Hypothesis 5: If advantages in executive control are related to bilingualism, differences in the working memory and the inhibitory control tasks should be associated with language proficiency. That is, high-proficiency bilinguals should perform better on the working memory and inhibitory control tasks than low-proficiency bilinguals. This prediction was partially supported. Participants who had a more balanced/equal exposure to English and Spanish showed a smaller Flanker effect than participants who were exposed much more to one language than the other. That is, the more often participants was exposed to both languages, the better inhibitory control they had on the Flanker task. This demonstrates advanced cognitive control associated with bilingualism.

Relation to Prior Research.

Color-Shape task. Prior and MacWhinney (2010) and Prior and Gollan (2011) both used color-shape tasks nearly identical to the one used in the current study. Prior and MacWhinney's participants were monolinguals and highly proficient bilinguals (various L1s, highly proficient in

English). Both groups had switching costs and mixing costs. The switching costs were significantly smaller for bilinguals, but there was no difference between groups in mixing costs. Overall reaction time was similar across participant groups. Prior and Gollan's participants were monolinguals and highly proficient Mandarin-English or Spanish-English bilinguals. All groups had similar switching and mixing costs. Overall, the monolinguals and Spanish-English bilinguals were faster than the Mandarin-English bilinguals.

Results of the current study mirrored those of Prior and Gollan. There was no difference in switching cost or mixing costs as a function of participant group, and SLLs were faster than BLs overall. Taken together, this indicates that highly proficient bilinguals do not show reduced mixing costs on the color-shape task. Why might some bilinguals show reduced switching costs, as in Prior and MacWhinney's study? One possibility is that their participants' L1 was not as well controlled as in Prior and Gollan's study or in my study. The discrepancy in switching costs could be due to a confounding factor specific to Prior and MacWhinney's participants, such as coming from a variety of language backgrounds. Their participants all spoke English, but the second language was not controlled; there were 15 different second languages reported. In comparison, all of Prior and Gollan's participants were Mandarin-English, Spanish-English, or Hebrew-English bilinguals. Additionally, it is possible that Prior and Gollan's participants came from more similar socioeconomic backgrounds—a factor that has been shown to influence bilinguals' performance on executive control tasks (Tare & Linck, 2011). Notwithstanding, findings are variable for the color-shape switch task and it is difficult to make conclusions with confidence.

Picture-naming task. As summarized in the introduction, several studies have used the picture-naming task to study switching costs (and sometimes mixing costs) in bilinguals. In one of the first studies, Meuter and Allport (1999) examined switching costs with highly proficient and less proficient bilinguals from a variety of language backgrounds. For balanced bilinguals, switch costs were symmetric (i.e., similar in size) for both languages. For unbalanced bilinguals, switching costs were asymmetric with larger switching costs for their more dominant language than theirless dominant language.

Costa and Santesteban (2004) found similar patterns for switching costs on a picture-naming task. Their participants were from a variety of language backgrounds. Two groups of SLLs showed asymmetric switching costs, such that is was more difficult to switch back to their native language than their second language. In contrast, participants who were highly proficient in Spanish and Catalan had similar switching costs across languages.

Costa et al. (2006) examined the asymmetric switching cost phenomenon on a picture-naming task as function of L2 age of acquisition. Participants were either early Spanish-Basque bilinguals or late Spanish-English bilinguals. Results showed that regardless of language background, participants had symmetrical switching costs across L1 and L2.

Prior and Gollan (2011) included simple trials to their digit naming task and were able to compare switching and mixing costs across languages across two groups of early bilinguals. The participants had symmetric switching costs across languages. In contrast, mixing costs were asymmetric, such that mixing costs to the non-dominant language were smaller than to the dominant language.

In the current study, the differences in switching costs and mixing costs were not statistically different as a function of participant group (SLL vs. BL) or language (English vs. Spanish), making it too early to conclude whether my results match prior research. Looking at the pattern for switching costs, both SLLs and BLs were slower to switch into Spanish than to

English. This pattern would make sense for the BLs if they were Spanish dominant and would match the asymmetrical switch costs found by Meuter and Allport (1999) and Costa and Santesteban (2004). However, there is no theoretical explanation as to why SLL would have a more difficult time switching to their L2. The observed power for my study was low (.059). Thus, it may be too early to make a definite conclusion regarding switching costs and subsequently, inhibitory control. Looking at the pattern for mixing costs, both SLLs and BLs had larger mixing costs for Spanish than English, but this was not statistically reliable (observed power = .545). Prior and Gollan found the opposite pattern. Unbalanced bilinguals had larger mixing costs for the dominant language (typically English in the current study). Because the results of the current study do not fit with what prior research would predict and because the observed power is low, additional data needs to be collected before making conclusions. If current pattern holds, it would indicate the SLLs have superior goal updating ability when using Spanish compared to BLs. As this result would not be theoretically supported, it could be the result of a confounding factor associated with the SLLs but to a lesser extent with the BLs. Once additional data is collected, and there is no significant difference between participant groups, this would indicate equivalent ability for updating across participant groups.

General Intelligence – Ravens Advanced Progressive Matrices Test. Participants in Prior and Gollan's (2011) study completed a matrices subtest from the Kaufman brief intelligence test, which was similar to the Ravens test. They found that highly proficient early Spanish-English bilinguals had slightly, but significantly, lower matrices reasoning scores compared to English monolinguals and early Mandarin-English bilinguals. When examining switching and mixing costs on the color-shape task for Spanish-English bilinguals only, they found no correlations with the matrices score. They did not report the correlations for the other two groups or with the language-switching task.

Using the Raven's APM scores, I repeated Prior and Gollan's analysis. The bilinguals in my study (early Spanish-English bilinguals, as in Prior and Gollan) had slightly lower scores than the SLLs – however, this difference was only marginally significant. Similar to Prior and Gollan, there were no significant correlations for either language group between Raven's scores and switching or mixing costs on the color-shape task.

In summary, the relation between general intelligence and processing costs on the color-shape task were similar in my study and in Prior and Gollan (2011). The Spanish-English bilinguals tended to have lower matrices scores, but this was not associated with switching or mixing costs.

Executive control measures. Soveri et al. (2011) conducted a set of multiple regression analyses in which they used the participants' current ages, L2 ages of acquisition, and amount of L2 use to predict reaction time performance on the Flanker Task (Flanker effect), Simon Task (Simon effect), visual-spatial n-back task (n-back effect), number-letter task (switching and mixing costs), and rates that participants reported switching between languages (tendency to switch between languages, contextual switches, and unintentional switches). Analyses revealed that a younger age of L2 acquisition was associated with a smaller Simon effect and greater L2 exposure was associated with a smaller Flanker effect (see Table 8 for a full summary).

I repeated Soveri and colleagues' regression analyses using my measures of current age, L2 AoA (Age of Acquisition), and English-Spanish exposure ratio measures. Results revealed that a younger age of L2 acquisition was associated with smaller mixing costs on the number-letter task (see Table 8 for full summary).

In summary, the background measures for my study and Soveri et al (2011) were associated with mixing costs on the number-letter task. Unlike Soveri et al., my background measures were not significant predictors of the Flanker and Simon effects. These differences could be explained by the type of participants. Participants in Soveri et al were older (30-70 years of age), used both languages daily, and had no difference in speaking reading, writing, or speech comprehension skills across languages. My participants were younger (18-36 years of age), and many were more proficient in English than they were in Spanish. It is typically harder to find executive control task differences in younger people (e.g., Bialystok & Martin, 2004). Additionally, if the use of a second language truly does enhance cognitive control, balanced bilinguals who frequently use both languages would be more likely to show this due to more switching between languages.

In summary, results rarely matched those of prior research. Predicted effects were either absent, as in the case of similar switching costs for SLLs and BLs on the picture-naming task, or the opposite of what was predicted, as in of larger mixing costs of BLs on the picture-naming task. Next I will discuss possible limitations of the study that could explain the pattern of results obtained.

Limitations

For the current study, low statistical power is a possible explanation for the lack of significant effects. The sample size of 21 participants per group was less than ideal. I had wanted to have at least 30 participants per group, but due to difficulty recruiting qualified participants for the second language learner group, I was unable to meet this goal. Prior and MacWhinney (2010) and Prior and Gollan (2011) had approximately 45 participants per group (e.g., monolingual, second language learners, bilinguals). However, previous studies that reported finding bilingual advantages had equivalent or considerably smaller sample sizes. Costa and Santesteban's 2004 study had only 12 participants per group. Costa et al. (2006) had 12 or less, and Soveri et al. (2011) had 38. The current study had total 42 participants. Based on prior research, the sample size was ample to replicate the expected effects. Furthermore, more recent studies have had even larger samples and failed to find bilingual advantages. For example, Paap and Greenberg (2013) had 122 bilinguals and 151 monolinguals complete the Simon task. Although not statistically significant, bilinguals actually had a *larger* Simon effect compared to monolinguals, which is evidence for monolinguals having better cognitive control (the opposite of what would be expected). This suggests that bilingual advantages in executive control may not be as robust as once thought. Although I have low power in the current study, it is plausible that my results are accurate and would hold with a larger sample size.

An alternative explanation for the absence of bilingual advantages is that the tasks in the current study were not valid measures of executive control. The executive control tasks were selected based on use in previous studies in which significant differences associated with proficiency were found. For example, the color-switch task was directly modeled from Prior and MacWhinney (2010)--the number of trials, the stimuli, and timing were identical. The visual-spatial n-back, the Flanker, the Simon, and the number-letter task were nearly identical to Soveri et al.'s (2011). Additionally, I was not entirely lacking concurrent validity. The two inhibitory control tasks were moderately correlated as were the two shifting tasks. The working memory tasks were not reliably correlated, but they could have been emphasizing different aspects of working memory. A recent meta-analysis by Linck, Osthus, Koeth, and Bunting (2013) looked that the relation between working memory and L2 comprehension and production from 729 samples (n = 3,707). They found a robust positive relationship between WM and L2 processing.

The effect of working memory on L2 processing was greater for the executive control aspect of WM (versus storage) and for verbal measures of WM (versus non-verbal). The WM tasks in the current study were nonverbal measures, which could explain the absence of a relationship between performance on them and the various measures related to language.

If the tasks are valid, then is it the participants? It is impossible to match on every individual difference. A study by Tare and Linck (2011) were able to individually match 35 bilinguals and 35 monolinguals (from a larger set of over 1100) on several factors including age, income, education level, intelligence, and verbal ability. Once matched, the two groups showed no difference in working memory or on task switching. However, bilinguals showed significantly worse inhibitory control compared to monolinguals—the *opposite* of what is typically shown. The researchers concluded that advanced inhibitory control in bilinguals in previous research may have been an artifact and actually driven by other factors. Participants in the current study were not matched on as many factors as in Tare and Linck's study, but participants in the present study all had some college education and were approximately matched on age, years of education, intelligence (Raven's), and several measures of English proficiency.

In summary, observed power in the study was low, but the sample size was larger than some previous studies' that found significant differences in executive control between BLs and SLLs. All of the tasks had been used in at least one previous study in which bilinguals showed an advantage in executive control abilities. The SLLs and BLs were well matched on English proficiency and several other key variables. Although there are a few limitations in the current study, they are not concerning enough to discount the current findings. Moreover, the study also has several strengths.

Strengths

Despite the current study having a few possible limitations, it also had it strengths. First, participants' language background was tightly controlled. All BLs acquired both languages before the age of 6. All SLLs were currently studying Spanish at the time of testing. And none of the participants had knowledge of a third language. Second, to my knowledge, the design of the study was more comprehensive than any previous research. While most prior studies tested only executive control, or only task switching performance, the current study 1) collected self-reported language proficiency and tested receptive and productive vocabulary objectively; 2) tested each component of executive control with two separate nonverbal measures per component; 3) tested both verbal and non-verbal task switching ability using simple and mixed blocks so that switching and mixing costs could be examined; and 4) included a measure of general intelligence. Considering the thoroughness of the design, a plausible conclusion is that there were no bilingual advantages. The low statistical power reported as a limitation might reflect the lack of a real difference between SLLs and BLs and not a limitation in the size of the subject sample.

Conclusion

The goals of the current study were to identify (a) which executive control components play a role during lexical retrieval and (b) whether increased executive control ability is related to higher proficiency in a second language. In short, none of the executive control components were associated with performance on the picture-naming task, and there was no evidence for enhanced executive control ability in bilinguals. Although it was not an original aim of the study, I was able to find evidence for differential processing for the two participant groups on the switching tasks. The relationship between performance on the nonverbal color-shape switching task and the verbal picture-naming switching task was moderated by language background. The nonverbal task was a good predictor for the verbal task for SLLs, but not for BLs. This suggests qualitatively different cognitive processing on the verbal task for the two groups. There is recent research that fits well with this finding.

Calabria, Hernández, Branzi, and Costa (2012) had trilinguals complete a nonverbal color-shape task and verbal picture-naming task in their L1 and L2 or in their L1 and L3. Participants were highly proficient in their L2 and of low proficiency in their L3. The authors correlated performance (accuracy and RT) across the two tasks and did not any find significant relationships between the verbal and non-verbal switching tasks. They concluded that there was a qualitative difference in how the participants switch between task sets in the two tasks and that there is no generalizability of the cognitive processes from one task to another. Color-shape task performance was governed by domain-general executive control, whereas picture-naming task performance was governed by bilingual language control that was not fully subordinate to the domain-general executive control.

Calabria et al.'s (2012) results are similar to what I found for BLs when I tested language proficiency as a moderating variable in the relationship between the picture-naming task and the color-shape task. The BLs in the current study did not show strong relationships between performance on the two tasks. Although Calabria et al. had a group of participants who completed the picture-naming task in a highly proficient language (L1) and a less proficient language (L3), their data are not directly comparable to the data for the SLLs in the current study. Previous research (Costa & Santesteban, 2004; Linck, Schwieter, & Sunderman, 2012) has shown that trilinguals behave similarly when switching from L1 to L2 and from L1 to L3. In other words, even in a less proficient language, trilinguals behave as if they were using their L2.

This suggests that once bilingual language control is developed with an L2, it is transferred to the L3 as well. Participants in the current study had yet to develop that bilingual language control skill. Therefore, they used more domain-general executive control and thus their performance was closely related across the two tasks. In summary, this suggests that bilinguals do develop enhanced control, but that is specific to the verbal domain, rather than being part of domain-general executive control.

Together, my finding of language proficiency as a moderating factor in the cognitive process on the picture-naming task and Calabria's results suggest that although bilinguals develop enhanced cognitive processing, it is likely domain-specific rather than domain-general. This would account for the lack of differences between the BLs and SLLs in the current study on the executive control tasks. Additional studies that have failed to find these bilingual advantages include Paap and Greenberg (2013) and Tare and Linck (2011), both of which strictly controlled for individual differences amongst participants. If my results hold with additional participants, my findings will not support the majority of prior research, but will be consistent with several recent studies that use participants groups that are well-matched on linguistic factors (e.g., experience with a language) and cognitive factors (e.g., Raven's).

Future Directions

There is still much to be done with the current study. First and foremost, additional participants are needed to verify that the lack of significant differences between groups does not reflect lack of statistical power. It also would be helpful to have a monolingual control group to get a true "baseline" on the executive control measures. Given that most students receive second language instruction in high school, obtaining a pure monolingual group would be difficult, over even impossible, at least in a diverse environment such as Chicago. SLLs were at such a low level of proficiency that is unlikely that they would have developed enhanced cognitive control in such a short amount of time. Nonetheless, a monolingual control group would make for a cleaner comparison.

Secondly, there are several additional analyses that could be conducted that were beyond the scope of this paper. These include further examining mixing and switching costs in the color-shape task and the picture-naming task as a function of trial sequence in a run. For example, is a repetition trial the first repetition after a switch or the second (e.g., AB*B* vs. ABB*B*)? First repetitions could be susceptible to lingering reactive inhibition from a previous switch trial. The second repetition trial would have the benefit of additional time for the person to experience the release from inhibition. This same approach could be taken with the inhibitory tasks as well. Additionally, participants' responses in the picture-naming task could be analyzed based on how many times an item had been seen before, as well as whether the most recent presentation was in the same or a different language. Lastly, with the large amount of data obtained, other statistical techniques might be more appropriate for teasing out possible group differences, such as hierarchical linear modeling.

In conclusion, the results of the present study should be considered preliminary. With additional participants and statistical power, this study has the potential to add helpful evidence to the ongoing debate about when and if bilingualism enhances cognitive control. As the results stand now, I am inclined to say that bilinguals have the possibility of developing enhanced executive control, but that being bilingual is not sufficient to experience this advantage. When participants are well matched, proficiency alone is not enough to guarantee better control. Being bilingual might help develop better language control, but this appears to be very domain-specific. How using two languages modifies domain-general control remains a murky topic. Lastly, it remains unclear what component(s) of executive control are responsible for any advantage. Models of executive control (e.g., Miyake et al., 200) propose separable components, yet we still have not been able to pin down something more specific than “general cognitive control” and inhibition in a broad sense.

V. REFERENCES

Baayen R.H., Piepenbrock R., & Gulikers L. (1995). The CELEX Lexical Database Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

Baddeley, A. (1996). Exploring the Central Executive. *Quarterly Journal of Experimental Psychology: Section A, 49*, 5-28.

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., ... & Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, *10*, 344-380.

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science*, *7*, 325-339.

Bialystok, E. (2008). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*, 3-11.

Bialystok, E. (2010) Global–local and trail-making tasks by monolingual and bilingual children: Beyond inhibition. *Developmental Psychology, 46*, 93-105.

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, Aging, and Cognitive Control: Evidence from the Simon Task. *Psychology and Aging, 19*, 290-303.

Bilker, W. B., Hansen, J. A., Brensinger, C. M., Richard, J., Gur, R. E., & Gur, R. C. (2012). Development of abbreviated nine-item forms of the Raven's Standard Progress. *Assessment, 19,* 354-369.

Calabria, M., Hernández, M., Branzi, F.M., & Costa, A. (2011). Qualitative differences between bilingual language control and executive control: evidence from task-switching. *Frontiers in Cognition, 2*, 1-10.

Callejas, A., Lupiàñez, J., Funes, M. J., & Tudela, P. (2005). Modulations among the alerting, orienting and executive control networks. *Experimental Brain Research, 167*, 27–37.

Carlson, S., Martinkauppi, S., Rämä, P., Salli, E., Korvenoja, A., & Aronen, H.J. (1998). Distribution of cortical activation during visuospatial *n*-back tasks as revealed by functional magnetic resonance imaging. *Cerebral Cortex, 8,* 743-752.

Colomé, K. (2001). Lexical activation in bilinguals' speech production: language-specific or language-independent? *Journal of Memory and Language, 45*, 721-736

Costa, A., & Caramazza, A. (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish–English and English–Spanish bilinguals. *Bilingualism: Language and Cognition, 2*, 231-244.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language, 50*, 491-511.

Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 1057-1074.

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language, 41*, 365-397.

Costa, A., Hernández, M., and Sebastian-Galles, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition, 106*, 59-86.

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*, 135-149.

Dempster, F. N. (1993). Resistance to interference: Developmental changes in a basic processing mechanism. *Emerging Themes in Cognitive Development: Foundations, 1*, 3-27.

Dunn, L.M., & Dunn, D.M. (2007). *The Peabody Picture Vocabulary Test, Fourth Edition*. Bloomington, MN: NCS Pearson, Inc.

Dunn, L. M., Padilla, E. R., Lugo, D. E., & Dunn, L. M. (1986). *Test de Vocabulario en Imágenes Peabody: Adaptación Hispanoamericana*. Circle Pines, MN: American Guidance Service.

Emmorey, K., Luk, G., Pyers, J., &Bialystok, E. (2008). The source of enhanced cognitive control in bilinguals: Evidence from bimodal bilinguals. *Psychological Science, 19*, 1201-1206.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16,* 143-149.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *Journal of Experimental Psychology: General, 133*, 101-135.

Gollan, T.H.. Montoya, R.I., & Werner, G.A. (2002). Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology, 16*, 562-576.

Gollan, T.H., Salmon, D.P., Montoya, R.I., Galasko, D.R. (2011) Degree of bilingualism predicts age of diagnosis of Alzheimer's disease in low-education but not in highly educated Hispanics. *Neuropsychologia, 49*, 3826-3830.

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

Harnishfeger, K. K. (1995). The development of cognitive inhibition: Theories, definitions, and research evidence. In F.N. Dempster & C.G. Brainerds (Eds*.), Interference and inhibition in cognition*, 175-204, New York: Academic Press.

Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition*, *1*, 213-230.

Koch, I., Prinz, W., Allport, A. (2005). Involuntary retrieval in alphabet-arithmetic tasks: Task-mixing and task-switching costs. *Psychological Research, 69*, 252-261

Kroll, J. F., & Dijkstra, A. (2002). The bilingual lexicon. In R. Kaplan (Ed.), *Handbook of applied linguistics* (pp. 301-321). Oxford, England: Oxford University Press.

Kroll, J. F. & Stewart, E. (1994). Category inference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*, 149-174.

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2013). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, 1-23.

Linck, J.A., Schwieter, J.W., & Sunderman, G. (2012). Inhibitory control predicts language switching performance in trilingual speech production*. Bilingualism: Language and Cognition, 15*, 651-662.

Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition, 114*, 29-41.

Marion, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*, 940-967.

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*, 49-100.

Monsell, S. (1996). Control of mental processes. *Unsolved mysteries of the mind: Tutorial essays in cognition*, 93-148.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*, 134-140.

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology, 81*, 111-121.

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language, 40*, 25-40.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126*, 220-246.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (pp. 1-18). New York: Plenum

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology, 66*, 232-258.

Prior, A. & Gollan, T.H. (2011). Good language-switchers are good task-switchers: evidence from Spanish-English and Mandarin-English bilinguals. *Journal of the International Neuropsychological Society, 17,* 682-691.

Prior, A. & MacWhinney, B.N. (2010). A bilingual advantage in task switching. *Bilingualism: Language and Cognition*, *13*, 253-262.

Raven, J., Raven, J. C., & Court, J. H. (1998). Manual for Raven's Progressive Matrices and Vocabulary Scales: Section 4. The Advanced Progressive Matrices. San Antonio, TX: Harcourt Assessment.

Rubin, O. & Meiran, N. (2005). On the Origins of the Task Mixing Cost in the Cuing Task-Switching Paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 1477-1491.

Smith, E. E., & Jonides, J. (1999). *Storage and executive processes in the frontal lobes. Science, 283*, 1657-1661.

Soveri, A., Rodriguez-Fornells, A., & Laine, M. (2011). Is there a relationship between language switching and executive functions in bilingualism? Introducing a within-group analysis approach. *Frontiers in Psychology, 2*, 1-8.

Tare, M., & Linck, J. (2011). Exploring bilingual cognitive advantages when controlling for background variables. Presented at the 52nd Annual Meeting of the Psychonomic Society, Seattle, WA.

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory, 17*, 635-654.

Table 1

Mean responses (and SE) for second language learners (SLL) and bilinguals (BL) on self-report background measures from the LEAP-Q, and score on the Raven's Advanced Progressive Matrices test.

	Participant Group			ANOVA results
Background	SLL		BL	
Reported Age (years)	21.0 (.8)	=	19.7 (2.1)	*F*(1,39) = 2.13, MSE = 9.39, p = .15
Years of Education	14.8 (.4)	=	14.8 (.5)	*F*(1,39) < 1.0, MSE = 4.26, ns
Raven's (prop correct)	.57 (.04)	=	.49 (.03)	F(1,39) = 2.69, MSE = .026, p = .11
Acquisition (in years)	SLL		BL	
English Age of Acquisition	0.6 (.5)	<	3.7 (.5)	*F*(1,39) = 15.66, *MSE* = 6.52, *p* < .001
English Age fluent speak	4.6 (2.2)	=	6.2 (3.1)	*F*(1,39) = 3.70, *MSE* = 7.22, *p* = .06
English Age began to Read	4.7 (.3)	=	5.8 (.4)	*F*(1,39) = 1.51, *MSE* = 8.69, *p* > .05
English Age fluent reader	7.3 (.4)	=	7.3 (.6)	*F*(1,39) < 1.0, *MSE* = 5.31, *p* = .931
Spanish Age of Acquisition	11.9 (.6)	>	1.3 (.6)	*F*(1,39) = 179.86, *MSE* = 6.52, *p* < .001
Spanish Age fluent speak	19.4 (.6)	>	5.1 (.7)	*F*(1,27) = 139.40, *MSE* = 8.53, *p* < .001
Spanish Age began to Read	14.1 (.8)	>	6.9 (.8)	*F*(1,39) = 62.49, *MSE* = 8.69, *p* < .001
Spanish Age fluent reader	18.2 (1.0)	>	9.3 (1.0)	*F*(1,27) = 27.74, *MSE* = 16.64, *p* < .001
Exposure/Usage	SLL		BL	
%Exposure to Spanish	13% (10%)	<	35% (2%)	*F*(1,39) = 45.79, *MSE* = .01, *p* < .001
%Time Choose Read Span	11% (3%)	=	16% (3%)	*F*(1,39) < 1.0, *MSE* = .02, *ns*
%Time Choose Speak Span	22% (5%)	<	40% (5%)	*F*(1,39) = 6.38, *MSE* = .05, *p* < .05
Self-Rated Proficiency	SLL		BL	
English Speaking	9.48 (.18)	=	9.38 (.18)	*F*(1,39) < 1, *MSE* = 1.72, *ns*
English Understand	9.57 (.15)	=	9.62 (.156)	*F*(1,39) < 1, *MSE* = 1.71, *ns*
English Reading	9.43 (.18)	=	9.52 (.18)	*F*(1,39) < 1.0, *MSE* = 1.80, *ns*
English Composite	28.48 (.42)	=	28.52 (.42)	*F*(1,39) < 1.0, *MSE* = 10.74, *ns*
Spanish Speaking	5.69 (.33)	<	8.19 (.33)	*F*(1,39) = 38.08, *MSE* = 1.72, *p* < .001
Spanish Understand	6.62 (.32)	<	8.95 (.32)	*F*(1,39) = 35.53, *MSE* = 1.71, *p* < .001
Spanish Reading	6.86 (.35)	<	8.12 (.35)	*F*(1,39) = 9.299, *MSE* = 1.80, *p* < .01
Spanish Composite	19.17 (.81)	<	25.26 (.81)	*F*(1,39) = 36.33, *MSE* = 10.74, *p* < .001

Note: Self-Rated Proficiency is a 10 point scale, with 1 labeled "very low" proficiency and 10 labeled "very high" proficiency. For mean scores, < indicates the rating for SLL is significantly less than BL, > indicates the rating for SLL is significantly greater than BL, and = indicates the rating for SLL is not significantly different from BL.

Table 2

Mean (SE) raw scores on the PPVT and TVIP receptive vocabulary tests and mean (SE) number of items produced as a function of language and participant group on the Verbal Fluency (VF) productive vocabulary test.

	Receptive Vocabulary		
Task	SLL	BL	Average
PPVT	209.0 (1.7)	191.3 (3.3)	200.1 (2.3)
TVIP	105.0 (3.8)	107.1 (2.5)	106.0 (2.3)
	English Verbal Fluency		
Task	SLL	BL	Average
Letter	14.1 (.8)	12.0 (.8)	13.1 (.5)
Category	18.7 (.6)	15.10 (.6)	16.9 (.4)
Average	16.3 (.6)	13.52 (.6)	15.0 (.4)
	Spanish Verbal Fluency		
Task	SLL	BL	Average
Letter	8.1 (.5)	10.2 (.5)	9.2 (.4)
Category	8.2 (.5)	12.3 (.5)	10.4 (.4)
Average	8.3 (.6)	11.24 (.4)	9.8 (.3)
	All Verbal Fluency		
Overall Average	12.2 (.4)	12.4 (.4)	12.4 (.3)
Letter Difference (English-Spanish)	5.3	0.9	5.0
Category Difference (English – Spanish)	10.2	2.8	7.4
Average Difference	8.0	1.8	6.2

Table 3

Mean (SE) accuracy and reaction time (ms) for correct trials only as a function of trial type, and participant group on the Color-Shape Task.

	Accuracy		
	Participant Group		
Trial Type	Second Language Learner	Bilingual	Total
Switch	.94 (.01)	.93 (.01)	.95 (.01)
Repetition	.95 (.01)	.94 (.01)	.95 (.01)
Simple	.97 (.01)	.96 (.01)	.97 (.01)
Average	.95 (.01)	.95 (.01)	.95 (.01)
Cost			
Switching	-.01 (.01)	-.01 (.02)	-.01 (.01)
Mixing	-.02 (.01)	-.02 (.01)	-.02 (.01)
	Reaction Time (Correct Responses)		
Trial Type	Second Language Learner	Bilingual	Total
Switch	1084 (47)	1176 (47)	1130 (33)
Repetition	944 (44)	1016 (44)	980 (31)
Simple	437 (21)	541 (21)	489 (15)
Average	822 (34)	911 (34)	867 (24)
Cost			
Switching	141 (24)	160 (22)	150 (16)
Mixing	508 (36)	474 (36)	490 (25)

Table 4A

Mean (SE) proportion of trials answered correctly as a function of language, trial type, and participant group on the Picture-Naming Task.

	Participant Group		
	Second Language Learner	Bilingual	Total
Trial Type		English	
Switch	.96 (.02)	.97 (.02)	.97 (.01)
Repetition	.95 (.02)	.97 (.02)	.96 (.01)
Simple	.95 (.02)	.96 (.02)	.96 (.01)
Average	.95 (.02)	.97 (.02)	.96 (.01)
Cost			
Switching	.01	.00	.00
Mixing	.01	.01	0.01
Trial Type		Spanish	
Switch	.97 (.01)	.97 (.01)	.97 (.01)
Repetition	.95 (.02)	.97 (.02)	.96 (.01)
Simple	.97 (.01)	.96 (.01)	.97 (.01)
Average	.96 (.01)	.97 (.01)	.97 (.01)
Cost			
Switching	.02	.01	.01
Mixing	-.02	.00	-.01
Trial Type		Total	
Switch	.96 (.01)	.97 (.01)	.97 (.01)
Repetition	.95 (.016)	.97 (.016)	.96 (.01)
Simple	.96 (.016)	.96 (.016)	.96 (.01)
Average	.96 (.01)	.97 (.01)	.96 (.01)
Cost			
Switching	.01	.00	.01
Mixing	-.01	.01	-.001

Table 4B
Mean (SE) reaction time (ms) for correct trials only as a function of language, trial type, and participant group on the Picture-Naming Task.

	Participant Group		
	English		
Trial Type	Second Language Learner	Bilingual	Total
Switch	931 (31)	979 (31)	955 (22)
Repetition	913 (31)	937 (31)	925 (22)
Simple	782 (27)	804 (27)	793 (19)
Average	875 (28)	907 (31)	891 (20)
Cost			
Switching	18 (13)	43 (13)	30 (9)
Mixing	131 (16)	133 (16)	132 (11)
	Spanish		
Trial Type	Second Language Learner	Bilingual	Total
Switch	903 (37)	988 (37)	946 (27)
Repetition	839 (31)	907 (31)	873 (22)
Simple	815 (26)	831 (26)	823 (18)
Average	907 (28)	909 (31)	881 (22)
Cost			
Switching	64 (13)	81 (13)	73 (10)
Mixing	24 (13)	76 (13)	50 (9)
	Average		
Trial Type	Second Language Learner	Bilingual	Total
Switch	917 (33)	984 (33)	950 (24)
Repetition	876 (30)	922 (30)	899 (21)
Simple	798 (25)	818 (25)	808 (18)
Average	864 (29)	909 (29)	886 (20)
Cost			
Switching	41 (9)	62 (11)	51 (7)
Mixing	77 (12)	104 (12)	91 (9)

Table 5

Correlations between accuracy (proportion correct) and reaction time (ms) measures for executive control tasks.

	Accuracy	Accuracy Diff	RT-CR	RT-CR Diff
Working Memory				
Visual Spatial N-Back Task				
Auto-Symmetry Span Task	Accuracy	Accuracy Diff (N2-N1)	RT-CR	RT-CR Diff (N2-N1)
Span Score	.17	.18	-.14	**.09**
Shifting				
Number-Letter Task				
Global-Local Task	Accuracy	Accuracy Diff (Switch - Rep)	RT-CR	RT-CR Diff (Switch - Rep)
Accuracy	**.40***	.10	-.07	.17
Accuracy	.29$^{\#}$	**-.31**	-.02	.12
RT	.18	.31$^{\#}$	**.05**	-.16
RT-CR Diff (Switch – Rep)	.32*	-.21	.60*	**-.08**
Inhibitory Control				
Flanker Task				
Simon Task	Accuracy	Accuracy Diff (Incong-Cong)	RT-CR	RT-CR Diff (Incong-Cong)
Accuracy	**.40****	.32*	.19	-.23
Accuracy Diff (Incong-Cong)	.25	**.29**$^{\#}$	.19	-.16
RT-CR	.12	-.07	**.58****	.08
RT-CR Diff (Incong-Cong)	-.18	-.16	-.31*	**.16**

Note: ** $p < .01$, * $p < .05$, $^{\#}$ $p \leq .10$; correlations expected to be significant are bolded.

Table 6

Top: correlations between inhibitory control task performance (Flanker and Simon effects), proficiency measures (ratio of English-to-Spanish exposure, differences between English and Spanish self-rated proficiency composite score, English-to-Spanish verbal fluency [VF] scores), and picture-naming task (PNT) switching costs. Bottom: correlations between working memory task performance (auto-symmetry spans score and visual-spatial n-back [VSNB] difference), and picture-naming task mixing costs.

	Inhibitory Control		Proficiency			PNT Switching Costs (ms)	
	Flanker Effect	Simon Effect	Eng:Span Exposure	Self-rated prof diff (E-S)	Eng:Span VF	English	Spanish
Flanker Effect		.16	.29#	.12	.22	.03	-.06
Simon Effect			.12	.01	.08	-.13	-.06
Eng:Span Exposure				.60**	.76**	-.13	-.05
Self-rated prof diff (E-S)					.60**	-.20	-.14
Eng:Span VF						-.08	.14
Eng Switching Costs (ms)							.09

	Working Memory		Proficiency			PNT Mixing Costs (ms)	
	S-Span Score	VSNB N2-N1 RT-CR diff	Eng:Span Exposure	Self-rated prof diff (E-S)	Eng:Span VF	English	Spanish
S-Span Score		.09	-.10	.13	.01	.01	-.09
VSNB N2-N1 RT-CR diff			-.04	-.17	-.02	-.10	-.08
Eng:Span Exposure				.60**	.76**	.00	-.38*
Self-rated prof diff (E-S)					.60**	-.08	-.57**
Eng:Span VF						.10	-.42**
Eng Mixing Costs (ms)							.09

Note: ** $p < .01$, * $p < .05$, # $p \leq .10$

Table 7

Comparison of Original Predictions to Outcomes

Hypotheses	Task	Prediction	Outcome: Accuracy	Outcome: Reaction Time
1A.	Color-Shape	Switching Costs: BL < SLL	BL = SLL	BL = SLL
1B.	Color-Shape	Mixing Costs: BL < SLL	BL = SLL	BL = SLL
2A.	Picture Naming	Switching Costs: BL < SLL SLL only - Switching Costs: Eng > Span BL only: Switching Costs: Eng = Span	BL = SLL Eng = Span Eng = Span	BL = SLL Repetition trials only : Eng > Span Eng = Span
2B.	Picture Naming	Mixing Costs: BL < SLL	BL > SLL#	Repetition trials only : Eng > Span
3.	Inhibition	Simon/Flanker Effect (Incong-Cong): BL < SLL	Flanker only: **BL < SLL**	Flanker only: **BL < SLL**#
4.	Working Memory	BL > SLL	BL = SLL	BL = SLL
5.	Proficiency and Executive Control	Positive correlation between proficiency and executive control	BL =SLL	Flanker effect: Eng: Span exposure positively associated w/ Flanker Effect

Table 8

Comparison and summary of the multiple regressions analyses from Soveri et al. (top) and the current study (bottom): background variables as predicates of processing cost in RTs on executive control tasks.

Soveri et al. (2011)					
Variable	Flanker effect	Simon effect	N-back effect	Number-Letter task	
				Switching Cost	Mixing Cost
	B	*B*	*B*	*B*	*B*
Constant	78.57**	-16.53	163.55	145.33	-104.56
Current Age	0	0.31	-0.55	3.23	4.13
L2 AoA	-0.34	8.95*	7.12	6.64	37.63
Everyday use	-0.43*	0.27	1.04	-0.49	1.93
R^2	.16	.22	.04	.15	.28
F	2.13	3.14*	0.46	1.85	3.95*

Current Study					
Variable	Flanker effect	Simon effect	N-back effect	Number-Letter task	
				Switching Cost	Mixing Cost
	B	*B*	*B*	*B*	*B*
Constant	76.89	11.63	226.53	278.50	22.56
Current Age	-1.09	0.56	-6.50	-4.14	2.25
L2 AoA	0.49	0.73	-4.16	-1.63	4.89*
Everyday use	0.61	0.02	2.18	3.61	1.43
R^2	.13	.03	.07	.10	.27
F	1.82	0.41	0.88	1.27	4.14*

Figure 1A

Predicted values for English simple trial RT illustrating the interaction of participant group and color-shape simple trial RTs that are one standard deviation above and below the mean.

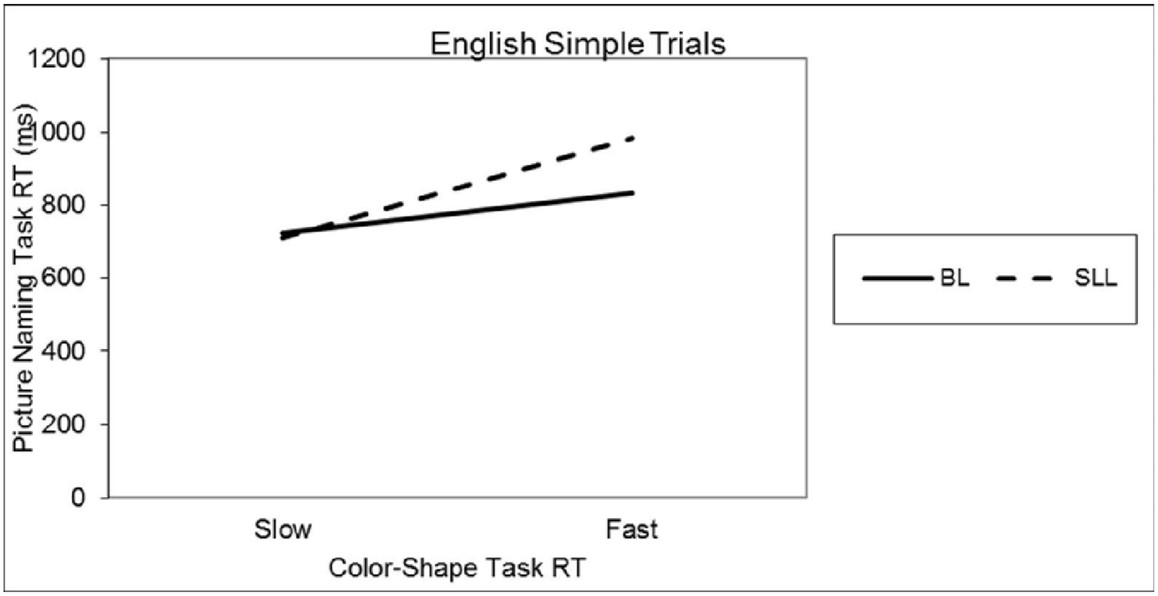


Figure 1B

Predicted values for Spanish simple trial RT illustrating the interaction of participant group and color-shape simple trial RTs that are one standard deviation above and below the mean.

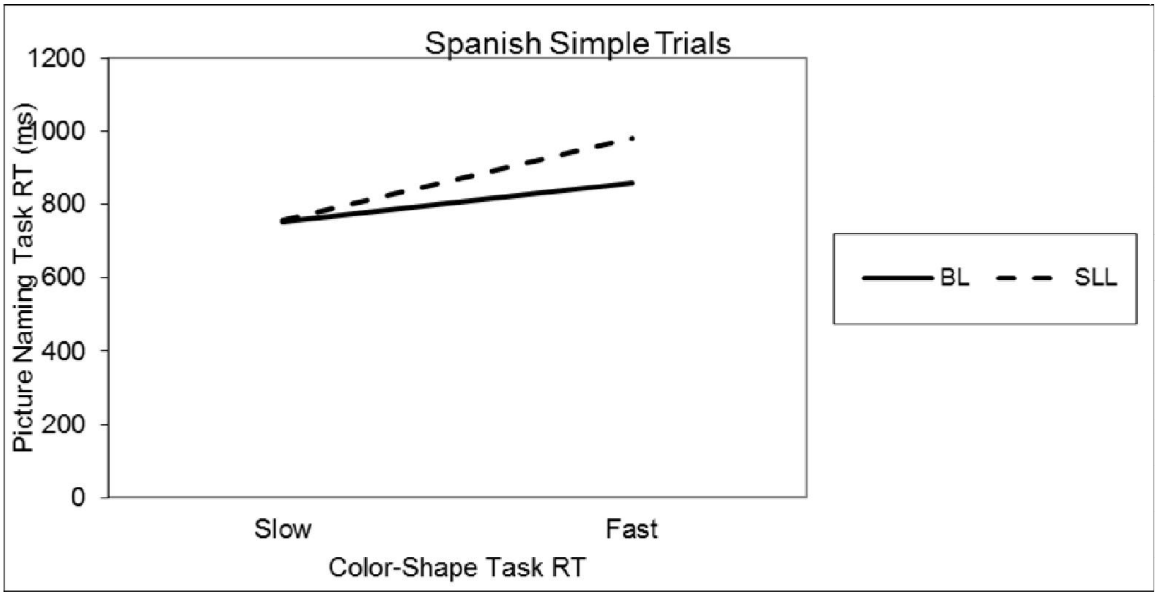


Figure 1C

Predicted values for English repetition trial RT illustrating the interaction of participant group and color-shape simple trial RTs that are one standard deviation above and below the mean.

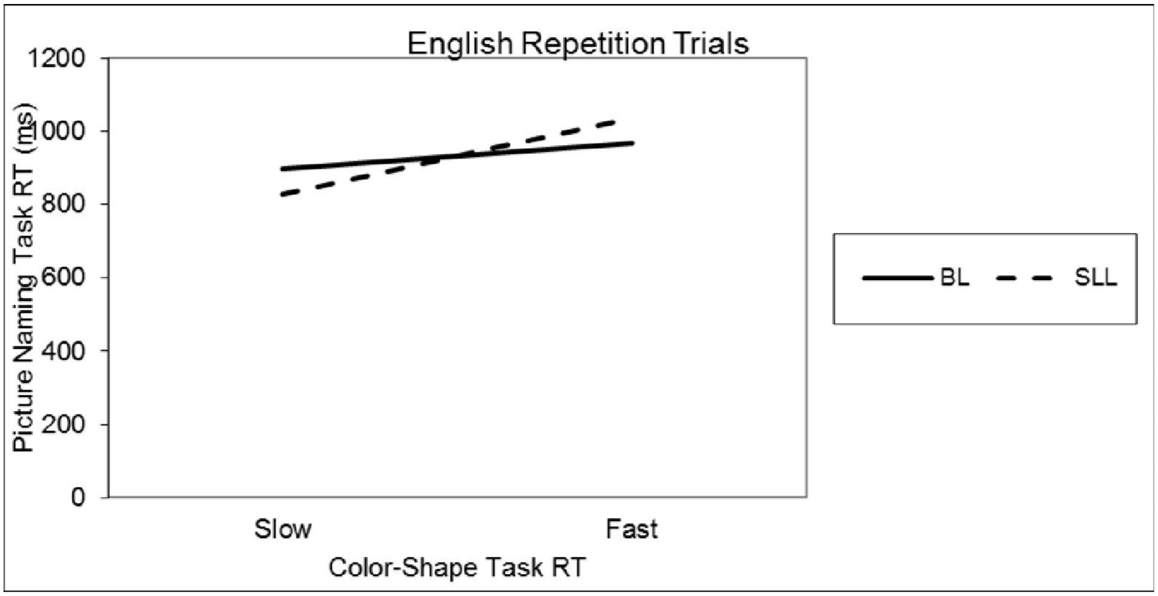


Figure 1D

Predicted values for Spanish repetition trial RT illustrating the interaction of participant group and color-shape simple trial RTs that are one standard deviation above and below the mean.

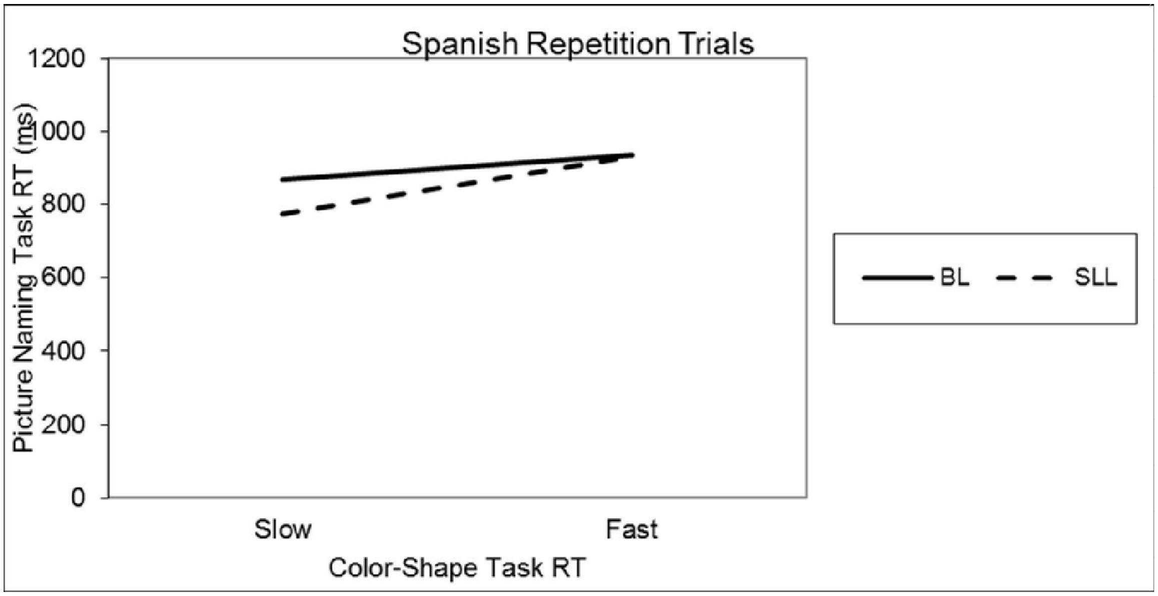


Figure 1E

Predicted values for English switch trial RT illustrating the interaction of participant group and color-shape simple trial RTs that are one standard deviation above and below the mean.

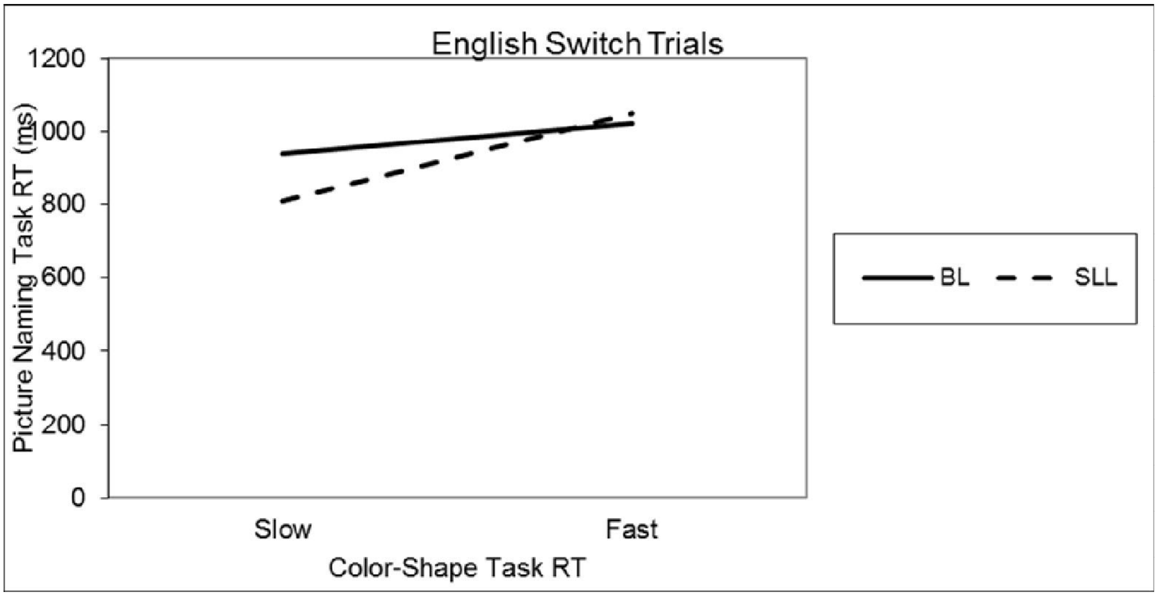


Figure 1F

Predicted values for Spanish switch trial RT illustrating the interaction of participant group and color-shape simple trial RTs that are one standard deviation above and below the mean.

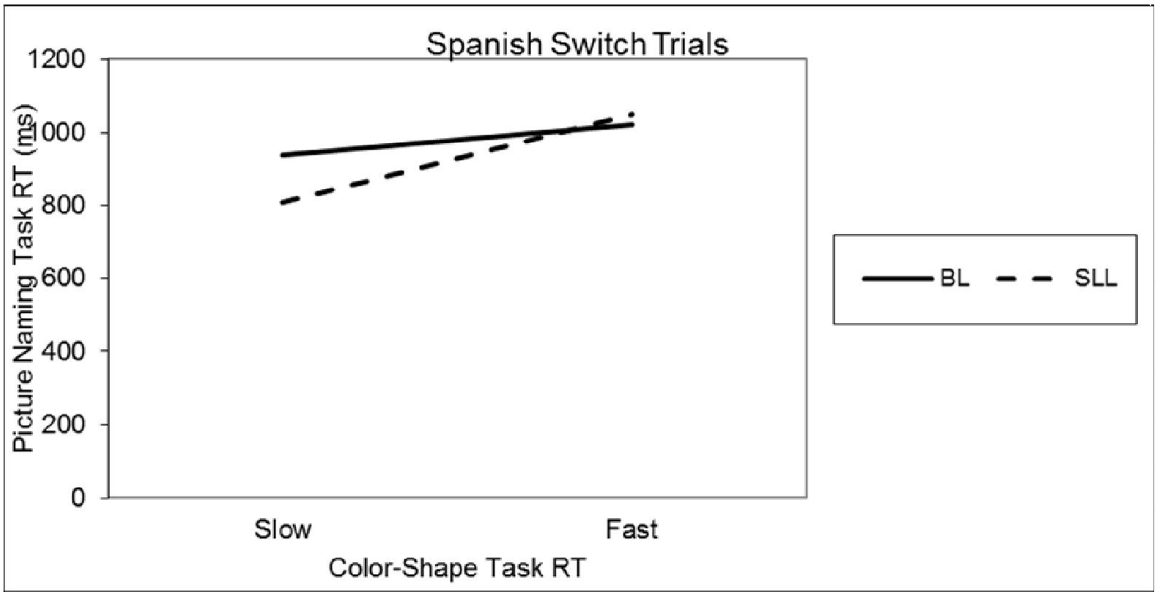


Figure 2A

Mean (SE) proportion correct as a function of trial type, and participant group on the Visual-Spatial N-back Task.

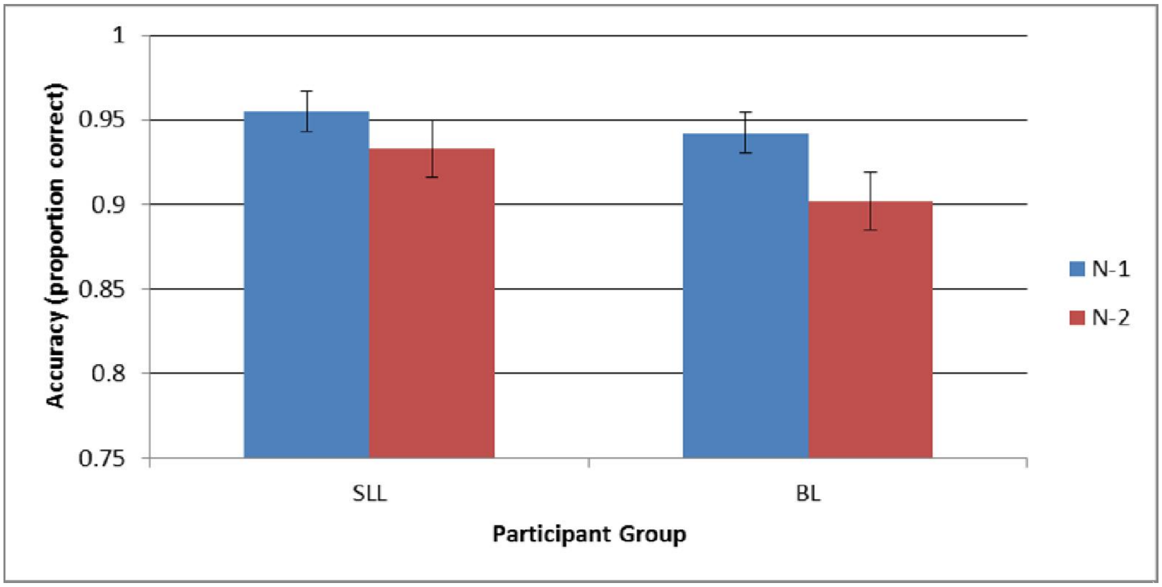


Figure 2B

Mean (SE) reaction time (ms) for correct trials only as a function of trial type, and participant group on the Visual-Spatial N-back Task.

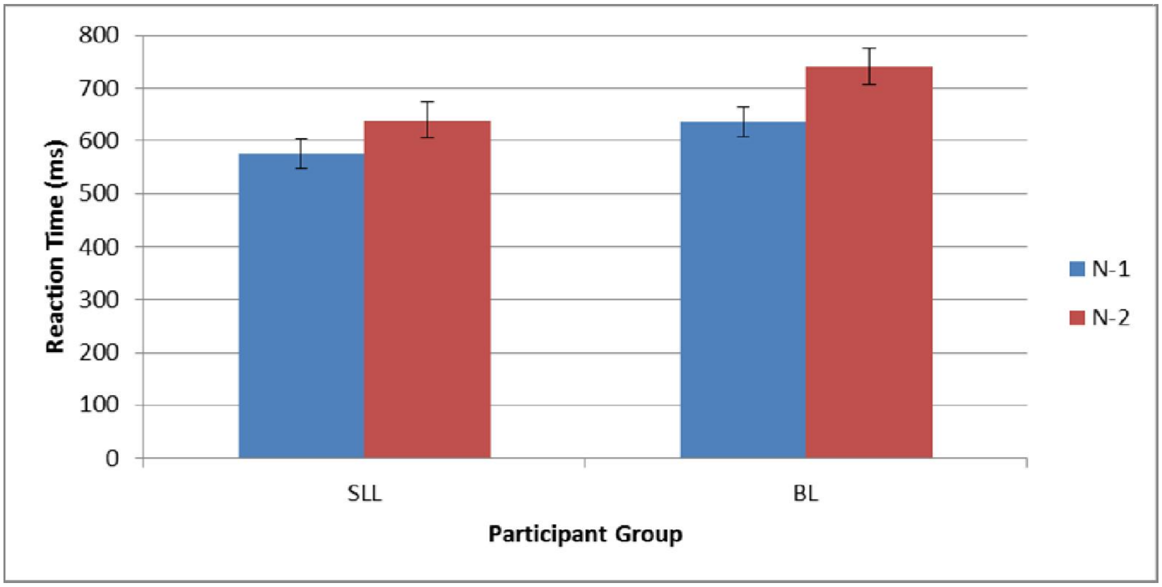


Figure 3A

Mean (SE) proportion correct as a function of trial type, and participant group on the Global-Local Task.

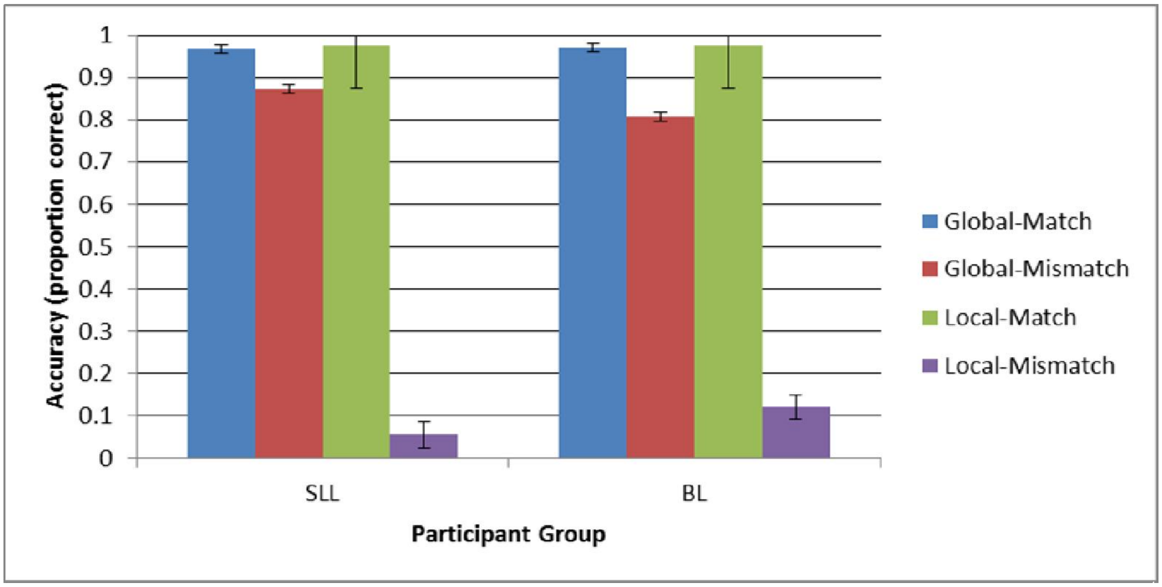


Figure 3B

Mean (SE) reaction time (ms) for correct and incorrect trials as a function of trial type, and participant group on the Global-Local Task.

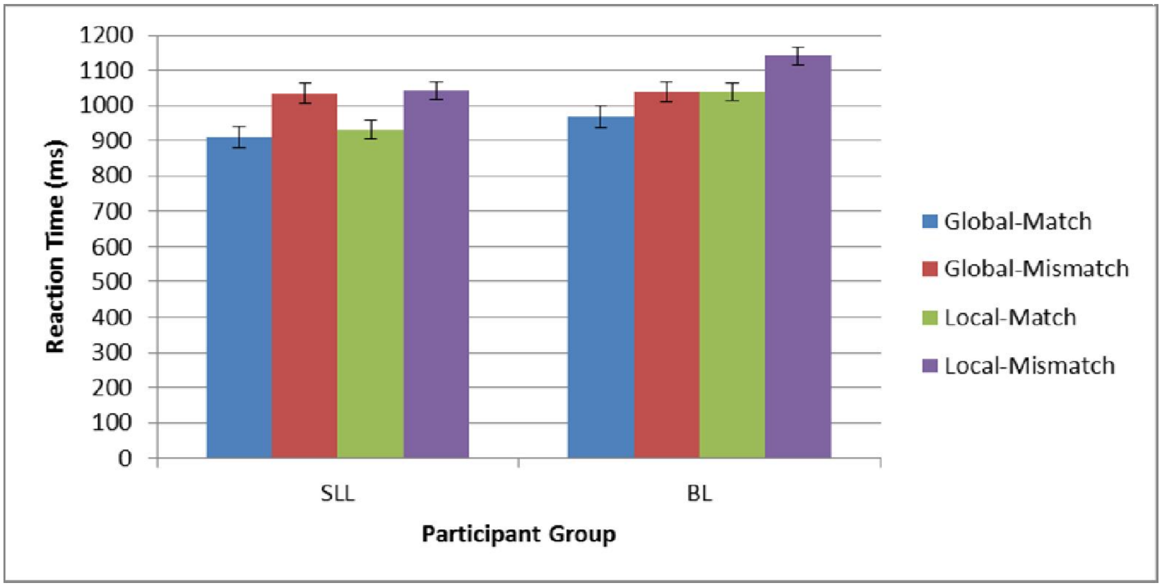


Figure 4A

Mean (SE) proportion correct as a function of trial type, and participant group on the Number-Letter Task.

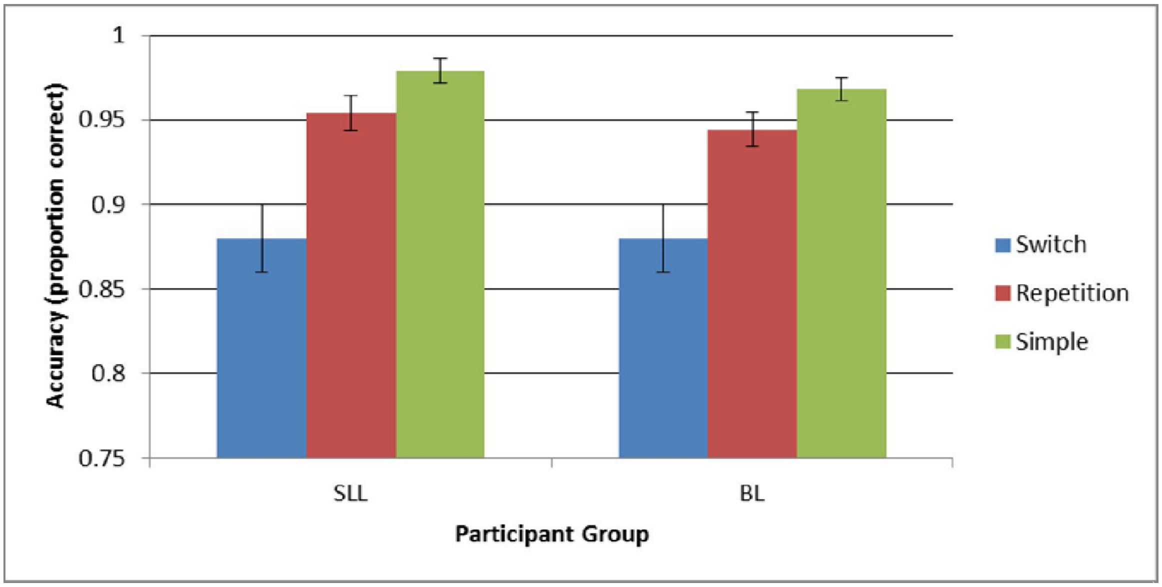


Figure 4B

Mean (SE) reaction time (ms) for correct trials only as a function of trial type, and participant group on the Number-Letter Task.

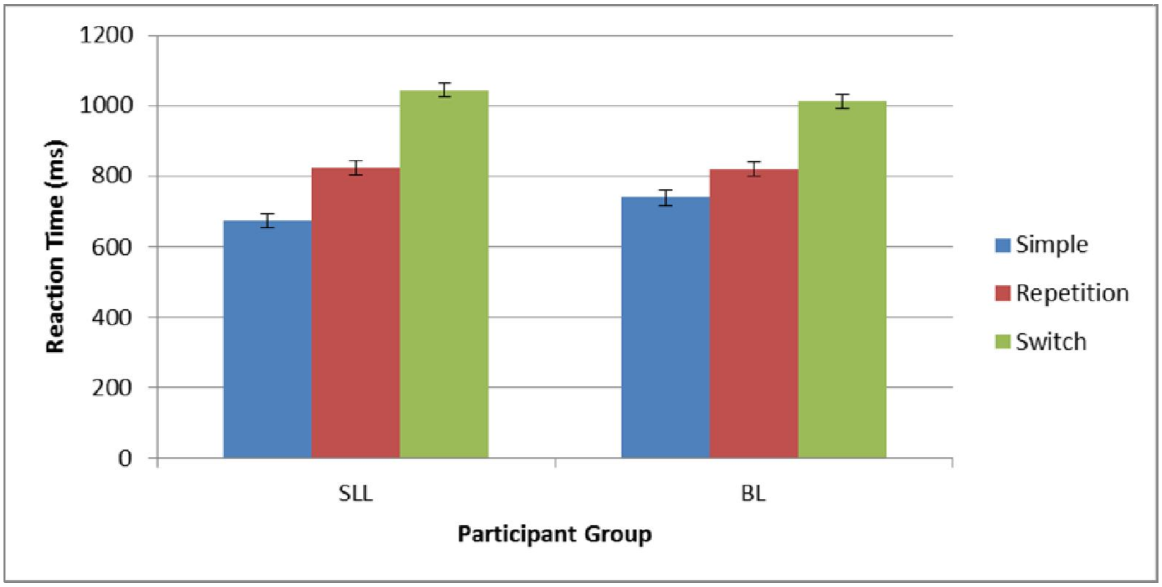


Figure 5A

Mean (SE) proportion correct as a function of trial type, and participant group on the Simon Task.

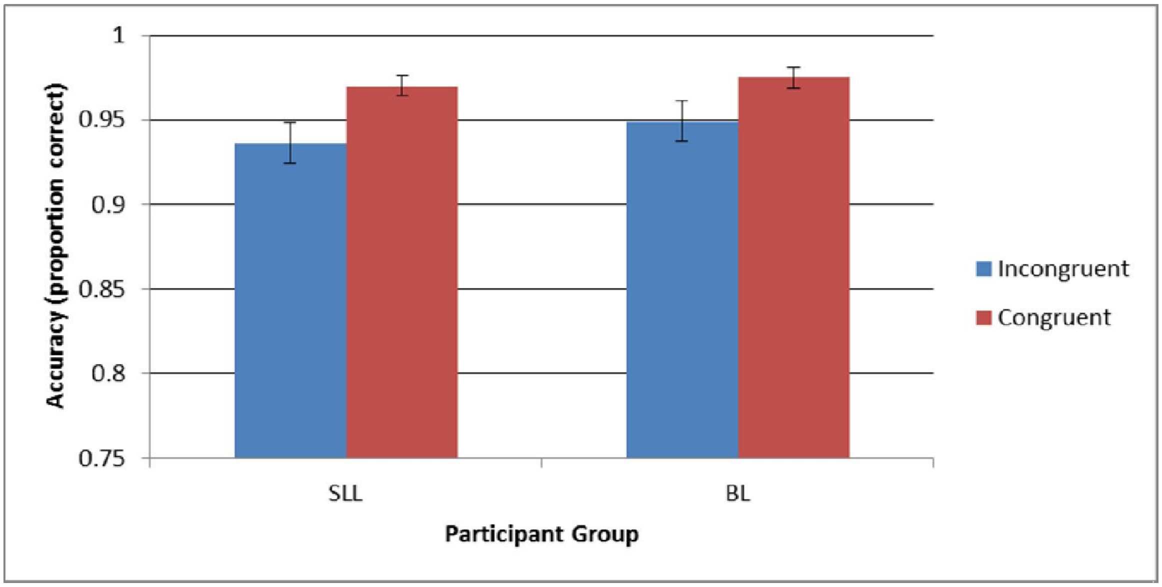


Figure 5B

Mean (SE) reaction time (ms) for correct trials only as a function of trial type, and participant group on the Simon Task.

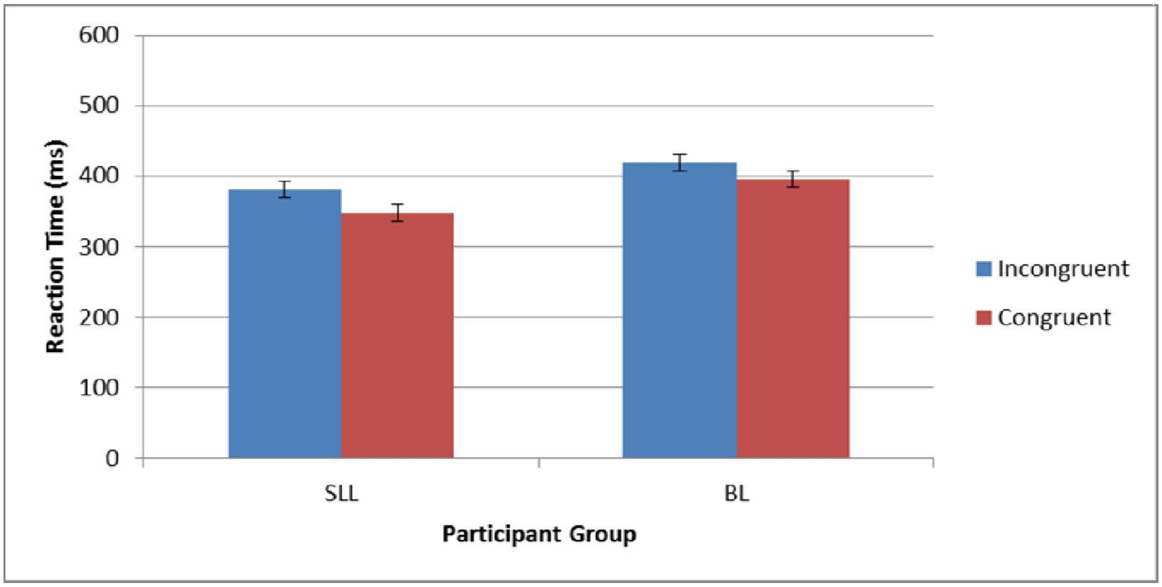


Figure 6A

Mean (SE) proportion correct as a function of trial type, and participant group on the Flanker Task.

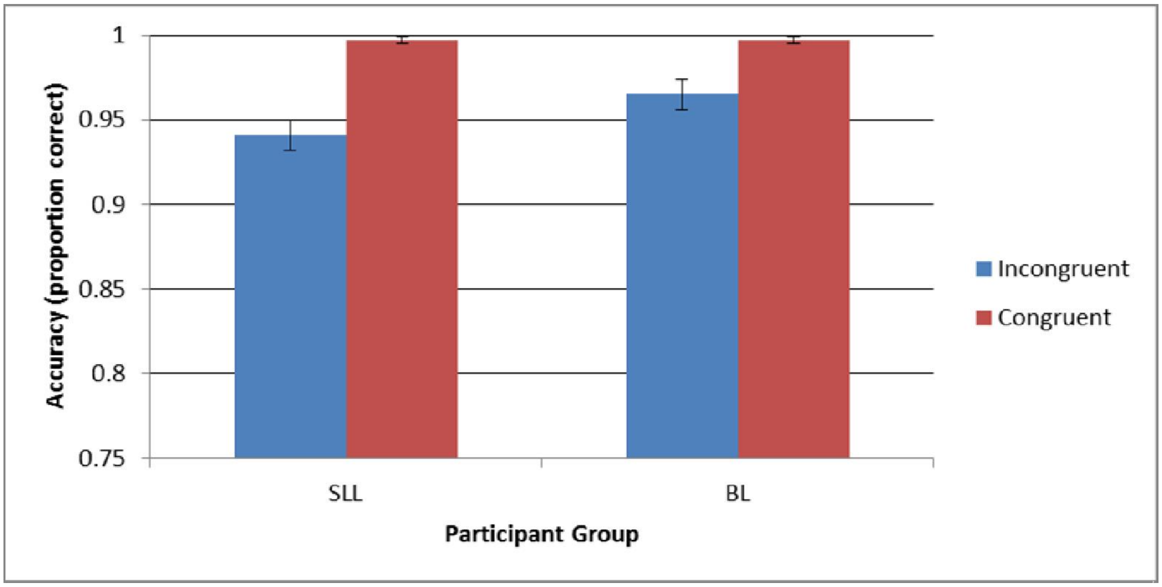
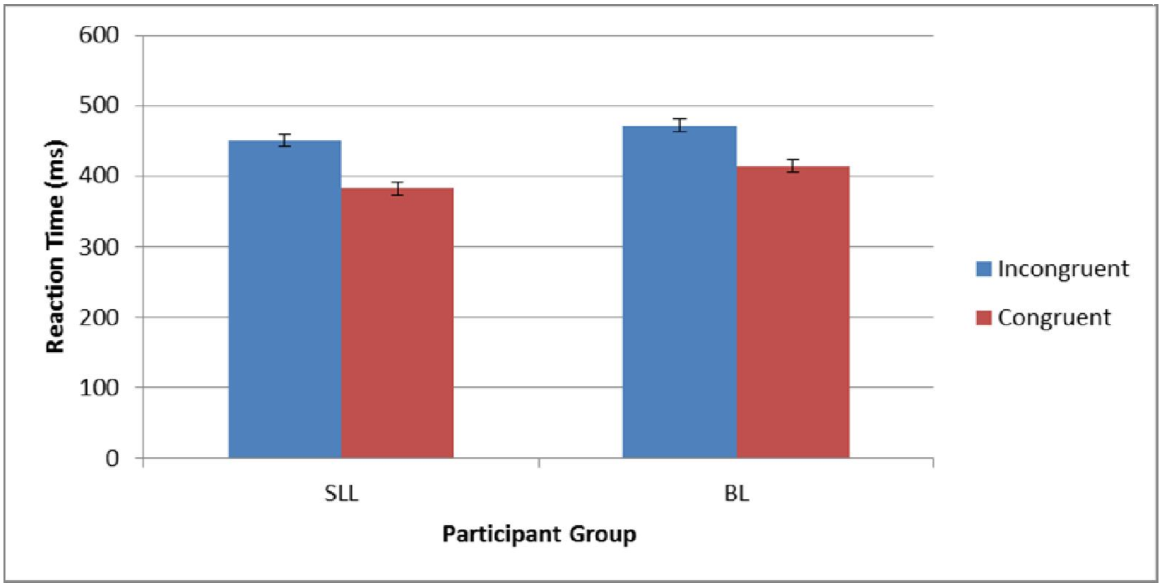


Figure 6B

Mean (SE) reaction time (ms) for correct trials only as a function of trial type, and participant group on the Flanker Task.



Appendix A: Language Experience and Proficiency Questionnaire (LEAP-Q)

Age:		Date of Birth:		Sex: Female ☐ Male ☐

(1) Please list all the languages you know **in order of dominance**:

1	2	3	4	5

(2) Please list all the languages you know **in order of acquisition** (your native-language first):

1	2	3	4	5

(3) Please list what percentage of the time you are currently and on average **exposed** to each language. (*Your percentages should add up to 100%*):

List languages here:					
List percentage here:					

(4) When choosing to read a text available in all your languages, in what percentage of cases would you **choose to read it in each of your languages**? Assume that the original was written in another language, which is unknown to you. (*Your percentages should add up to 100%*):

List languages here:					
List percentage here:					

(5) When choosing a language to speak with a person who is equally fluent in all your languages, what percentage of time would **you choose to speak each language**? Please report percent of total time (*Your percentages should add up to 100%*):

List languages here:					
List percentage here:					

(6) Please name the **cultures with which you identify**. On a scale from 0-10 (0 = no identification, 5 = moderate identification, 10 = complete identification) please rate the extent to which you identify with each culture:

List cultures here:					
List extent of identification here:					

(Examples of possible cultures include US-American, Chinese, Jewish-Orthodox, etc.)

(7) How many years of formal education do you have? __________ years
Please check your highest education level (or the approximate US equivalent to a degree obtained in another country)

☐ Less than High School
☐ High School
☐ Professional Training
☐ Some College
☐ College
☐ Some Graduate School
☐ Masters
☐ Ph.D./M.D./J.D.
☐ Other:_________________

(8) Date of immigration to the USA , if applicable (MM-DD-YYYY): ____
If you have ever immigrated to another country, please provide name of country and date of immigration here:
____

(9) Have you had a vision problem ☐, hearing impairment ☐ , language disability ☐, or learning disability ☐? (Check all applicable. If yes, please explain, including corrections:____

All questions below refer to your knowledge of **ENGLISH**.

(1) Age when you…:

	began *acquiring* English	became *fluent* in English	began *reading* in English	became fluent *reading* in English
List age here:				

(2) Please list the number of years and months you **spent in each language environment**:

	Years	Months
A country where ENGLISH is spoken		
A family where ENGLISH is spoken		
A school and/or work environment where ENGLISH is spoken		

(3) On a scale from 0 – 10 (where 0 = none, 5 = adequate, 10 = perfect), please rate your level of **proficiency** in speaking, understanding, and reading English:

Ability:	Speaking	Understanding spoken language	Reading
Proficiency:			

(4) On a scale from 0 – 10 (where 0 = not a contributor, 5 = moderate contributor, 10 = most important contributor), please rate how much the following factors **contributed to you learning** English:

	Rating		Rating
Interacting with friends		Language tapes/self-instruction	
Interacting with family		Watching TV	
Reading		Listening to the radio	

(5) On a scale from 0 – 10 (where 0 = never, 5 = half the time, 10 = always), please rate the extent to which you are **exposed** to English in the following contexts:

	Rating		Rating
Interacting with friends		Language tapes/self-instruction	
Interacting with family		Watching TV	
Reading		Listening to the radio	

(6) In your perception, how much of a **foreign accent** do you have in English (circle one)?

0------------1-------------2------------3-------------4-------------5------------6------------7-------------8------------9-------------10

None Moderate Pervasive

(7) Please rate how frequently others identify you as a non-native speaker based on your accent in English (circle one):

0------------1-------------2------------3-------------4-------------5------------6------------7-------------8------------9-------------10

Never half the time Always

All questions below refer to your knowledge of **SPANISH**.

(1) Age when you…:

	began *acquiring* Spanish	became *fluent* in Spanish	began *reading* in Spanish	became fluent *reading* in Spanish
List age here:				

(2) Please list the number of years and months you **spent in each language environment**:

	Years	Months
A country where SPANISH is spoken		
A family where SPANISH is spoken		
A school and/or work environment where SPANISH is spoken		

(3) On a scale from 0 – 10 (where 0 = none, 5 = adequate, 10 = perfect), please rate your level of **proficiency** in speaking, understanding, and reading Spanish:

Ability:	Speaking	Understanding spoken language	Reading
Proficiency:			

(4) On a scale from 0 – 10 (where 0 = not a contributor, 5 = moderate contributor, 10 = most important contributor), please rate how much the following factors **contributed to you learning** Spanish:

	Rating		Rating
Interacting with friends		Language tapes/self-instruction	
Interacting with family		Watching TV	
Reading		Listening to the radio	

(5) On a scale from 0 – 10 (where 0 = never, 5 = half the time, 10 = always), please rate the extent to which you are **exposed** to Spanish in the following contexts:

	Rating		Rating
Interacting with friends		Language tapes/self-instruction	
Interacting with family		Watching TV	
Reading		Listening to the radio	

(6) In your perception, how much of a **foreign accent** do you have in Spanish (circle one)?

0------------1-------------2-------------3-------------4-------------5-------------6-------------7-------------8------------9-------------10

None Moderate Pervasive

(7) Please rate how frequently others identify you as a non-native speaker based on your accent in Spanish (circle one):

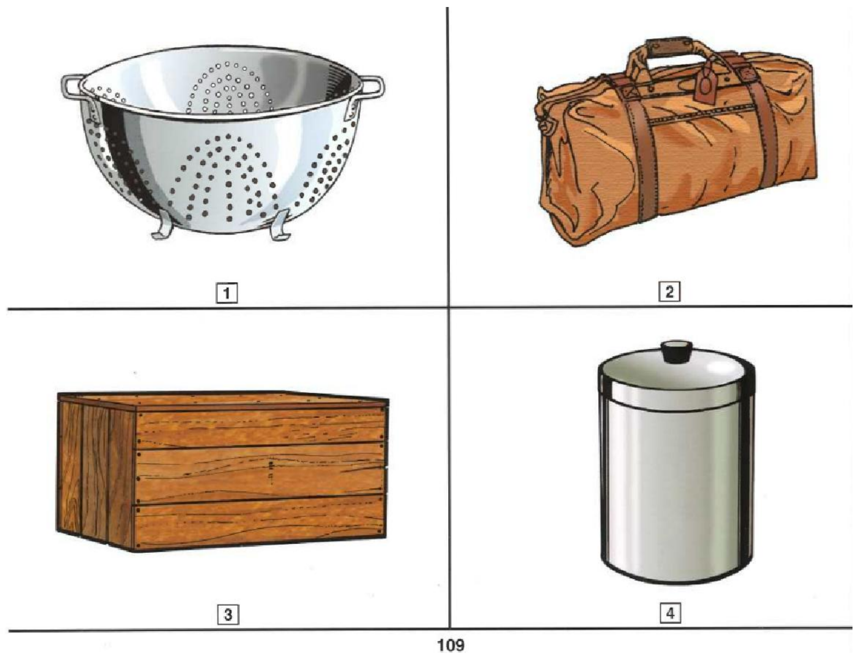
0------------1-------------2-------------3-------------4-------------5-------------6-------------7-------------8------------9-------------10

Never half the time Always

Appendix B: PPVT/TVIP

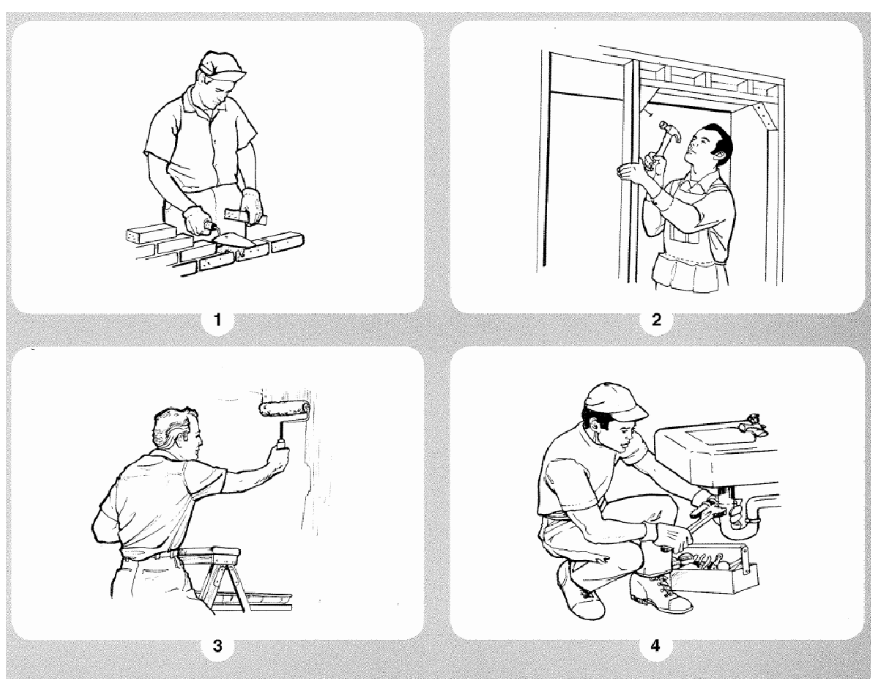
PPVT Example:

Word: Luggage
Answer: 2



TVIP Example:

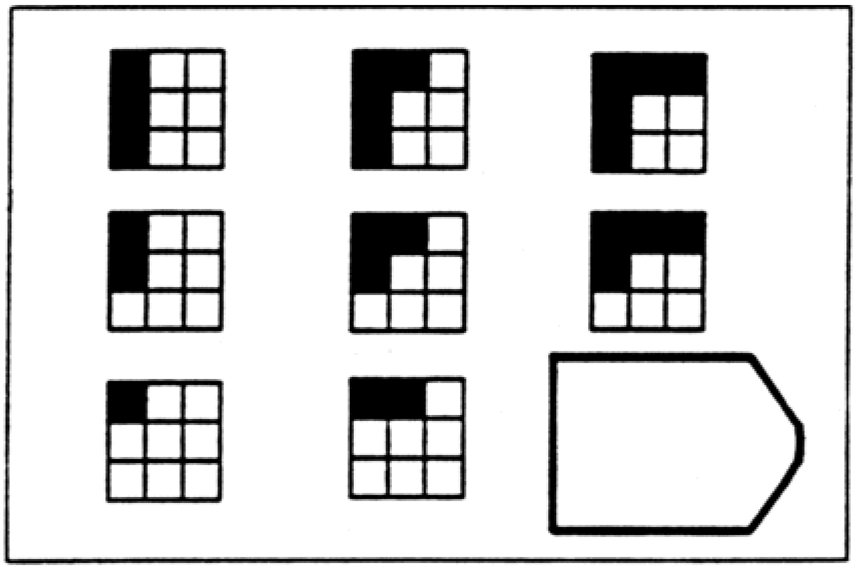
Word: Carpintero
Answer: 2

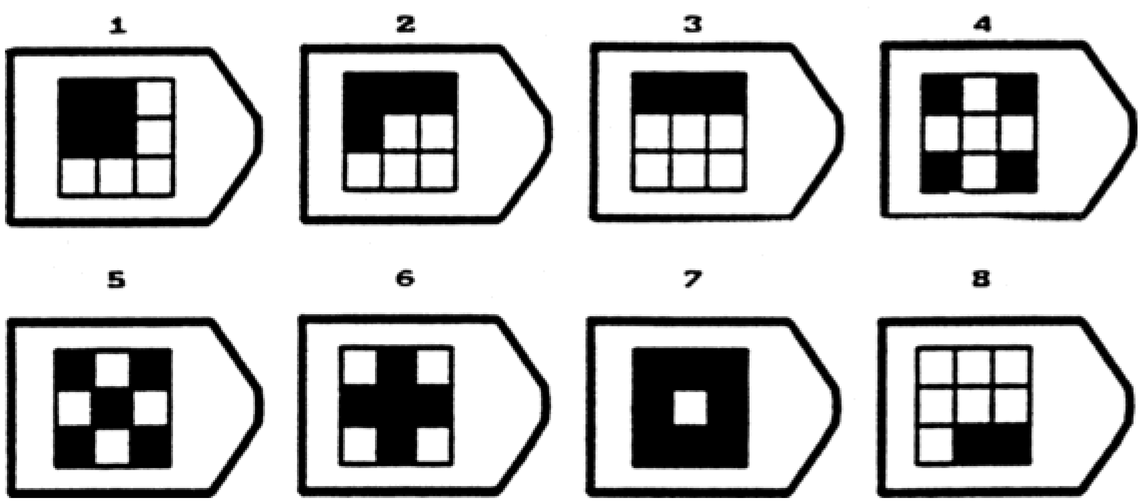


Appendix C: Raven's Advanced Progressive Matrices

Instructions: Select the piece below that best completes the pattern going across and down

Example Stimuli (correct response is 3).





Appendix D: Color-Shape Task

Instructions:

- When given the cue to identify the shape (three black squares), press the RIGHT button if the shape is a CIRCLE and the LEFT button if the shape is a TRIANGLE.
- When given the cue to identify the color (a rainbow), press the RIGHT button if the color is RED and the LEFT button if the color BLUE.

Stimuli from left to right: Blue circle, blue triangle, red circle, red triangle

Cues: shape identification (left) and color identification (right)

Appendix E: Picture-Naming Task Images and Words

Note: picture names are below each image with the English word first, followed by the Spanish word.

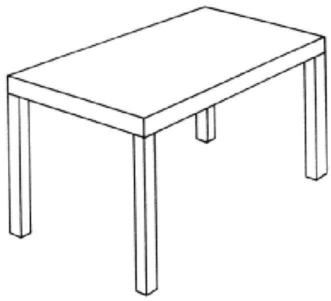
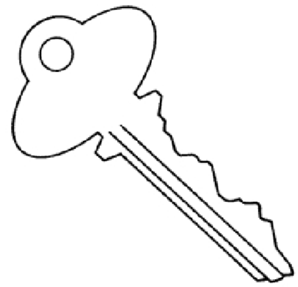


table - mesa



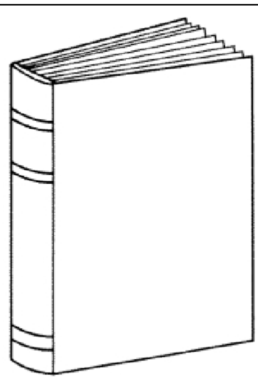
key - llave

fire - fuego

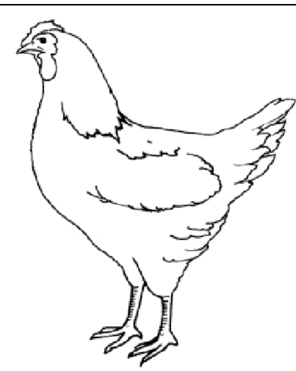
king - rey

rock - piedra

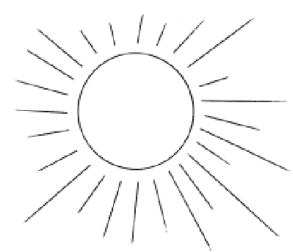
woman - mujer



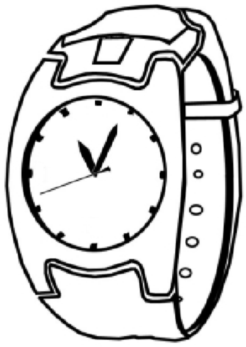
book - libro



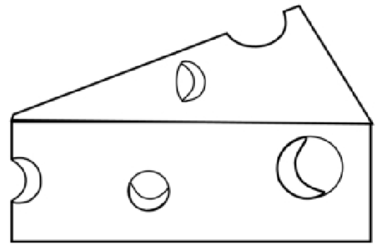
chicken - gallina



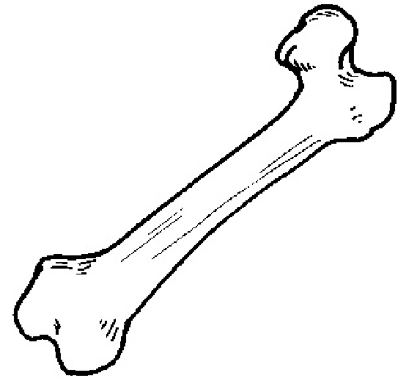
sun - sol



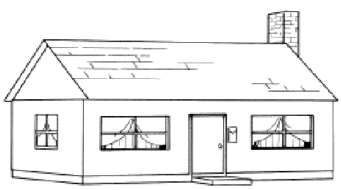
watch - reloj



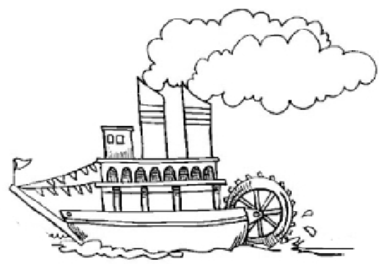
cheese - queso



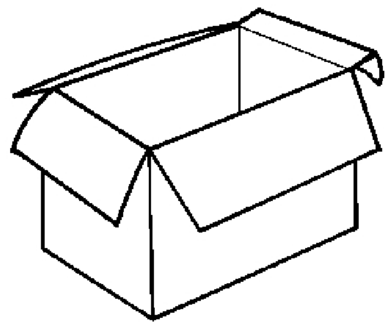
bone - hueso



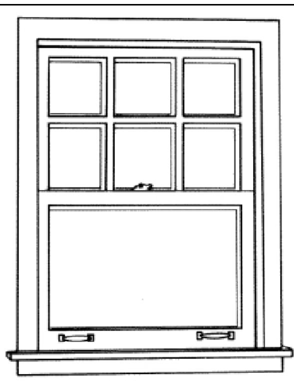
house - casa



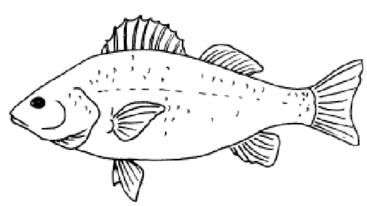
boat - barco



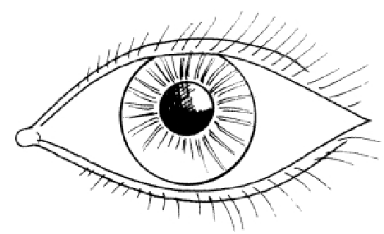
box - caja



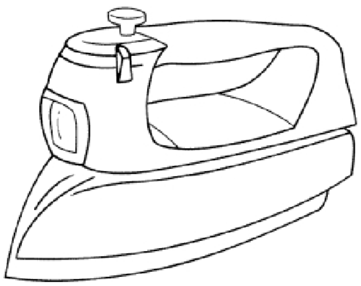
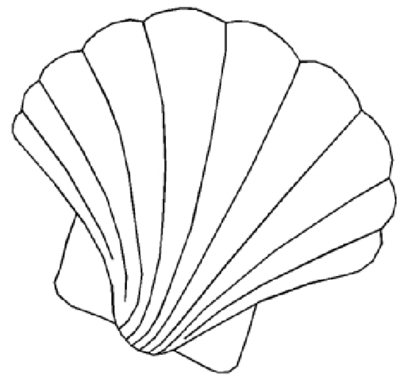
window - ventana



fish - pez

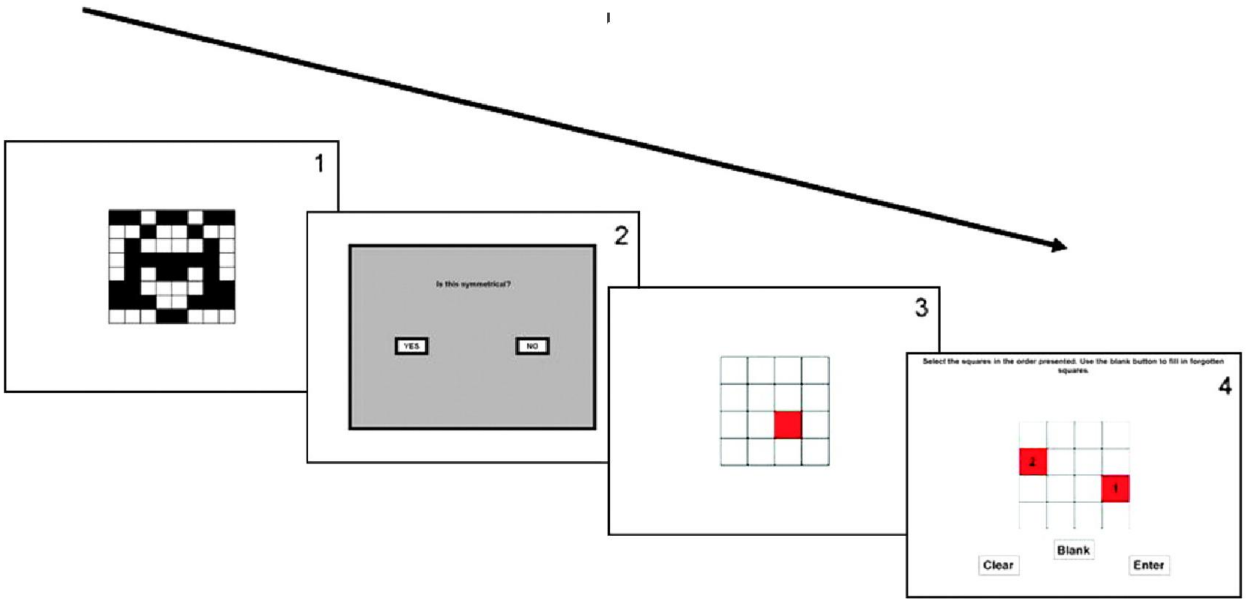


eye - ojo

		
fly - mosca	tree - árbol	hand - mano
		
iron - plancha	cloud - nube	shell - concha

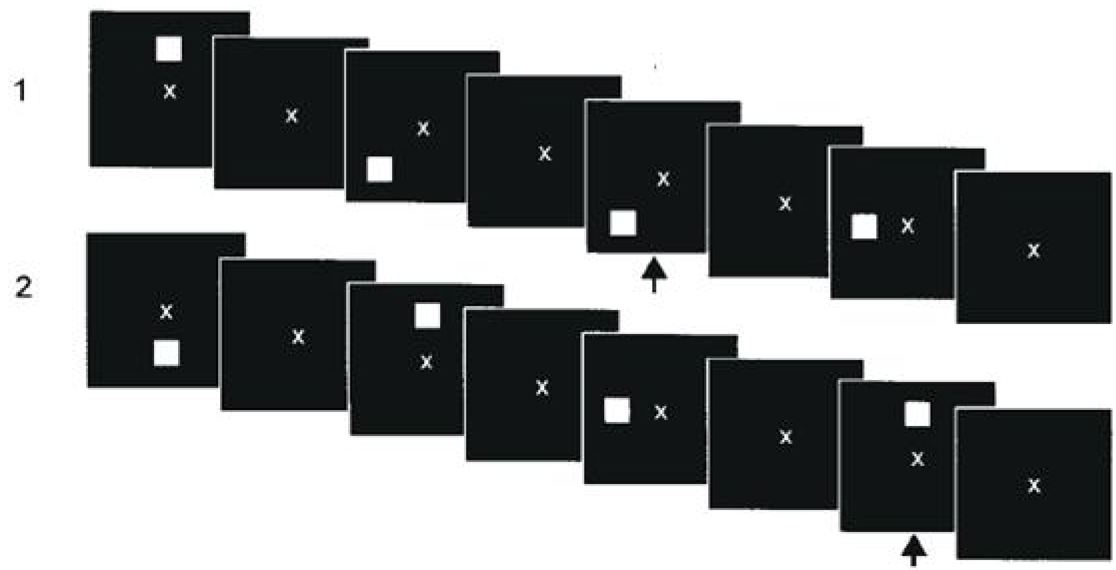
Appendix F: Automated-Symmetry Span

Example trial: Screen 1 is the to-be-judged symmetry image, Screen 2 is where participants indicate whether the image on screen was symmetrical, Screen 3 is the presentation of the to-be-remembered square, Screen 4 is the response square where the participants indicates the sequence of the squares in the set.



Appendix G: Spatial N-Back Test

Example Stimuli: The arrow on Line 1 indicates a target trial, where the location of the square on the current trial (n) matches that of n-1. The arrow on Line 2 indicates a target trial, where the location of the square on the current trial (n) matches that of n-2. Participants should respond "yes."



Appendix H: Number-Letter Task

Stimuli Sets

Numbers:

- Even: 2, 4, 6, 8 (Button 1)
- Odd: 3, 5, 7, 9 (Button 2)

Letters:

- Vowels: A, E, I, U (Button 1)
- Consonants: D, L, S, R (Button 2)

Target Stimuli (based on position)

- Top Row: Number
- Bottom Row: Letter

Instructions: Press the button that corresponds with the target response

- If the Number-Letter pair is in the TOP box, then identify the number as ODD or EVEN
- If the Number-Letter pair is in the BOTTOM box, then identify the letter as a VOWEL or CONSONANT

Example stimuli - Note: Target appears in red here for illustrative purposes only. All target and nontarget items will be presented as black to participants; correct response (CR).

Example 1 (CR = Button 1)	
L	2

Example 2 (CR = Button 1)	
I	6

Example 3 (CR = Button 2)	
A	7

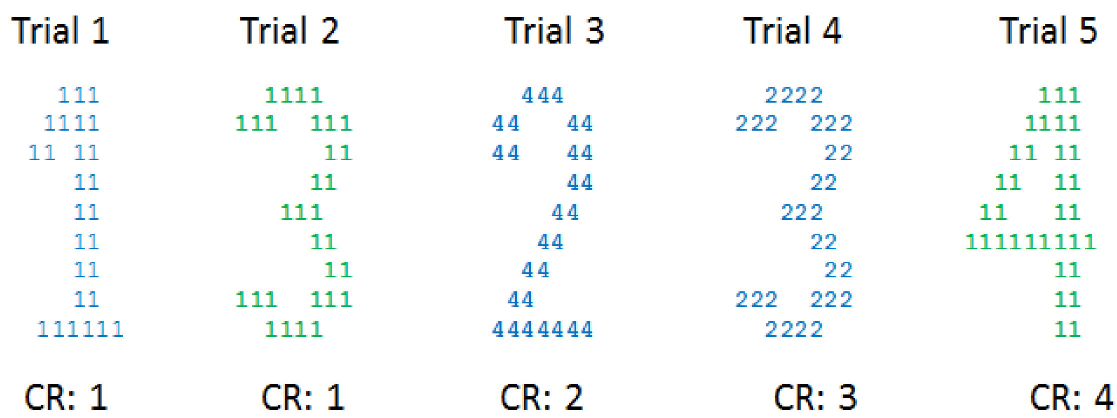
Example 4 (CR = Button 2)	
D	5

Appendix I: Local-Global Task

Instructions: Press the button (1, 2, 3, or 4) that corresponds with the target number

- If the figure is BLUE, then identify the GLOBAL number
- If the figure is GREEN, then identify the LOCAL Number

Example stimuli with correct responses (CR): from left to right -- Trial 1 (global-match), Trial 2 (local-mismatch), Trial 3 (global-mismatch), Trial 4 (global-mismatch), Trial 5 (local-mismatch).

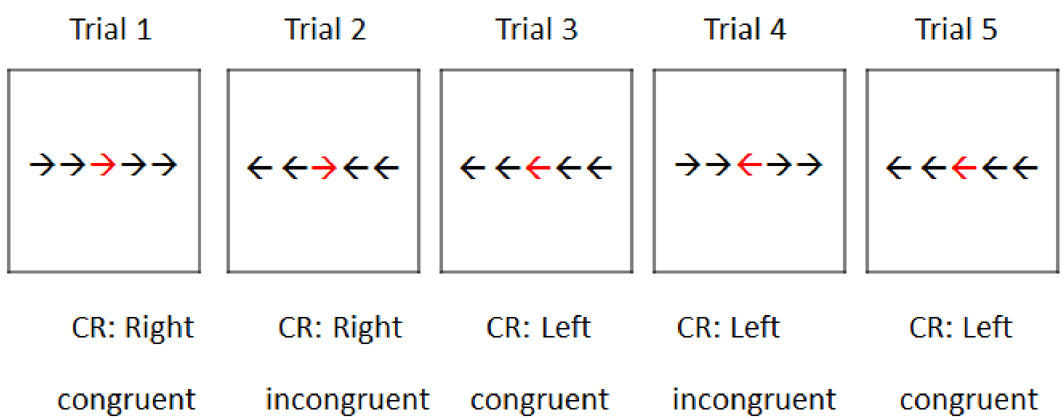


Appendix J: Flanker Task

Instructions:

- Press the button that corresponds to RIGHT arrow when the arrow in the MIDDLE is pointing RIGHT
- Press the button that corresponds to the LEFT arrow when the arrow in the MIDDLE is pointing LEFT

Example stimuli with correct response and trial type labels.

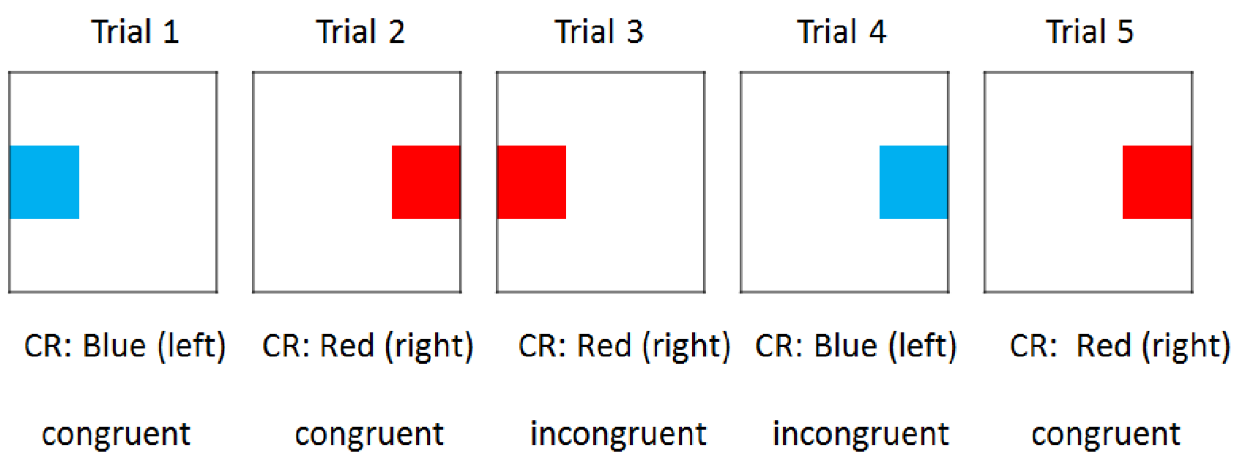


Appendix K: Simon Task

Instructions:

- Press the blue (LEFT) button when you see the BLUE square.
- Press the red (RIGHT) button when you see the RED square

Example stimuli with correct response and trial type labels.



Appendix L: Key Terms for Switch Tasks

Tasks – Nonverbal: Identify shape (circle vs. square) or identify color (red vs. blue); Verbal: Name picture in English or Spanish

- **Block types**:
    - **Simple block**: 1 task only, no switching (e.g., identify shape only; name in Spanish only)
        - **Trial type:**
            - SINGLE-TASK TRIAL (non-switch only, always the same task)
    - **Mixed block**: two tasks with cued switching (e.g., identify shape or identify color as cued)
        - **Trial types:**
            - REPETITION TRIAL (repeat task from previous trial): Trial 1: shape, Trial 2: shape, Trial 3: shape
            - SWITCH TRIAL (switch task from previous trial): Trial 1: shape, Trial 2: color, Trial 3: shape
- **Costs** (e.g., in RT on accurate trials)
    - **Switching costs** (trials *within* a mixed block): repetition trial < switch trial
        - Asymmetrical switch costs = difference in size of switch costs *between tasks*
            - Typically, more difficult to switch back to the more dominant task
        - Symmetrical switch costs = switch cost size equivalent for *both tasks*
            - Possibly evidence for no use of IC or equal amounts of IC use
    - **Mixing costs** (*non-switch trials across* blocks): single-task trial < repetition trial
        - Smaller mixing costs across participants associated with better updating of working memory ability
        - Asymmetrical mixing costs = difference in size of mixing costs *between tasks*
        - Symmetrical switch costs = size of mixing cost equivalent for *both tasks*

Curriculum Vitae

Joanna C. Bovee
boveejc@gmail.com

Education

Ph.D. 2014	University of Illinois at Chicago, Chicago, IL Concentration: Cognitive Psychology Minor: Psyscholinguistics
M.A. 2009	University of Illinois at Chicago, Chicago, IL Concentration: Cognitive Psychology
B.A. 2006	Lake Forest College, Lake Forest, IL Summa Cum Laude, Phi Beta Kappa Double Major: Psychology (with Honors) & Spanish Language and Literature

Honors and Awards

2011	Provost's Dissertation Research Award ($2518), University of Illinois at Chicago
2006, 2011	University Fellowship, University of Illinois at Chicago
2006	Sterling Price Williams Prize in Psychology, Lake Forest College
	McPherson Prize for Excellence in Scholarship – Spanish, Lake Forest College
2005	Maria Velez de Berliner Spanish Essay Award, 3rd place, Lake Forest College
2004	Member, Psi Chi, Honor Society
2004	Member, Phi Sigma Iota, Foreign Language Studies Honor Society

Publications

Raney, G. E., Campbell, S. J., & Bovee, J. C. (2014). Using Eye Movements to Evaluate the Cognitive Processes Involved in Text Comprehension. *JoVE (Journal of Visualized Experiments)*, (83).

Bovee, J., Fitz, C., Yehl, G., Parrott, S., & Kelley, M. R. (2009). Applied part-set cuing. In M. R. Kelley (Ed.), *Applied Memory* (pp 73-78). Hauppage, NY: Nova.

Bovee, J. C. (2007). Effects of transcendental meditation on blood pressure: A literature review. *Modern Psychological Studies*, *11,* 1-11.

Kelley, M. R., & Bovee, J. C. (2007). Part-Set Cuing & Order Retention. *Advances in Psychological Research, 51*, 133-148.

Manuscripts in Press

Raney, G.E. & Bovee, J.C. (in press). Reading integration in bilingual speakers. In R.R. Heredia, J. Altarriba, & A.B. Cieslicka, (Eds.) *Methods in Bilingual Reading Comprehension Research*. Springer

Bovee, J.C. & Raney, G.E. (in press). Language Proficiency Affects Word Processing, Attentional Resource Allocation, and Comprehension. *Applied Psycholinguistics.*

Manuscripts in Preparation

Bovee, J. C, Raney, G. E., & Daniel, F (in preparation). Reading ability and text comprehensibility: Missing links in understanding the letter detection task.

Morgan-Short, K. & Bovee, J.C. (in preparation). Second language instruction background and executive control in bilingual adolescents.

Raney, G.E., Daniel, F., Bovee, J. C. Lynch, F. & Vadakara, T. (in preparation) *Repeating phrases across seemingly unrelated narratives: A similarity-based explanation of text repetition effects.*

Professional Presentations

Raney, G.E., Bovee, J.C., Miller, K., Campbell, S.J., Fayz, L. & Brill-Schuetz, K.. *Phonological Similarity Effects Vary in Native and Non-Native Speakers.* Presented at the Annual Meeting of the Midwestern Psychological Association. Chicago, IL, May 2014.

Bovee, J.C., Brill, K., Raney, G. E. & Morgan-Short, K.. *Immersion in Elementary School Enhances Inhibitory Control.* Presented at the Annual Meeting of the Psychonomic Society. Toronto, ON, Canada, November 2013.

Raney, G.E., Ali, A., & Bovee, J.C. *Errors During Letter Detection: Where are They?* Presented at the Annual Meeting of the Psychonomic Society. Minneapolis MN, November 2012.

Bovee, J.C., Morgan-Short, K., Brill, K. & Raney, G. E. *Age of second language acquisition predicts enhanced executive control in bilingual adolescents.* Presented at the Annual Meeting of the Psychonomic Society. Seattle, WA, November 2011.

Lee, K., Raney, G.E., Bovee, J.C., & Daniel F. *False Cognate Detection: Parallel Activation of Native and Second Language Lexicons.* Presented at the Annual Convention of the Association for Psychological Science. Chicago, IL, May 2012.

Raney, G.E., Gold, L. & Bovee, J.C. *Reading and Listening Produce Similar Comprehension and Transfer Benefits*. Presented at the Annual Convention of the Association for Psychological Science. Chicago, IL, May 2012.

Bovee, J. C. & Raney, G. E. *Bilingual word processing across varying levels of proficiency.* Presented at the UIC Bilingual Forum. Chicago, IL, April 2011.

Bovee, J. C., Raney, G. E., & Daniel, F. *Levels of Comprehension During Letter Detection: The Roles of Reading Ability and Text Comprehensibility.* Presented at the Annual Meeting of the Psychonomic Society. St. Louis, MO, November 2010.

Bovee, J. C. & Raney, G. E. *Novice Language Learners Show a Reversed MLE During Letter Detection.* Presented at the Annual meeting of the Psychonomic Society. Chicago, IL, November 2009.

Daniel, F., Raney, G. E., & Bovee, J. C. *Titles Facilitate High-Level Processing But Not Low-Level Word Frequency Effects When Reading Vague Texts.* Presented at the Annual Meeting of the Psychonomic Society. Chicago, IL, November 2008.

Bovee, J. C., Raney, & G. E. *Non-native readers show long lasting effects of early language experience on word processing.* Presented at the Annual meeting of the Society for Text and Discourse, Memphis, TN, July 2008.

Bovee, J. C., Raney, G. E. Daniel, F. Vadakara, T. & Lynch F. *Letter Detection Impairs Higher-Level Text Comprehension.* Presented at the Annual meeting of the Midwestern Psychological Society. Chicago, IL, May 2008.

Kelley, M. R., Briggs, J. E., Chambers, J. C. Blegen, E., Koch, D., & Bovee, J. C. *Ironic Effects of Censorship: Generating Censored Lyrics Enhances Memory.* Presented at the Annual Meeting of the Psychonomic Society. Long Beach, CA, 2007.

Raney, G. E., Daniel, F., Bovee, J. C., Lynch, F., & Vadakara, T. Paper Presentation: *Transfer Across Seemingly Unrelated Narratives: A Similarity Based Explanation.* Presented at the Annual Meeting of the Psychonomic Society. Long Beach, CA. November, 2007.

Kelley, M. R., & Bovee, J. C. *Part-Set Cuing & Order Retention.* Presented at the Annual Meeting of the Psychonomic Society. Houston, TX. November, 2006.

Blumenfeld, H. K., Marian, V., Kaushanskaya, M., Rabin, A., Cone, N., & Bovee, J.C. *Relating Self-Reported and Behavioral Language Skills in Bilinguals.* Presented at the American Speech-Hearing Association Convention, San Diego, California. November, 2005.

Invited Presentations

Bovee, J.C., Brill, K., Raney, G. E. & Morgan-Short, K.. *The Long Lasting Effects of Early Language Immersion on Cognitive Control.* Presented at the Robert B. Glassman Memorial Brain, Mind, & Behavior Symposium, Lake Forest College, Lake Forest, IL, November 2013.

University Presentations

Bovee, J.C. Benefits of Bilingualism: Searching for the Source. Presented at the Cognitive Psychology Brown Bag Series. University of Illinois at Chicago. Chicago, IL. April 2013.

Bovee, J.C. *Language Proficiency Affects Word Processing, Attentional Resource Allocation, and Comprehension.* Presented at the Cognitive Psychology Brown Bag Series. University of Illinois at Chicago. Chicago, IL. November, 2009

Bovee, J.C. *Letter Detection Impairs Higher-Level Comprehension*. Presented at the Cognitive Psychology Brown Bag Series. University of Illinois at Chicago. Chicago, IL. April, 2007

Research Experience

The Role of Executive Control Processes in Bilingual Semantic Access (Dissertation in progress)

Committee: Gary Raney, Kara Morgan-Short, Viorica Marian, Jim Pelligrino, and Susan Goldman.
The purpose of this project is to identify which executive control processes are used when bilinguals attempt to retrieve words from their mental lexicons. The influence of cognitive individual differences and language proficiency are also explored.

Second Language Acquisition in the Oak Park/River Forest School District (Fall 2009 – present)

Supervisor: Kara Morgan-Short, Ph.D., University of Illinois at Chicago.
The purpose of this project is to explore how early second language acquisition experience may lead to long-term effects on language proficiency and more general cognitive skills. My responsibilities include experimental design, stimuli development, participant recruitment, and collecting, analyzing, and interpreting the data.

The Influence of Language Proficiency on Word Processing Ability and the Allocation of Attentional Resources (August 2006 - present)
The purpose of this project is to explore how language proficiency interacts with a reader's ability to process words and allocation attentional resources. My responsibilities include designing the experiments, developing stimuli, collecting, analyzing, and interpreting that data.

Applied Part-Set Cuing (Researcher, August 2007 – Fall 2009).
Supervisor: Matthew Kelley, Ph.D., Lake Forest College.
The project investigates how the memory-related phenomenon of part-set cuing inhibition, which has been traditionally studied in lab settings, applies in the real world. My responsibilities include designing the experiments, developing the stimuli, analyzing the data, and writing up the findings for publication.

Eye Tracking (Research Assistant, 2007 – present)
Supervisor: Gary Raney, Ph.D., University of Illinois at Chicago.
The focus of this research is on text comprehension. Studies have ranged from the effect of false cognates during sentence processing to the effects of transfer benefits across texts. My responsibilities include: programming experiments, testing participants, training and supervising other research assistants, and cleaning and analyzing data.

Memory for Known Information in Text (Research Assistant, 2006 – present)
Supervisor: Gary Raney, Ph.D., University of Illinois at Chicago.
This project explores memory for information in a text that is known to be encoded. My responsibilities include: testing participants, coding and scoring data, supervising undergraduate research assistants, conducting and interpreting statistical analyses, and helping prepare a manuscript for publication.

Exploring Part-Set Cuing Inhibition & Facilitation in Order Memory (Senior Thesis with Distinction, August 2005 – May 2006)
Chair: Matthew R. Kelley, Ph.D., Lake Forest College
This project replicated and expanded upon previous work on part-set cuing inhibition and facilitation when testing memory for order by manipulating list length and cue types. My responsibilities included: experiment design; creating and implementing testing methods; collecting and coding data; conducting and interpreting statistical analyses; and helping prepare a poster presentation and a manuscript for publication.

Language Experience and Proficiency Questionnaire (LEAP-Q) (Research Assistant, June 2005 – November 2005)
Supervisor: Viorica Marian, Ph.D., Northwestern University, Evanston, IL.
This project was funded by the National Science Foundation's Research Experience for Undergraduates Program in Culture, Language, & Thought, Northwestern University
This project related self-reported and behavioral language skills in bilinguals. My responsibilities included: developing a Spanish grammaticality judgment task, administering standardized tests of English language abilities, such as the PPVT and the Woodcock-Johnson, and their Spanish equivalents; and assisting in the coding and analysis of data for future presentations and publications.

Relevant Skills and Training

ERP (completed 1 semester course on ERP methodology)
Eye Tracking (EyeLink 1000, SR Research Experiment Builder, & Data Viewer)
E-Prime
SuperLab
SPSS
Excel (Advanced)
Qualtics

Languages

English (Native)
Spanish (Advanced)
Italian (Beginner-Intermediate)

Teaching Experience

FALL 2008, FALL 2009, FALL 2013, SPRING 2014 - TA for PSCH 353: Laboratory in Cognition and memory, Supervisor: Gary Raney Responsibilities included posting information on Blackboard, attending lectures, creating and grading lab assignments, holding office hours, and overseeing student research projects.

SPRING 2009 - TA for PSCH 354: Knowledge Acquisition, Supervisor: Stellan Ohlsson
Responsibilities included posting grades on Blackboard, collecting and grading assignments, proctoring exams, and holding office hours.

FALL 2009 - TA for PSCH 353: Laboratory in Cognition and memory, Supervisor: Gary Raney
Responsibilities included posting information on Blackboard, attending lectures, creating and grading lab assignments, holding office hours, and overseeing student research projects.

Fall 2010 to Present - TA for SUBJECT POOL, Supervisor: Stew Shankman/Gary Raney
Responsibilities include training new researchers, lecture to PSCH 100 students on instructions and procedures, organizing mass testing, administering mass testing and assigning credit, corresponding with researchers and students, online sign-up system management.

Departmental Service

Cognitive Division TA

Academic Years 2007-2008 and 2008-2009, Supervisor: Stellan Ohlsson
Responsibilities included assisting in aspects of faculty searches that included graduate students (e.g., lunches, advertising candidate talks, and distributing candidate evaluations); coordinating graduate student applicant visiting day (arranging accommodations, meals, and activities).

Committee on Graduate Studies Student Representative (August 2010 – August 2011)

Eleceted position. Responsibilities include leading new graduate student orientation, planning social events, and serving as a liaison between faculty and graduate students within the department,

Graduate Student Council Psychology Department Representative (August 2010 – August 2011)
Elected position. Responsibilities include acting as a liaison between the graduate student council and the graduate students in the psychology department and representing the graduate students at graduate student council meetings.

Undergraduate Mentor
Responsibilities include mentoring undergraduates enrolled in PSCH 397 & 399 (Directed and Independent Research) who are working on research projects. This includes teaching students how to develop a study idea and appropriate stimuli, test participants, analyze data, and prepare the results for either poster or paper presentations. Students have gone on to win such university awards or grants such as the Nancy Hirschberg grant (sponsored by the Psychology Department), the Nancy Hirschberg research paper award, and the undergraduate research symposium at UIC.