Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Margo Menkes March 28, 2017

Self-Generation of Novel Information through Integration of Cross-Language Episodes

by

Margo W. Menkes

Patricia J. Bauer
Adviser

Department of Psychology

Patricia J. Bauer

Adviser

Alena G. Esposito

Committee Member

Donald Tuten

Committee Member

Robyn Clarke

Committee Member

2017

Self-Generation of Novel Information through Integration of Cross-Language Episodes

By

Margo W. Menkes

Patricia J. Bauer

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Department of Psychology

2017

Abstract

Self-Generation of Novel Information through Integration of Cross-Language Episodes
By Margo W. Menkes

Adults must integrate and extend upon separate episodes of learning to build a knowledge base. Through the productive process of integration, adults can self-generate novel understandings that go beyond explicitly taught information. Adults who acquire information in multiple languages encounter separate episodes across languages, creating a situation in which episodes share low similarity at the surface level. This low surface similarity of cross-language episodes may make it more difficult to recognize their relatedness, potentially hindering the integration process. The primary question of interest in the present study is whether presenting episodes across languages creates a cost to the self-generation and integration processes. The present research investigates the extent to which bilingual adults can integrate across languages as well as within them, and investigated individual differences that might explain integration performance. Young adults who self-identify as Spanish-English bilinguals were presented with cross-language fact pairs as well as within English and within Spanish fact pairs. Participants were assessed for derivation of new knowledge through integration of separate but related factual episodes. Participants self-generated novel information through integration of separate but related episodes within English and across languages. They also showed self-generation of novel information through integration of within Spanish episodes, although, at a lower success rate than within English. Spanish and English vocabulary levels explained variability in cross-language and within Spanish integration performance. Additionally, it was observed that repeated presentation of the research paradigm facilitated performance. The results provide evidence that the process of self-generation through integration does suffer in adults under more challenging testing conditions; however, presenting facts across languages does not create a significant cost to the integration process.

Self-Generation of Novel Information through Integration of Cross-Language Episodes

By

Margo W. Menkes

Patricia J. Bauer

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Department of Psychology

2017

Acknowledgements

Thank you to the Memory at Emory laboratory for their help and guidance with various steps of the research process.

Table of Contents

Introduction

Adults learn from various sources through different experiences. A major question of interest is how individuals accumulate knowledge through these experiences. To build knowledge, learners must acquire information in separate episodes and integrate them. One must integrate separate traces of knowledge based on their related content in order to build a knowledge base. What is more, to do so efficiently, one must extend beyond the explicitly taught information. A growing body of research investigates the process of integration of separate episodes and demonstrates that children and adults both productively extend knowledge through integration of separate but related factual episodes (Bauer & Jackson, 2015; Bauer & San Souci, 2010; Varga & Bauer, 2013). It has been shown that integration is more challenging when children must recognize relatedness across separate episodes that appear different at the surface level (Bauer et al., 2012). That is, it is more difficult to recognize the shared deep structure of related facts when their surface level features are dissimilar. One situation in which episodes have low surface similarity that may challenge integration is when content is presented in different languages. Although recent research has begun to investigate cross-language integration in children, it is not yet known how changing surface similarity by presenting episodes across languages impacts the integration process in adults. Therefore, the present research investigates integration of separate but related episodes across languages.

Integrating information learned in separate episodes under different circumstances is crucial to building a knowledge base. For example, one might learn in one episode that *The first vitamin found was Niacin*, and in another episode *Polish biochemist Casimir Funk discovered Niacin*. These separate traces of knowledge can be integrated and extended upon to derive the novel information, *The first vitamin to be found was Niacin*. In this way, learners can extend

beyond two facts directly acquired at different times or in different contexts and make logical inferences. This productive process allows for efficient learning as a knowledge base is built up through different experiences.

There is evidence that self-generation of new knowledge through integration is an important way adults build a knowledge base. Bauer and Jackson (2015) found that new information self-generated through integration is rapidly incorporated into the semantic knowledge base. In an encoding phase, adults read novel facts in a single sentence format on a screen one at a time. Separate but related "stem" facts could be integrated to form a novel understanding; however, participants were not told at any point that facts were related. For example, participants read separate stem facts *The world's largest biome is the boreal forest* and *The boreal forest is the Taiga*. The facts were separated in time and no mention of their relatedness was made. At test, experimenters measured neural activity in response to fact presentation using event-related potentials (ERPs; scalp recorded electrical activity associated with information processing). Fact presentation included newly generated integration facts (that could be derived through integration of two separate stem facts previously presented in encoding—in the previous example, *the world's largest biome is the Taiga)*, novel facts (that were likely previously unknown information to adults), and well-known facts (that were likely to be well-known information to adults). At the first 400ms presentation of the fact, neural responses to the newly generated facts was intermediate between that to novel and well-known. At the second 400ms presentation, neural responses to the newly generated facts no longer differed from well-known facts, and both differed from novel facts (Bauer & Jackson, 2015). This rapid transition from novel to well-known information after one 400ms long presentation

demonstrates that semantic memory can be extended through the productive process of integration.

Self-generation of new knowledge through integration contributes to the long-term accumulation of knowledge. Self-generated knowledge has been shown to remain accessible over time in both children and adults. At encoding, individuals were presented with separate but related stem facts. At test, an experimenter asked open-ended questions that prompt participants to self-generate a novel fact derived through integration of stem fact pairs. In a young adult population, newly self-generated knowledge was retained after a 1-week delay (Varga & Bauer, 2017a). Additionally, self-generation performance is related to standard measures of academic achievement. Varga, Esposito, and Bauer (2017) studied integration performance in young adults and children in relation to educational success. In adults, self-generation performance was associated with both standardized measures of scholastic aptitude (SAT scores) and academic achievement (college GPA). The relationship between this productive process and educational success demonstrate its importance for learning and academic outcomes.

The process of extending knowledge through integration of separate episodes has been extensively examined in children by presenting two separate but related novel facts each in the context of a short passage. At encoding, 4-page long illustrated passages are read aloud to children. At test, children are asked questions that assess their self-generation of new knowledge through integration of the separate learning episodes. From this body of research, we know that newly self-generated knowledge is retained after a 1-week delay, reminder cues facilitate this long-term accessibility, hints facilitate integration through self-generation, and performance is related to reading comprehension (Varga & Bauer, 2013; Varga, Stewart, & Bauer, 2016; Bauer, Varga, King, Nolen, & White, 2015; Bauer, Blue, Xu, & Esposito, 2016).

Important to the current investigation, research with children has shown that self-generation is facilitated by high surface similarity between episodes of learning (Bauer, King, Larkina, Varga, & White, 2012). Separate but related novel stem facts were presented in the context of two passages that either shared a main character, such as a ladybug (high surface similarity), or had two distinct main characters, such as a ladybug and a rabbit (low surface similarity). Integration performance suffered when related facts were presented in contexts sharing lower similarity: 37% of children successfully generated the integration fact in the low surface similarity condition compared to 63% in the high surface similarity condition. Bauer et al. (2012) proposed an explanation for this lower performance: presenting facts across contexts sharing low surface similarity makes it more difficult to recognize their relatedness. Whether manipulations of surface similarity would similarly impact integration processes in adults is not known.

In the present research, we began investigation of surface similarity effects in adults. Rather than present facts in the context of story passages, we manipulated surface similarity by presenting facts in separate languages. Presenting episodes in separate languages can similarly create low surface similarity, which may result in lower integration performance.

Testing effects of surface similarity by manipulating the language in which stem facts are presented has high ecological validity, with clear implications for learning and education. For individuals regularly exposed to multiple languages, distinct languages create different contexts in which individuals may learn. Therefore, it is of interest whether individuals can still successfully integrate information if it is acquired in separate languages. One example of adults who may experience this is international university students. Young adults who travel internationally to pursue a higher education must build on existing knowledge from home and

previous schooling—which may have been learned in a different language. The U.S. has seen steadily increasing numbers of international students over recent years, and is a global leader in educating international students (Banjong & Olson, 2016). The majority of international students in U.S. higher education are from countries where the primary language is not English (Banjong & Olson, 2016). This population must integrate new information they learn in English with information acquired in a potentially different native language. However, no previous research has addressed cross-language integration in adults. That is, we do not know whether adults integrate separate learning episodes acquired across languages in the same way that they do within just one language. The present research investigated the extent to which adults who self-identify as bilingual can integrate separate episodes of learning across different languages of instruction. Additionally, the current study also addressed how individual differences in language experience and proficiency drive differences in self-generation performance.

Presenting each of two stem facts in separate languages may make it more difficult to recognize the relatedness of facts and the opportunity to integrate. Behavioral research shows enhanced memory performance for associate pairs that were congruent (*classroom-chalk*) rather than incongruent pairs (*tennis court-soup ladle*) in terms of co-occurrence in the world (van Kesteren et al., 2013). Associative memory performance increased with subjective congruency of associative pairs. Lower proficiency bilinguals may subjectively perceive sentences across languages as semantically incongruent, posing a difficulty to recognize cross-language fact pairs as related. Indeed, explicit awareness for relatedness of facts is associated with the proportion of integration facts successfully self-generated (Varga & Bauer, 2017b). That is, those who recognized the opportunity to integrate showed more successful self-generation performance. It may be more difficult to recognize this relatedness when facts are presented across languages.

Consistent with the suggestion that cross-language episodes are more difficult to integrate, recent research investigated self-generation of new information through cross-language integration in children. Esposito and Bauer (2017a) found that children in Grades 1 and 2 participating in a Spanish-English dual-language classroom successfully self-generated new knowledge through integration of separate factual knowledge episodes learned in different languages. However, when facts were presented in a single-sentence format (without the context of a story) to 3rd and 4th graders, performance for stem facts presented across languages suffered in comparison to the English-English condition for 3rd graders. Fourth graders did not differ in performance for cross-language episodes compared to the English-English condition. That is, a cost to integration was observed only in 3rd graders when stem facts were presented across languages. One explanation for this is the difference in experience students have had in the dual-language setting. With an additional year of experience in the dual-language classroom, 4th graders may be more accustomed to acquiring information in two languages. It is expected that adults with different language experiences will also differ in cross-language integration performance. Those who are less experienced or less proficient may face difficulty in recognizing the opportunity to integrate across languages.

Individuals with lower proficiency may also suffer compared to single-language conditions because of the challenge in understanding and encoding stem facts in a language they have not mastered. Adult bilinguals show a cognate effect in sentence reading tasks, supporting a non-selective view of bilingual activation (van Assche, Duyck, Hartsuiker & Diependaele 2009). Van Assche et al. (2009) tracked Dutch-English bilinguals eye movements while reading words that are cognates in Dutch and English (*schaap-sheep*). Participants read cognate words faster than control words in a native language sentence context, suggesting their second language is

also active and influences reading in their native language. It is possible that the single sentence fact presentation format of the current study will lead to parallel activation of languages, therefore facilitating integration of cross-language facts. However, speed of cross-language activation may depend on second-language proficiency level. In bilinguals low in L2 proficiency, L2 representations are activated more slowly than L1 representations (see van Heuven & Dijsktra, 2010 for a review). Successful integration performance was related to speed of response (Varga & Bauer, 2017b). Therefore, lower proficiency bilinguals may show lower integration performance in response to cross-language fact pairs or fact pairs within their weaker language.

Although recent work has begun to study cross-language integration in school-age children, research has not yet investigated the extent to which adults can integrate across separate language episodes. The current research addresses the impact of this low surface similarity manipulation, in the mode of two separate languages, on the integration of separate but related episodes.

Present Study

The purpose of the current research was to study self-generation of new factual knowledge learned in cross-language episodes. The primary question of interest was whether integration performance across languages differs from integration performance within a language. Additionally, the current research investigated how individual differences in language experience and proficiency might relate to differences in integration performance. The current study is the first to assess cross-language integration in an adult population. Participants in the present study were college undergraduate students self-identifying as Spanish-English bilinguals, including both native Spanish speakers and advanced Spanish language students. Participants varied in relative Spanish and English proficiency, as well as in bilingual experience.

The current study adapted the method used by Esposito and Bauer (2017a) for use with adults in a laboratory setting. Their method was designed to study the integration of cross-language integration in elementary students in the dual-language classroom. In the current study, participants were presented with separate but related novel facts in a single sentence format. Each participant was presented with fact pairs across four language conditions: (1) both stem facts were presented in English, (2) both stem facts were presented in Spanish, (3) Stem 1 was presented in English and Stem 2 in Spanish, and (4) Stem 1 was presented in Spanish and Stem 2 in English. Integration performance was assessed for facts presented across languages in comparison to within languages. Self-generation through integration was assessed with open-ended questions that elicit the integration fact, or the novel understanding that can be derived by integrating the two facts. Additionally, language experience and proficiency in both Spanish and English were assessed in order to study those variables in relation to integration performance. Language experience and proficiency were measured using a self-report survey, and proficiency levels in Spanish and English were also established using a norm referenced language measure.

Hypotheses

I hypothesized that if stem facts were presented across languages, then integration performance would suffer: fact pairs presented within one's native language would yield higher integration performance than cross-language fact pairs and fact pairs within their non-native language. However, I expected variability in the extent to which this cross-language manipulation poses a challenge. I predicted individual differences in language proficiency and experience to predict integration performance: greater bilingual experience and proficiency would predict higher performance in cross-language integration.

Method

Participants

The sample consisted of 32 participants (16 women, 16 men, M_{age} = 20.33 years, *SD* = 4.51, age range: 18 to 23 years) drawn from the Emory University undergraduate population. Participants were recruited from a pool of Emory undergraduates receiving course credit for an introductory psychology course or from Spanish language courses. Eligibility for participation required that participants either be native Spanish speakers or be enrolled in a Spanish course above a specified minimum level. This minimum course level teaches reading comprehension and communicative speaking skills. Placement into this course is determined by a score of 440 or above on the WebCAPE (Computer Adaptive Placement Exam) in combination with the student's previous Spanish coursework. This qualifier was to ensure participants had sufficient Spanish reading and speaking competency to complete the study, but still allowed for a range of proficiency levels. Assuming all Emory undergraduates enrolled in English taught courses have a certain proficiency in English, those who also have Spanish proficiency are therefore considered Spanish-English bilinguals for the purposes of this study. Participants either volunteered or received course credit at the end of the hour-long session. All participants provided written informed consent prior to initiation of the study session. The Emory University institutional review board approved the study protocols and procedures.

Materials

Stimuli. Stimuli were 32 novel fact pairs consisting of 64 stem facts. 4 of these fact pairs were used by Varga and Bauer (2017a) and 10 pairs were used by Esposito & Bauer (2017b). Stem facts were determined to be novel information to Emory undergraduates based on pilot testing with 16 participants. Each fact pair included two separate but related novel facts that

could be integrated to generate a novel integration fact. For example: Stem 1 reads "*Saffron is the most expensive spice"* and Stem 2 reads "*The most expensive spice smells like honey."* These two stem facts can be integrated to generate the novel fact "*Saffron smells like honey."* However, the participant was never presented with any integration facts. It was demonstrated in pilot testing that both stem facts were necessary for self-generation of the integration facts. That is, Emory undergraduates did not self-generate the integration fact when presented with only one of the two stem facts. 50 potential fact pairs were narrowed down to 32 that were used in the present study.

Stimuli were formatted for presentation on a monitor. Each fact was presented in a single sentence, ranging from 3-15 words in length. Stimuli included a textual presentation of facts on the screen accompanied by a recording of the fact being read aloud. Text was white size 44 Times New Roman font on a gray background with a black border. For all facts in a given language, recordings were created by a single native speaker. That is, in the cross-language conditions participants heard two speaker's voices, whereas in within language conditions participants heard only one speaker's voice. This decision was based on the observation that child integration performance did not differ whether one or two speakers read facts within English (Esposito & Bauer, 2017a).

Stimuli also included an engagement question after each stem fact. These questions asked what the previous fact was about: for example, "*Was this fact about plants or animals?"* These questions served to encourage the participant to pay attention. Engagement questions were asked in the language of the fact they followed (i.e., if the fact was in Spanish, the engagement question also was in Spanish).

Measures. Language measures included the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian, Blumenfeld, & Kaushanskaya, 2007) and The Woodcock-Muñoz Language Survey®-Revised Normative Update (*WMLS®-R NU*). Both the LEAP-Q and *WMLS®-R NU* were used in order to capture language proficiency both subjectively and objectively with a standardized measure.

The LEAP-Q is a self-report measure appropriate for adults assessing individual language status. Domains the LEAP-Q assesses include acquisition history, present language use, language preference, and proficiency ratings across four domains of language use: speaking, understanding, reading, and writing). An example item reads, "Please list all the languages you know in order of acquisition." Factor analyses suggested items were internally valid (Marian et al., 2007). Multiple regression and correlation analyses established criterion-based validity and demonstrate the measure as a reliable indicator of language performance (Marian et al., 2007).

The *WMLS®-R NU* is a norm-referenced measure that establishes language proficiency level in both Spanish and English. It includes areas of reading, writing, listening, and comprehension and is appropriate for ages 2 to 90+. The experimenter administered both Test 1, vocabulary, and Test 2, analogies, in both Spanish and English. These tests measure verbal intelligence and verbal reasoning. Raw scores in each language for both Test 1 and Test 2 were recorded. For analyses with total vocabulary level in English and Spanish, score on Test 1 and Test 2 were added together. Thus, variables of English vocabulary level and Spanish vocabulary level were created for each participant.

Buffer Tasks. These tasks included the Trail-Making Test, Digit Span Reverse, and Digit Cancellation, respectively. The first and last were paper and pencil tasks, and the second task was a verbal memory recall task. For the purposes of the current study, these activities were only

included as filler tasks and to allow the participant a break from the screen. Performance on these tasks was not analyzed.

Procedure

Participants were tested individually in a laboratory room. All sessions were conducted by a single female experimenter; her Spanish proficiency was ensured by her extensive coursework in Spanish and study abroad experience. The hour-long session included (1) verification of completion of the LEAP-Q, (2) exposure to stimuli and test for integration, and (3) administration of the Woodcock-Muñoz Language Survey-Revised Normative Update. The experimenter first verified completion of the LEAP-Q. If the LEAP-Q was not completed, the participant completed the LEAP-Q then. Participants then completed 4 blocks of stimuli, one in each language condition, with buffer tasks between each block (Trail Making Test, Digit Span Reverse, and Digit Cancellation). Each block consisted of two phases: (1) encoding and (2) test. In the encoding phase, participants read novel facts on an eye tracking monitor while hearing a pre-recorded audio reading of the fact. In the test phase, participants were given a paper with all 8 integration questions related to the fact pairs of the current block (the use of the eye-tracking equipment was incidental to the present research. Analysis of the data is beyond the scope of the present investigation).

32 fact pairs were presented across 4 different blocks. Each block contained 16 stem facts, yielding 8 integration facts. Participants were presented with each block in a distinct language condition: (1) both stem facts were presented in English (hereafter E/E), (2) both stem facts were presented in Spanish (S/S), (3) Stem 1 was presented in English and stem 2 in Spanish (E/S), and (4) Stem 1 was presented in Spanish and Stem 2 in English (S/E). Participants were presented with these four language conditions in 1 of 4 different orders. Stimuli were

counterbalanced so that across participants, each fact was presented in each language condition an approximately equal number of times. Within a single block, both stem facts of a fact pair were presented. Stem 1 of all 8 fact pairs was displayed first, then followed immediately by each Stem 2 in a consistent order. This structure served to minimize language switching to one time within cross language blocks.

Participants were never explicitly informed that the facts are related in any way, nor were they ever presented with any integration facts. Before each encoding phase, participants were told they would be asked questions about the facts at the end. To further encourage attention to the task, after each stem fact participants saw and responded aloud to an engagement question. The experimenter recorded whether engagement questions were answered correctly or not.

At test, participants were first given all integration questions in an open-ended format to prompt self-generation. To continue with the previous example: the related integration question was *What does Saffron smell like?* Participants wrote down answers. Once they indicated they had finished, the experimenter instructed them to flip the sheet over. On the other side of the paper, the same 8 integration questions appeared in a forced-choice format. Each forced-choice question had three choices, only one of which was the correct answer. Alternative choices included words that the participant was presented with in other, unrelated facts, but were on topic for the given question. That is, the correct answer could not be recognized as the only choice the participant had previously seen, nor as the only plausible answer. Rather, questions were designed to require integration of both stem facts. For example, *What does saffron smell like? Coffee, freshwater, or honey?* Open-ended and forced-choice integration questions were written in the language of the second stem fact presented. That is, for the block in which Stem 1 was in English and Stem 2 was in Spanish, integration questions were written in Spanish. This decision

was based on the finding that by 8 years of age, children begin to integrate upon presentation of the second stem fact (Bauer, Blue, Xu, & Esposito, 2016).

Following Block 4, the experimenter administered the Woodcock-Muñoz Language Survey-Revised Normative Update. All participants first were tested in Spanish, and then in English. This order of tests was to ensure that all participants finished their session with a positive experience, even if they struggled with the Spanish test due to lower proficiency level.

Scoring

Participants received a self-generation score for each language condition based on the number of open-ended integration questions they answered correctly. Participants also received a total integration score for each language condition that included correct answers in open-ended and forced-choice formats. Vocabulary level was scored using the standard instructions of the Woodcock-Muñoz Language Survey-Revised Normative Update.

Creation of Language Experience Groups

Participants were categorized into 2 groups for some analyses: native English monolinguals and native bilinguals/native Spanish speakers. The first group was defined by exposure to only English prior to school age. That is, these individuals did not begin acquiring Spanish until age 6 or later. The second group was defined by exposure to Spanish (Spanish only or Spanish and English) prior to school age. That is, individuals in this group began acquiring Spanish, either in advance of or simultaneous to acquisition of English—making them early Spanish Speakers. Hereafter these groups will be referred to as *Late Spanish speaker* and *Early Spanish speaker,* respectively. All participants in these two groups were exposed from birth to either English, Spanish, or both. Five participants who had no exposure to English or Spanish

from birth were excluded from analyses comparing these two groups. See Table 1 for a list of ages at which each participant began acquiring English and Spanish.

There are many possible variables to consider in categorizing individuals on bilingual status, including: age of acquisition, age of fluency, language dominance, or exposure to each language. Due to the nature of the Emory student population, however, these controlling variables mask much of the variability in the sample's language experience. Categorizing participants based on age of Spanish acquisition allowed this variability to be more visible and created two nearly-balanced groups to compare (15 early Spanish speakers, 12 late Spanish speakers).

Results

The results are reported in four sections. First, we report performance on self-generation through integration of separate but related learning episodes generally and across language conditions. We then examined differences between language conditions. We then analyzed factors that might drive differences in performance between blocks, as well as individual differences in performance. We examined how differences in language experience, English and Spanish vocabulary level, and order of blocks predicted performance on self-generation through integration. All analyses were conducted using SPSS Statistics package (Version 23). All statistical tests reported as significant were below an alpha level of 0.05.

Self-Generation through Integration Across Language Conditions

In the forced-choice format, participants selected 94% of correct integration facts on average (*M proportion* = 0.94, *SD* = 0.06), with performance ranging from 81 % to 100% correct. Because forced-choice performance showed ceiling effects and there was little variability to analyze, it will not be discussed further.

In contrast with forced-choice performance, self-generation performance showed more variability. Participants self-generated the novel integration facts for 64% of fact pairs on average (*M proportion* = 0.64, *SD* = 0.18). That is, participants answered correctly 64% of open-ended questions they were asked, across all language conditions. Individual self-generation performance ranged from 25% to 94% correct across the sample.

Self-Generation though Integration as a Function of Language Condition

Following general performance across language conditions, self-generation performance (in response to open-ended questions) within each language condition was examined. Mean and standard deviation of performance in each language condition across the sample is reported in Table 2, and mean and standard deviation of performance by language experience group is reported in Table 3. These analyses excluded the 5 participants who did not fall into one of the two language groups.

To examine differences in performance across language conditions, we conducted a mixed factor repeated measures ANOVA with language experience (Early Spanish speaker, Late Spanish speaker) as the between-subjects factor and language condition (E/E, S/E, E/S, S/S) as the within-subject factor. The dependent measure was the self-generation score in each block of trials, calculated based on the number correct out of 8 open-ended questions. The ANOVA revealed there was no main effect of language experience on self-generation score. There was a main effect of language condition on performance, $F(3, 75) = 4.50$, $p = 0.006$, partial $\eta^2 = 0.15$. That is, there was a significant difference in mean performances between language conditions. However, there was not an interaction between language condition and language experience.

Post hoc comparisons examining the differences in language conditions using a Tukey Least Significant Differences (LSD) test indicated only 1 pairwise comparison was significant:

E/E performance was significantly higher than S/S performance (SE = 0.32, $p < 0.001$). That is, within English integration performance was higher than within Spanish across the sample. No other pair of language conditions showed significantly different performance, $p > 0.05$ for all other pairwise comparisons.

Individual Differences in Self-Generation Performance

We conducted further analyses to examine different variables that may have influenced the observed differences in performance across language conditions and between individuals.

Vocabulary. We tested regression models examining English and Spanish vocabulary levels as predictors of self-generation score in each language condition. Results are reported in Table 4. We conducted separate analyses predicting performance in each language condition. Additionally, we tested a model with cross-language performance collapsed. Model 1 predicting E/E performance was not significant. That is, neither English nor Spanish vocabulary predicted performance in the E/E condition. Model 2 predicting S/E performance was significant, with English vocabulary as a significant predictor. Model 3 predicting E/S performance was not significant. Performance in S/E and in E/S were not significantly different from each other, $t(31) = -0.29$, $p = 0.77$, allowing us to collapse the two scores into cross-language accuracy. Given the small sample size, collapsing cross-language performance allowed us to increase power for this analysis. Model 4 predicting cross-language accuracy was significant; however, neither English nor Spanish vocabulary individually were significant predictors of performance. Lastly, Model 5 predicting S/S performance was significant, with only Spanish vocabulary as a significant predictor.

In summary, Spanish vocabulary level only predicted S/S performance, and English vocabulary level only predicted S/E performance. English vocabulary predicted performance

when Stem 2 and the integration question were both in English for the S/E language condition, however, not for the E/E language condition.

Order of Stimuli. Next, we examined effects of the order of stimuli presentation on self-generation performance. See Table 5 for a summary of performance in each presentation block. Language condition was counter-balanced so that each one was presented an equal number of times in each block number position. Therefore, differences in performance across block numbers cannot be explained by language condition. These analyses were conducted based on anecdotal evidence that participants began to show awareness of the paradigm (that stem facts were related and questions required integration of two facts) with its repeated presentation. We conducted a repeated measures ANOVA with block number (Block 1, Block 2, Block 3, Block 4) as a within-subjects factor. The dependent measure was self-generation score, calculated based on the number correct out of 8 open-ended questions. The ANOVA revealed a significant effect of block number on performance, $F(3, 93) = 5.43$, $p = 0.002$, partial $\eta^2 = 0.15$. Post hoc comparisons using a Tukey LSD test indicated significantly lower performance in Block 1 than in Block 2 ($SE = 0.40$, $p = 0.04$), Block 3 ($SE = 0.43$, $p = 0.003$), and Block 4 ($SE = 0.36$, $p = 0.002$). That is, self-generation performance in the first block a participant was presented with was significantly lower than any subsequent block. All other pairwise comparisons were not significant ($p > 0.05$).

Discussion

The purpose of the present study was to investigate integration of separate but related episodes of learning in bilingual adults. It was observed that bilingual adults successfully self-generated novel information through integration of separate but related episodes, both within English and across Spanish and English. They also self-generated novel information through

integration of separate episodes within Spanish, though at significantly lower success rate than within English.

Across the sample, self-generation performance was significantly higher in the E/E condition compared to the S/S condition, with cross-language performance falling in between the two. That is, integration performance within English was nominally more successful than both cross-language conditions, across the entire sample. These differences in performance across language conditions appear to be driven by individual differences in language experience and proficiency. Although significant differences were not revealed, language experience groups did show nominal differences in performance. Regression analyses revealed Spanish and English vocabulary levels can explain within Spanish and cross-language self-generation performance. Additionally, order of stimuli presentation had an effect on self-generation performance across blocks.

As in previous adult research, in the present research, integration performance suffered under more challenging testing conditions (Varga & Bauer, 2017b). In the present study, these more challenging testing conditions took the form of presenting facts within a language other than the dominant language (the dominant language of the sample being English). Trends in the data suggest presenting information across languages also challenges integration nominally. However, further research is necessary to investigate the cost of cross-language integration. Overall, the present study is consistent with the child cross-language integration research (Esposito & Bauer, 2017a) in that adults as well as children successfully integrate across languages and within languages. The present research suggests that low surface similarity between episodes does not challenge integration for adults in the way that it does children (Bauer

et al., 2012). That is, low surface similarity of cross-language episodes still yielded successful integration in adults and only nominally lower performance than within English.

It was hypothesized that integration performance within one's native language would be higher than across languages or within one's non-native language. Analyses comparing groups based on native language and non-native language were not conducted because these variables did not yield equal or approximately equal groups. However, analyses were conducted comparing groups that were created based on early exposure to Spanish. The conceptually equivalent prediction based on the language experience groups would be that early Spanish speakers would outperform late Spanish speakers integrating within Spanish, and that late Spanish speakers would perform highest integrating within English. This hypothesis was partially supported: across the sample, integration performance was highest within English, integration within Spanish yielded the lowest performance, and cross-language integration performance fell in between. However, the only statistically significant difference was between E/E and S/S performance. As was predicted, native English speakers performed highest when integrating within their native language. Interestingly however, even early Spanish speakers followed the pattern of performing their highest when integrating within English and poorest within Spanish. That is, presentation of stimuli within Spanish created a cost to integration for the entire sample. A significant difference in performance was not observed across language experience groups. In line with the hypotheses, an examination of mean performance reveals that early Spanish speakers outperformed late Spanish speakers integrating within Spanish and across languages when the second stem fact and integration question were presented in Spanish. Moreover, those exposed only to English before age 6 outperformed early Spanish speakers

when the second stem fact and integration question was in English. However, these differences were nominal and failed to reach statistical significance.

Regarding the cost of presenting episodes across languages, the findings again followed the predicted trend but failed to reach statistical significance. Cross-language integration performance, on average, was lower than within English integration performance. Esposito and Bauer (2017a) proposed that experience level in the dual-language setting may explain why 3rd graders showed a cost to integration when episodes were presented across languages while 4th graders did not. In the present study, language experience (as operationalized by Spanish exposure prior to school age or lack thereof) cannot explain differences in performance—language experience groups did not differ in self-generation performance significantly. The high level of experience required for participation in the study may be a reason for the success observed in cross-language integration across the sample. Moreover, it is possible the nature of the sample being English dominant did not allow for observation of differences between language experience groups, as discussed further in Limitations.

It was also hypothesized that individual language differences in language proficiency would drive differences in performance. This hypothesis was also partially supported. Spanish vocabulary level was a predictor of within Spanish integration performance, however; English vocabulary level did not predict within English integration performance as in previous research (Esposito & Bauer, 2017a). This may be due to the E/E condition yielding high performance and showing little variability in the current study. English vocabulary level predicted cross-language integration performance when both the second stem fact and integration question were presented in English. Additionally, both Spanish and English vocabulary level predicted cross-language accuracy.

The current study is the first to examine cross-language integration of separate but related episodes in adults, as well as the first study of integration within a language other than English. The present research adds to the literature by demonstrating that self-identifying bilingual adults of a certain proficiency level in both English and Spanish can derive novel understandings through integration of separate but related episodes across languages. Additionally, it adds to the literature on self-generation of novel information through integration by demonstrating task improvement from the first presentation block to later ones. It is the first study of its kind to show that repeated presentation of this research paradigm can facilitate performance. It is likely that over multiple trials, adults began to recognize they were being presented with two separate but related facts that could be integrated together. This study suggests the importance of considering this phenomenon when designing future studies investigating integration of separate but related episodes, particularly with adults. Future research might employ a between-subjects design rather than within-subjects design to avoid practice effects.

The findings of the current study have implications for adults acquiring information in multiple languages. Previous research has established the relevance of the integration and self-generation processes for the long-term accumulation of knowledge in adults and its relation to academic outcomes (Bauer & Jackson, 2015; Varga & Bauer, 2017b; Varga, Esposito & Bauer, 2017). The present study demonstrates that international students and other students learning in multiple languages can derive novel understandings through integration of separate episodes across their two languages as well as within their dominant language. The present study additionally provides evidence that acquiring separate traces of information across languages does not necessarily hinder derivation of new knowledge; it suggests students whose native

language is not English can successfully integrate information acquired in their first language with information acquired in English-taught university settings.

Limitations

Ceiling effects in forced-choice performance across language conditions posed limitations; the low variability in performance hindered the present study from examining performance across language conditions and how individual factors might explain these differences. Within English open-ended performance was also high and showed low variability, limiting the study's ability to demonstrate how English vocabulary might predict within English integration performance. In future research, stimuli for adults should be designed to be difficult enough to see sufficient variability. Additionally, low variability in English vocabulary level (M = 74.03, SD = 3.18) compared to Spanish vocabulary level (M = 72.34, SD = 5.15) across the sample may have limited the predictive utility of English vocabulary.

Participants included in the study were limited to adults at a specified proficiency level of Spanish based on their course level or native language status. Therefore, the results likely do not extend to adults at a lower level of proficiency in one of their languages. As discussed previously, the nature of the sample posed additional limitations. At a university where the language of instruction is English, all participants were largely exposed to English. Even if one grew up as a native Spanish speaker, they were in an English dominant environment at the time of testing. Although 15 of 32 participants were grouped as early Spanish speakers, this included anyone exposed to Spanish before school age in order to create approximately equal groups. Only 6 participants identified themselves as native Spanish speakers and 3 identified themselves as Spanish dominant. This sample limited the potential to see differences in performance between dominant English speakers and dominant Spanish speakers. In order to examine

effects of native language on self-generation through integration performance more closely, the study should be repeated with a sample that is currently exposed largely to Spanish or is Spanish dominant. In such a sample, participants likely would perform highest integrating within Spanish and lowest integrating within English.

Conclusion

In conclusion, bilingual adults can integrate and extend upon facts presented across languages as well as within their dominant language. Self-generation through integration performance is higher within one's dominant language or the language they are primarily exposed to than in their less dominant language. The present research provides insight into how bilingual adults accumulate knowledge and can build a knowledge base with information acquired in each language.

References

Bauer, P. J., Blue, S. N., Xu, A., & Esposito, A. G. (2016). Productive Extension of Semantic Memory in School-Aged Children: Relations with Reading Comprehension and Deployment of Cognitive Resources. *Developmental Psychology, 52,* 1024-1037.

Bauer, P. J., & Jackson, F. L. (2015). Semantic elaboration: ERPs reveal rapid transition from novel to known. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(1), 271.

Bauer, P. J., King, J. E., Larkina, M., Varga, N. L., & White, E. A. (2012). Characters and clues: Factors affecting children's extension of knowledge through integration of separate episodes. *Journal of Experimental Child Psychology*, *111*(4), 681–694.

Bauer, P. J., & San Souci, P. S. (2010). Going Beyond the Facts: Young Children Extend Knowledge by Integrating Episodes. *Journal of Experimental Child Psychology*, *107*(4), 452–465.

Bauer, P. J., Varga, N. L., King, J. E., Nolen, A. M., & White, E. A. (2015). Semantic elaboration through integration: Hints both facilitate and inform the process. *Journal of Cognition and Development*, *16*(2), 351-369.

Esposito, A. G., & Bauer, P. J. (2017a). Building a knowledge base across languages: Self-generating new factual knowledge in a dual-language classroom. Manuscript in preparation.

Esposito, A. G., & Bauer, P. J. (2017b). Component cognitive skills supporting self-generation through integration. Manuscript in preparation.

Marian, 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 (4), 940-967.

van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological Science*, *20*(8), 923-927.

van Heuven, W. J., & Dijkstra, T. (2010). Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models. *Brain Research Reviews*, *64*(1), 104-122.

van Kesteren, M. T., Beul, S. F., Takashima, A., Henson, R. N., Ruiter, D. J., & Fernández, G. (2013). Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: from congruent to incongruent. *Neuropsychologia*, *51*(12), 2352-2359.

Varga, N. L., & Bauer, P. J. (2013). Effects of delays on 6-year-old children's self-generation and retention of knowledge through integration. *Journal of Experimental Child Psychology*, *115*(2), 326–341.

Varga, N. L., & Bauer, P. J. (2017a). Using ERPs to dissociate the neurocognitive processes underlying knowledge extension through memory integration. Manuscript in revision.

Varga, N. L., & Bauer, P. J. (2017b). Young adults self-generate and retain new factual knowledge through memory integration. Manuscript in revision.

Varga, N. L., Esposito, A. G., & Bauer, P. J. (2017). Knowledge Extension through Memory Integration: An Explanation of Variability in an Educationally Relevant Phenomenon. Manuscript in preparation.

Varga, N. L., Stewart, R. A., & Bauer, P. J. (2016). Integrating across episodes:

Investigating the long-term accessibility of self-derived knowledge in 4-year-old children. *Journal of Experimental Child Psychology, 145*, 48-63.

Table 1.

Summary of Language Experience for All Participants

Participant Number	*Age of Acquisition: English*	*Age of Acquisition: Spanish*
01	0	5
02	6	5
03	0	7
04	5	15
05	0	11
06	4	0
07	0	13
08	0	6
09	0	5
10	0	6
11	0	4
12	0	12
13	12	0
14	0	12
15	0	5
16	0	7
17	0	4
18	10	10
19	0	4
20	0	16
21	0	0
22	0	0
23	2	7
24	4	0
25	4	13
26	8	7
27	0	2
28	5	0
29	0	5
30	2	2
31	0	12
32	0	18

Table 2.

Forced-Choice and Open-Ended Integration Performance in Each Language Condition

Language Condition	*Mean Forced-Choice*	*SD Forced-Choice*	*Mean Open-Ended*	*SD Open-Ended*
E/E	7.85	0.36	6.07	1.66
S/E	7.52	0.80	5.30	1.96
E/S	7.67	0.62	5.41	2.08
S/S	7.30	0.95	4.52	1.81

Note: The descriptive statistics are based on $N = 27$ (12 late Spanish speakers, 15 early Spanish speakers).

Table 3.

Self-generation through Integration Performance in Each Language Condition, by Language Experience Group

Language Condition	*Language Experience*	*Mean Performance*	*SD Performance*
E/E	Late Spanish	6.08	1.73
	Early Spanish	6.07	1.67
S/E	Late Spanish	5.50	1.57
	Early Spanish	5.13	2.26
E/S	Late Spanish	5.00	2.73
	Early Spanish	5.73	1.39
S/S	Late Spanish	4.25	1.71
	Early Spanish	4.73	1.91

Note: The descriptive statistics above are based on $N = 27$ (12 late Spanish speakers, 15 early Spanish speakers).

Table 4.

Regression Models

Model	*Predictor*	*β*	R^2
1: E/E			
	English Vocab	0.33	
	Spanish Vocab	0.09	
			0.12
2: S/E			
	English Vocab	0.37*	
	Spanish Vocab	0.25	
			0.22*
3: E/S			
	English Vocab	0.11	
	Spanish Vocab	0.26	
			0.08
4: Cross-language			
	English Vocab	0.29	
	Spanish Vocab	0.32	
			0.20*
5: S/S			
	English Vocab	-0.11	
	Spanish Vocab	0.47*	
			0.22*

Note: * < .05. All Coefficients listed are standardized (*β*). The descriptive statistics above are based on *N* = 32 (12 late Spanish speakers, 15 early Spanish speakers, 5 others).

Table 5.

Self-Generation through Integration Performance in Each Presentation Block

Block	*Mean Performance*	*SD Performance*
1st	4.25	2.02
2nd	5.09	1.99
3rd	5.66	1.83
4th	5.50	1.90

Note: The descriptive statistics above are based on $N = 32$ (12 late Spanish speakers, 15 early Spanish speakers, 5 others).