Distribution Agreement

Signature:

______________________ ____________
Nicole Leigh Varga Date

Constructing a Knowledge Base through Memory Integration: Cognitive and Neural Factors Involved

By

Nicole L. Varga
Doctor of Philosophy

Psychology

Patricia J. Bauer, Ph.D.
Advisor

Jocelyne Bachevalier, Ph.D.
Committee Member

Robyn Fivush, Ph.D.
Committee Member

Joseph Manns, Ph.D.
Committee Member

Phillip Wolff, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Constructing a Knowledge Base through Memory Integration: Cognitive and Neural Factors Involved

By

Nicole L. Varga

B.S. (Distinguished Honors), Ursinus College, 2010

M.A., Emory University, 2012

Advisor: Patricia J. Bauer, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Psychology
2016

Abstract

Constructing a Knowledge Base through Memory Integration: Cognitive and Neural Factors Involved

By Nicole L. Varga

The construction of a knowledge base fundamentally relies on memory integration—the combination of information acquired within or across separate learning episodes. Without the ability to integrate information learned at different times and in different places, building a domain of knowledge would not be possible. We also exhibit the striking capacity to extend far beyond what has been directly learned in order to generate new thoughts, ideas, and understandings never directly specified. Self-generative learning through memory integration is pervasive in human cognition; acts ranging from basic creativity to the derivation of scientific theories depend upon it. Though it is widely assumed that we all engage in this form of knowledge extension and that we do so regularly, the question of how information learned in separate episodes becomes integrated in memory to form new knowledge is far from understood. In the present research, behavioral, event-related potential (ERP), and academic measures were used to address: (1) the extent to which college-aged students successfully self-generate and retain knowledge newly derived through memory integration; (2) the distinct neurocognitive processes underlying this behavior, and (3) how variability in self-generation through integration contributes to real-world academic outcomes. Study 1 provided the first empirical demonstration that knowledge extension through memory integration supports the long-term accumulation of integrated knowledge in adults. Moreover, substantial individual differences were observed, which were linked to whether individuals spontaneously recognized the opportunity to integrate. ERP measures in Study 2 extended the behavioral results of Study 1 and showed that the opportunity to integrate is recognized within 400 msec of experience of a separate yet related learning episode, which then initiates a cascade of subsequent processes that support the integration and further extension of newly acquired knowledge. In Study 3, we found that variability in knowledge extension through integration is associated with measures of scholastic aptitude (SAT) and academic achievement (GPA). Together, the findings inform our understanding of the cognitive and neural factors associated with self-generation and retention of new factual knowledge through integration of separate yet related episodes of new learning, and provide insight into how to promote this educationally relevant learning phenomenon.

Constructing a Knowledge Base through Memory Integration:

Cognitive and Neural Factors Involved

By

Nicole L. Varga

B.S. (Distinguished Honors), Ursinus College, 2010

M.A., Emory University, 2012

Advisor: Patricia J. Bauer, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Psychology
2016

Acknowledgements

This research was funded by NIH HD67359, to Patricia J. Bauer, and by Emory College of Arts and Sciences. I would like to thank Manas Winfield for assistance with stimulus generation, and Amanda Broyles for her help with data collection and other aspects of the research. I would also like to thank my husband, Matt Nixon, for his support throughout the completion of the dissertation. Lastly, I also extend my gratitude to the individuals who took part in the research, without whom this work would not have been possible.

Table of Contents

List of Tables

Study 1

Study 2

Study 3

List of Figures

Study 1

Study 2

Figure 5. Grand averaged ERP waveforms across the frontal-central sites (Panel A: Fz, F3/F4, FC1/FC2, Cz, C3/C4, and CP1/CP2) and parietal sites (Panel B: CP1/CP2, CP5/CP6, Pz, P3/P4, and P7/P8) at encoding.

Figure 6. Grand averaged ERP waveforms across the parietal electrode sites (CP1/CP2, CP5/CP6, Pz, P3/P4, and P7/P8) at the test for self-generation.

Study 3

Figure 1. Frequency distribution of self-generation through integration performance among participants.

Figure 2. Standardized regression coefficients for the relation between self-generation through integration (SG-Int) and SAT score as mediated by verbal comprehension (Panel A) and working memory (Panel B).

Figure 3. Standardized regression coefficients for the relation between self-generation through integration (SG-Int) and GPA (inverse-transformed scores) as mediated by verbal comprehension (Panel A) and working memory (Panel B).

General Introduction

Learning through experience (e.g., direct observation) is a major source of what we know. We also exhibit the striking capacity to extend far beyond what has been directly learned. In fact, much of what we know we have not learned directly, but rather, we have generated for ourselves by integrating knowledge acquired across our experiences. For instance, in one episode an individual may learn that Tamu Massif is the largest volcano in the world. In a separate episode she or he may then learn that the largest volcano in the world is located in Japan. Integration of the two distinct memory traces may then support self-generation of the new knowledge that Tamu Massif is located in Japan, a fact heretofore not directly specified. Self-generative learning through memory integration is pervasive in human cognition; acts ranging from basic creativity to the derivation of scientific theories depend upon it. Yet empirically, the study of memory for specific episodes and of general knowledge acquisition has been quite separate. Researchers have been concerned with either (a) episodic memory: how individuals retain information about unique experiences located in a particular place and time, or (b) semantic knowledge: how information is represented or further extended through logical processes (e.g., induction, deduction, analogy). Until recently, the complementary question of how information learned in separate episodes becomes integrated in memory to form new semantic knowledge has gone largely unexplored.

Despite this gap, researchers across psychology, cognitive science, and cognitive neuroscience implicitly adopt the semantic memory framework in which conceptual knowledge is viewed as *redescriptions* of episodic experience into an abstract, amodal form (see Barsalou, 2012 for a comprehensive review). Indeed, this feature of semantic memory may help explain the absence of episodic information specifying *when* and *where* its contents were acquired (Bauer & Jackson, 2015). Moreover, the extent to which individuals are able to further extend their knowledge has been shown to depend upon the degree to which knowledge becomes integrated in memory (Chi, Hutchinson, & Robin, 1989; Chi & Koeske, 1983). Yet the question of *how* information acquired across episodes is integrated in memory such that it can be used productively is only beginning to be understood.

We have conducted a number of studies on self-generation of new knowledge through integration of separate episodes with children 4 and 6 years of age (Bauer, King, Larkina, Varga, & White, 2012; Bauer & San Souci, 2010; Bauer, Varga, King, Nolen, & White, 2015; Varga & Bauer, 2013; Varga, Stewart, & Bauer, 2016), children 7 to 10 years of age (Bauer, Blue, Xu, & Esposito, in press; Bauer & Larkina, in press), and college students (Bauer & Jackson, 2015). For children, true yet novel facts ("stem" facts) are conveyed through text passages (i.e., constituting different "episodes" of experience) read aloud by an experimenter; related passages are separated by unrelated passages and buffer activities. Importantly, separate yet related passages can be integrated to self-generate a novel, indirectly learned understanding (i.e., an "integration" fact). For example, there were two stem facts pertaining to dolphins (i.e., dolphins talk by clicking and squeaking; dolphins live in groups called pods) that could be combined to generate a novel integration fact (i.e., pods talk by clicking and squeaking). Children are tested for self-generation of new knowledge through integration of the separate episodes, first in open-ended format ("How does a pod talk?") and when self-generation fails, via forced-choice. Children's self-generation of the integration facts in open-ended testing ranges from 13% for 4-year-olds, to 87% for 8-year-olds. Four-year-olds' total performance (across open-ended and forced-choice) approaches 70%. Performance in a 1-stem control condition in which children are exposed to only half of the information required to generate the novel integration facts is substantially lower, and in forced-choice testing does not differ from chance, making clear that integration of separate episodes is necessary for productive extension of the new knowledge. Furthermore, children not only generate new knowledge through integration, they also retain it over time, with virtually no loss of access after 1 week (Varga & Bauer, 2013; Varga et al., 2016). These patterns suggest that the paradigm captures an ecologically valid learning mechanism.

For purposes of the present research, we used an adult-appropriate adaptation of the developmental paradigm. Rather than through text passages, college students were exposed to novel stem facts as they read long lists of seemingly unrelated information, one sentence at a time on a monitor. After a buffer activity, they were tested for self-generation of new knowledge through integration. In Bauer and Jackson (2015), testing was through forced-choice; subjects selected the correct responses on 56% of trials, compared to 27% correct in a 1-stem control condition (25% expected by chance). In the present research, we tested for replication of Bauer and Jackson (2015) and extended it in several ways. In Study 1, we investigated factors known to influence accessibility of separate episodes to one another, namely, the degree of lag imposed between to-be-integrated stem facts. Moreover, to provide a fuller picture of how a real-world semantic knowledge base is built over time, we extended beyond forced choice to provide the first test of self-generation of new knowledge in young adults, as well as whether newly self-generated knowledge is retained over time. In Study 2, we utilized the event-related potential (ERP) technique to test the hypothesis that self-generation of new knowledge through memory integration consists of several distinct neurocognitive subprocesses. Finally, as Barsalou and Prinz (1997) initially postulated, a brief glimpse into the inventions, ideas, and scientific breakthroughs achieved throughout history would seem to suggest that some individuals exhibit a truly exceptional knack for self-generating new knowledge. Conversely, other people may exhibit a lesser proclivity for this capacity. Indeed, to foreshadow the dissertation findings, young adults exhibit substantial individual differences in the extent to which they engage in, and are successful at, self-generating new knowledge through integration. In an effort to explain such wide variability in what is assumed to be a fundamental and pervasive learning mechanism, in Study 3 we sought to delineate the role of domain-general cognitive abilities on successful self-generative learning through integration. In light of the potential consequences such individual differences might have on educational outcome, we also investigated the association between task performance and real-world metrics of academic success, including college GPA and SAT. Because the theoretical motivation for these questions, and relevant empirical findings, are discussed at length in the individual manuscripts, I will not elaborate on them here.

There is one exploratory aspect of the dissertation that was not intended to be included in the manuscripts but which appears in Addenda 1-3 at the end of the thesis. Specifically, the dissertation proposal included a plan to classify the types of reasoning problems included in the stimulus set, and to provide descriptive measures of performance by reasoning type. Post-hoc classification indicated that the stimuli enabled integration through a number of logical relations, including substitution based on the principle of equivalence, transitive inference, and general to specific deductions. Addenda 1 and 2 provide descriptions of initial self-generation by each fact type in Experiments 2 and 3 of Study 1, respectively. Due to vastly unequal numbers of stimuli in each reasoning category, the paradigm was not suited for statistical tests of differences across problem types. Nevertheless, examination of descriptive measures of mean proportion correct suggests that performance was comparable across the logical relations identified.

In summary, the overarching purpose of the dissertation research was to illuminate the behavioral, neural, and cognitive factors associated with the self-generation of new factual knowledge through memory integration. The question was pursued using a newly designed analogue of the paradigm previously employed with children, which was suitable for assessing knowledge extension both in an open-ended format as well as while ERPs were recorded. Because we were interested in delineating a process model that is predictive of *successful* versus *unsuccessful* self-generation, we addressed this question in adults: a population in which we could maximize the number of facts presented, thereby allowing us to assess a range of performance and to obtain a sufficient number of ERP trials per subject to enable valid analysis. With the exception of Experiments 1 and 2 from Study 1, the participants and procedures included across the three manuscripts were the same, whereas the data featured in each were unique. As such, there is a certain amount of repetition throughout the method section of each manuscript. For the moderate effect sizes expected, G*Power 3.1.7 indicated a sample of approximately 98 was needed to achieve power of 80%. Specifically, this sample size allowed for: (1) test of individual difference predictors via correlation/regression analysis on the entire sample, and (2) examination of the neural ERP components underlying successful versus unsuccessful knowledge extension in only those individuals who comprised the middle of the performance distribution (i.e., who exhibited close to 50/50% performance, thus ensuring that there were a sufficient number of successful and unsuccessful trials per subject). Thus, to satisfy the criteria for ERP group analyses (Study 2) and individual differences analyses (Study 3), the final *N* was 120 adults (half women). Together, the studies provide important insight into the cognitive and neural processes associated with knowledge extension through integration, their timing, and how variability contributes to real-world academic outcomes.

Study 1

Knowledge extension through memory integration:

Factors and conditions that promote the long-term accumulation of knowledge

Nicole L. Varga & Patricia J. Bauer

Emory University

Knowledge extension through memory integration: Factors and conditions that promote the long-term accumulation of knowledge

The construction of a knowledge base fundamentally relies on memory integration—the combination of information acquired within or across separate learning episodes. Indeed, without the ability to integrate information learned at different times and in different places, building a domain of knowledge would not be possible. Moreover, the formation of an integrated knowledge base also permits flexible extension beyond direct experience, enabling self-generation of new thoughts, ideas, and understandings. Prior research on productive knowledge extension provides important insight into the mechanisms involved in memory integration in nonhuman animals (e.g., Bunsey & Eichenbaum, 1996; Tse et al., 2007; 2011), in adults (e.g., Bauer & Jackson, 2015; Kumaran Summerfield, Hassabis, & Maguire, 2009; Zeithamova, Dominick, & Preston, 2012a), and in children (e.g., Bauer, King, Larkina, Varga, & White, 2012; Bauer, Varga, King, Nolen, & White, 2015). Yet although memory integration is presumed to serve as the key mechanism through which a knowledge base is formed (Bauer, 2012; Bauer & Varga, 2016; Preston & Eichenbaum, 2013; Siegler, 1989), the long-term retention of knowledge newly derived through memory integration has not been examined in adults. Furthermore, few studies with adults have investigated how this phenomenon operates under conditions that mimic those encountered in the world outside the laboratory. To address these gaps, in the present research, we examined whether knowledge extension through integration is affected by the amount of temporal spacing imposed between separate yet related episodes of new learning (Experiment 1), whether memory integration underlies more flexible forms of knowledge extension than have previously been tested (Experiment 2), and whether factual knowledge newly derived through integration becomes incorporated into the semantic knowledge base as evidenced through long-term accessibility (Experiment 3).

Knowledge extension through memory integration has been the focus of substantial research in nonhuman animals (e.g., Bunsey & Eichenbaum, 1996; Dusek & Eichenbaum, 1997) and human adults (Bauer & Jackson, 2015; Kumaran et al., 2009; Preston, Shrager, Dudukovic, & Gabrieli, 2004; Schlichting, Zeithamova, & Preston, 2014; Shohamy & Wagner, 2008; Sweegers, Takashima, Fernández, & Talamini, 2014; Zeithamova et al., 2012a; Zeithamova & Preston, 2010). The primary methods used to study this topic include transitive inference and associative inference, both of which necessitate integration of overlapping yet arbitrary stimulus pairs (although see Bauer & Jackson, 2015, discussed below, for an exception). For instance, in transitive inference, subjects learn a set of premises through trial-and-error and reinforcement (e.g., A > B, B > C, C > D, D > E, E > F), such as odors in rats (e.g., Dusek & Eichenbaum, 1997) or visual patterns in humans (e.g., Heckers, Zalesak, Weiss, Ditman, & Titone, 2004). Once a criterion level of performance is reached, subjects then are tested via forced-choice selection for knowledge of both directly trained pairs (e.g., A > B) and of untrained, indirectly learned associations (e.g., B > E). Success on the untrained pairs is contingent on integration across premises in order to represent the hierarchy of relations. Whereas transitive inference requires repeated exposures to elicit integration, associative inference enables examination of knowledge extension through integration under single-trial learning conditions. Specifically, subjects are presented with temporally distributed stimulus pairs (e.g., AB: *Chair & Basketball*, BC: *Basketball & Blender*) that do not form a hierarchy. They then are tested for integration of the overlapping episodes via a forced-choice transfer test (e.g., A:*Chair* = ?) (e.g., Zeithamova & Preston, 2010). Converging evidence from these paradigms indicates that the capacity to form and flexibly express novel relational understandings is conserved across species, thus underscoring the significance of this process.

Despite the presumed importance of memory integration for the acquisition of knowledge, this mechanism has not been investigated under conditions that mimic those encountered in everyday learning situations in which the target of learning is factual knowledge as opposed to arbitrary paired associates. Therefore, for the purposes of the present research we adopted an ecologically valid method used to examine the development of knowledge extension through integration in children (Bauer et al., 2012, 2015; Bauer, Blue, Xu, & Esposito, in press; Bauer & Larkina, in press; Bauer & San Souci, 2010; Varga & Bauer, 2013, 2014; Varga, Stewart, & Bauer, 2016). In this approach, children are taught true but novel, related “stem” facts in the context of separate story passages (e.g., *Dolphins talk by clicking and squeaking; Dolphins live in groups called pods*) and then are tested for self-generation of new knowledge through integration of the target information (e.g., *How does a pod talk?*). Relative to the inference paradigms employed with nonhuman animals and adults, this method has the advantage of being about real-world facts, and thus is directly relevant to the issue of how a semantic knowledge base is built over time. Indeed, studies with 4- and 6-year-olds indicate that knowledge newly derived through memory integration persists in memory over time (Varga et al., 2016; Varga & Bauer, 2013), thus making it an appropriate methodology for examination of the long-term accumulation of an integrated knowledge base in adults.

The present research builds on a study by Bauer and Jackson (2015) which employed an adult analogue of the knowledge extension paradigm used with children. Specifically, adults incidentally learned true but novel stem facts (*Apple seeds are called pips; Cyanide is found in pips*) and then were tested for self-derivation of new factual knowledge through integration of the target information (*Apple seeds contain ___?*). During the learning phase, individuals read 120 individual sentences, each of which conveyed a fact. Ninety of the facts were well-known (*Washington D.C. is the capital of the United States*), 20 could be integrated to form 10 novel integration facts (i.e., 2-stem condition), and 10 conveyed one-half of the information necessary to form 10 novel integration facts (i.e., 1-stem control condition). At the time of test, participants were shown incomplete facts that had not been presented previously and were asked to fill in the final word of each sentence via forced-choice selection. Of the 40 facts tested, 10 were Well-known, 10 were derived through integration of the 2-stem facts, 10 were derived through integration of the 1-stem facts, and 10 were novel. Bauer and Jackson (2015) found that young adults derived the novel integration fact on 56% of the trials in which participants could integrate (2-stem condition), which is consistent with patterns observed in other paradigms (e.g., Schlichting et al., 2014). Moreover, performance on the 2-stem integration facts was intermediate to well-known and novel facts, signifying its transitional status within semantic memory. Importantly, when only a single stem fact was provided (1-stem condition), performance did not exceed chance levels (27% with 4 choice alternatives). Therefore, exposure to both stem facts from a target pair was necessary to reliably produce the integration facts, thus indicating that the integration facts were novel. Based on the utility of this paradigm, in the present research we tested for replication of Bauer and Jackson (2015) and extended the approach to investigate knowledge extension through integration under several ecologically valid learning conditions.

In Experiment 1 we extended Bauer and Jackson (2015) by examining the effect of a factor known to influence accessibility of separate episodes to one another (and thus which logically should influence productive extension beyond direct experience), namely, the lag (or temporal distance) between to-be-integrated stem facts (e.g., Kahana & Howard, 2005). Results of prior research make clear that successful memory integration is accomplished through reactivation of prior, related episodes upon experience of new information (Zeithamova et al., 2012a). To continue with the previous example, to successfully integrate separate yet related information in memory, previously learned material (*Apple seeds are called pips*) must be retrieved while processing novel yet related information *(Cyanide is found in pips*), and the overlapping content among the traces must be bound in memory (*pips*). As such, we hypothesized that successful knowledge extension through integration of previously learned information would become more difficult as the lag between separate yet related learning episodes increased. Support for this prediction also comes from the finding that temporal spacing between newly and previously learned information has been shown to impair knowledge extension through integration in young children (Varga & Bauer, 2013). Specifically, when a 1-week delay was imposed between experience of related stem facts, 6-year-olds self-generated the novel integration fact on only 23% of the trials, compared to 60% when no delay was imposed. Importantly, the decrements were similarly observed when only those trials in which both stem facts necessary for self-generation were recalled, thereby suggesting that differential memory could not account for diminished performance. To compare the effects of lag to results reported in prior research with adults, in the present study we manipulated the temporal distance between to-be-integrated stem facts within a single learning session. Accordingly, the primary aim of Experiment 1 was to test the differential effect of short lags (1-10 intervening facts) versus long lags (40-80 intervening facts) on subsequent knowledge extension through integration. In addition, aggregation of performance across the lag conditions also provided a direct test for replication of Bauer and Jackson (2015).

In Experiment 2 we moved beyond forced-choice measures to examine whether memory integration supports knowledge extension under more demanding testing conditions. Current theories propose that memory integration underlies a number of complex cognitive feats, such as creativity and imagination, behaviors that undoubtedly rely on flexible forms of knowledge extension (Schacter et al., 2012; Schlichting & Preston, 2015). Yet to date, studies with adults have only examined knowledge extension through memory integration using forced-choice measures. Thus, in Experiment 2 we designed a novel stimulus set that could be used to test the frequency with which young adults *self-generate* integrated knowledge in an open-ended form. This question is of both practical and theoretical significance. From a practical perspective, individuals are frequently called upon to demonstrate their knowledge without the provision of options from which to choose, such as when explaining something recently learned to a colleague. From a theoretical perspective, superior memory in forced-choice measures (in comparison to open-ended measures) is well-documented for directly experienced events (e.g., Haist, Shimamura, & Squire, 1992). Indeed, some have proposed that forced-choice permits accurate responding based on a weaker memory trace (see Squire, Wixted, and Clark, 2007 for review). Moreover, whereas knowledge extension through forced-choice is evident in children as young as 4 years of age, it is not until ages 6 and 8 years that children reliably *self-generate* novel information in memory integration paradigms (Bauer & Larkina, in press; Bauer & San Souci, 2010). Experiment 2 provides the first empirical test of whether adults are successful at extending knowledge under more demanding, open-ended test conditions.

To foreshadow the results of Experiment 2, young adults extended knowledge through integration under the more demanding conditions of open-ended testing. In Experiment 3, we tested whether newly self-generated knowledge is retained. For knowledge extension through integration to be psychologically, cognitively, and educationally meaningful, its products must persist in memory over time. Indeed, current theories regarding the nature of knowledge acquisition cite memory integration as the key mechanism through which a knowledge base is formed (e.g., Bauer, 2012; Bauer & Varga, 2016; Preston & Eichenbaum, 2013). Yet although the extant literature with adults undoubtedly captures integrative mechanisms presumed to be involved in the long-term accumulation of knowledge, this outcome has not been examined directly. As discussed previously, because the most common knowledge extension paradigms employed with adults rely on arbitrary stimuli that are unlikely to be incorporated into the knowledge base, this literature has not addressed the role of memory integration in the long-term accumulation of knowledge. Based on findings that young children have strong memory for real-world knowledge self-derived through integration (Varga et al., 2016; Varga & Bauer, 2013), we expected to observe a similar pattern in young adults. Moreover, Bauer and Jackson (2015) showed that integrated knowledge is incorporated into semantic memory within a single study session. However, we do not know whether that information was retained in memory over time. We addressed this question in Experiment 3. We tested retention after a delay of 1 week because this is a period of time over which other age groups have demonstrated memory for self-derived knowledge (Varga et al., 2016; Varga & Bauer, 2013).

Finally, we also used Experiment 3 to begin the task of identifying the source(s) of individual differences documented both in the present research (Experiments 1 and 2) and in prior research with other knowledge extension paradigms. For instance, using an incidental acquired equivalence knowledge extension task, Shohamy and Wagner (2008) observed a range of performance from 38% to 100% correct in successful knowledge extension through integration in young adults. They also found that better performers exhibited significantly faster responses during the test for knowledge extension, thereby suggesting that integrative encoding might have been initiated at the time of learning. Interestingly, however, only two of the 24 participants reported explicit awareness of the opportunity to integrate. Guided by these findings, in Experiment 3, we assessed whether both explicit awareness of the task structure and response speed at the time of test were associated with individual differences in the capacity to self-generate new factual knowledge through integration.

Taken together, the present research contributes valuable insight into the later accessibility of knowledge newly self-generated through memory integration, in addition to the factors and conditions that affect this learning mechanism. We chose to focus the inquiry on adults because the overarching aim was to characterize knowledge extension through integration under conditions commonly encountered in everyday life, with particular emphasis on factors that have only previously been systematically investigated in children.

Experiment 1

Method

Participants

Participants were 17 adults between 18-21 years (*M* = 19.61, *SD* = 0.80; 11 females) enrolled in undergraduate psychology courses at a private university. Based on self-report, the sample was 18% African American, 12% Asian, and 70% Caucasian. None of the participants was of Hispanic descent. Data from one participant were excluded due to failure to reach above-chance accuracy on facts intended to be well-known (see below). The protocol and procedures were approved by the university Institutional Review Board. Written informed consent was obtained prior to the start of the study session. Participants were compensated with course credit.

Stimuli

The stimuli were the same facts employed in prior, related research (Bauer & Jackson, 2015). Specifically, the encoding phase materials consisted of 120 sentences, each ranging from 5-15 words in length. Forty sentences conveyed facts that participants were likely to know and which, based on Virginia state academic standards, should have been acquired in elementary school ("*a globe is a round model of the earth*"), middle school ("*a parallelogram is a quadrilateral with two sets of parallel sides*"), high school ("*World War II began with the German invasion of Poland*"), or college ("*Kierkegaard and Nietzsche are the fathers of Existentialism*"). The remaining 80 sentences consisted of 40 pairs of related facts (i.e., stem facts) that subjects were unlikely to know ("*the first jigsaw puzzles were called dissections"* and *"the first jigsaw puzzles were used to teach children about geography*"). Through integration of separate yet related stem facts, participants could derive a novel integration fact never directly specified (*"dissections were used to teach children about geography"*). Half of the related stem facts were presented in a Short-lag condition (*N* = 20 pairs) and the other half were presented in a Long-lag condition (*N* = 20 pairs). Short-lag stem facts were separated by 1-10 intervening sentences (*M lag* = 5 sentences), whereas, Long-lag stem facts were separated by 40-80 intervening sentences (*M lag* = 60 sentences).

The test phase materials consisted of 80 sentences that had not been previously presented, each ranging from 5-9 words in length. Some facts were expected to be well-known (*"the Tyrannosaurus Rex is now extinct"; N* = 20) whereas other facts served as novel control facts unrelated to any previously encountered information (*"snails can sleep three years without eating"*; *N* = 20). The remaining sentences were facts derived through integration of previously presented stem facts, presented in Short-lag (*N* = 20) and Long-lag (*N* = 20) conditions. Prior research employing these stimuli has demonstrated that the target integration facts are novel to young adults and that both facts from a given pair are necessary for derivation of the integration fact. Specifically, when only one of the stem facts from a target pair was presented (Bauer & Jackson, 2015; Experiment 2) or when one of the stem facts was replaced with a "Pseudo-stem" that was related but could not be combined to derive the novel integration fact (Bauer & Jackson, 2015; Experiment 1b), behavioral and neural responses to the integration facts did not differ from Novel facts, respectively.

Procedure

Encoding phase. Participants were informed that we were interested in the perceived difficulty of various facts. Participants then read 120 sentences on a laptop and made judgments about when in school the facts might have been learned: elementary school, middle school, high school, or college (see Figure 1, Panel A; Bauer & Jackson, 2015). Sentences were presented using Microsoft PowerPoint® and appeared on the screen for 100 ms per character. Participants had an unlimited amount of time to provide their grade-level judgment and did so by stating it aloud. Difficulty judgments were collected to ensure that participants were attending to the information and were not considered in the data analysis. Each pair of related facts appeared in the Short-lag and Long-lag condition equally often across the sample. Moreover, stem fact order was counterbalanced such that each fact from a target pair appeared in the first or second serial position an approximately equal number of times. At no time were participants told that any sentences were related. None of the test-phase facts were presented.

Test phase. After the encoding phase, participants engaged in buffer activities lasting approximately 15 minutes. Participants then read 80 new facts (20 per condition). Unlike the encoding phase presentation, sentences in the test phase were presented in the form of questions (see Figure 1, Panel B). Specifically, the final word of each sentence was replaced by a decision screen with four answer choices, one of which was correct (the other three choices were conceptually related distracters). Participants were instructed to select the word that accurately completed the fact. Stimuli were presented and response times were measured using ASA computer software (A.N.T. Software B.V., Enschede, the Netherlands). Participants had an unlimited amount of time to make a choice and did so via a button-press response. Response time was recorded from the onset of the decision screen until the button-press. To ensure comprehension of the task, participants completed four practice trials requiring use of all buttons prior to beginning the test phase. Finally, to assess prior knowledge of the facts tested, at the end of the session participants heard each of the 80 test phase facts aloud and were asked to indicate whether they knew the fact prior to participation in the study. This familiarity check was completed immediately after the test phase as not to contaminate forced-choice fact performance.

Results

Performance across Fact Conditions

We first examined how knowledge extension performance in the present sample compared to that observed in Bauer and Jackson (2015). Participants received a score of 1 or 0 (correct or incorrect) for each fact tested. When accuracy was collapsed across the Short-lag and Long-lag conditions, participants answered 58% of the Integration fact questions correctly (*M* = 23.24, *SD* = 6.62), providing a direct replication of the 56% previously reported by Bauer and Jackson (2015). We next examined, the primary question of interest, namely, whether performance differed across the four fact conditions, and whether derivation of the target integration facts differed as a function of the lag manipulation. To examine this question, a one-way repeated measures analysis of variance (ANOVA) was conducted for mean number of correct responses and mean response time across the four conditions. For all ANOVAs, Greenhouse-Geisser corrections were applied in cases of violation of sphericity. Bonferroni corrections were applied to *t*-values for post-hoc analyses with multiple comparisons.

As depicted in Figure 2, Panel A, examination of accuracy revealed a main effect of condition, $F(2.05, 32.79) = 89.24$, $p < .001$, $p\eta^2 = .848$, with greater accuracy observed for Well-known facts compared to all other facts (*p*s < .001) and poorer accuracy observed for Novel facts compared to all other facts (*p*s < .001). Although accuracy on the Short-lag and Long-lag integration facts significantly differed from that of Well-known and Novel facts, accuracy in the Short- and Long-lag integration conditions did not differ ($p = 1.00$). Not only was performance on Novel facts lower than all other facts, it also was not different from chance levels (25% correct), $t(16) = 1.99$, $p = .07$. In contrast, performance significantly differed from chance for Well-known facts ($t(16) = 64.09$, $p < .001$), Short-lag facts ($t(16) = 8.47$, $p < .001$), and Long-lag facts ($t(16) = 6.52$, $p < .001$). Thus, participants reliably derived the novel integration facts but successful knowledge extension did not differ as a function of the lag manipulation employed.

As reflected in in Figure 2 (Panel B), a parallel pattern of results was observed when mean reaction time was examined across the four fact conditions. That is, a main effect of condition was found, $F(3, 48) = 47.27$, $p < .001$, $p\eta^2 = .747$, such that responses to Well-known facts were significantly faster compared to response times to all other facts (*p*s < .001) and responses to Novel facts were significantly slower compared to response times to all other facts (*p*s < .001). Consistent with the accuracy findings, mean reaction time to the Short-lag and Long-lag integration facts did not differ ($p = 1.00$).

Thus, as information becomes better known, response speed increases along with accuracy, yet there is no difference as a function of lag.

Prior Knowledge of the Integration Facts

To examine possible effects of prior knowledge of the test stimuli on performance, we analyzed participants' self-report measures. For each of the 80 test phase facts, participants received a score of 1 or 0 (reported known or unknown, respectively). Examination of self-reported prior knowledge by stimulus type revealed that participants claimed to know 98% of the Well-known facts (*M* = 19.67; *SD* = 0.69), but only 9% of the integration facts (*M* = 3.72; *SD* =3.05), and 7% of the Novel facts (*M* = 1.44; *SD* = 1.15). In a correlational analysis, there was a moderate, marginally significant positive relation between self-reported prior knowledge and performance on Well-known facts (*r*(15) = .46, *p* = .06). No such relation was observed for the novel stimuli (Short-lag, Long-lag, and Novel), *r*(15) = .17, *p* = .51. As a final check of whether the present results were impacted by prior knowledge, we repeated all analyses taking self-report ratings into account (i.e., by eliminating from analysis on a subject-by-subject basis Well-known facts of which subjects reported no knowledge and Integration and Novel facts of which subjects had prior knowledge). The secondary analyses did not change the pattern of results reported above.

Discussion

The results from the present experiment indicate that the degree of temporal spacing between related facts does not affect integration across them, at least as tested through the lag manipulation employed here. Importantly, however, performance on the Integration facts was intermediate to that of Well-known and Novel facts and exceeded chance levels, suggesting that individuals did indeed derive new knowledge through integration. Moreover, because performance replicated findings from prior research (Bauer & Jackson, 2015), results of the present study are unlikely to be explained by unique features of the current sample. Based on this evidence, it is reasonable to conclude that the lag between separate yet related traces of information does not affect the capacity to extend beyond information acquired within a single session, at least as tested in this research. Nevertheless, it is also possible that lag effects might only be detected under conditions that assess self-derivation through integration under more demanding test conditions (i.e., open-ended vs. forced-choice testing). That is to say, if forced-choice permits accurate responding based on a weaker memory trace (see Squire et al., 2007), then any effect that lag might have had at encoding could have been overcome by the facilitative effect of highly supportive testing conditions. To address this possibility, in Experiment 2 we designed and validated a new stimulus set to assess knowledge extension through integration in open-ended form.

Experiment 2

Method

Participants

Participants were 31 adults between 18-24 years (*M* = 19.63 years, *SD* = 1.26; 24 females) drawn from the same population as in Experiment 1; none of the participants had taken part in Experiment 1. An additional 3 participants took part in the study but were excluded from analysis due to failure to comply with task instructions (*N* = 1) or to meet the native English criteria (*N* = 2). Based on self-report, the sample was 36% African American, 13% Asian, 48% Caucasian, and 3% mixed racial descent. Ten percent of the participants were of Hispanic descent. The protocol and procedures were approved by the university Institutional Review Board. Written informed consent was obtained prior to the start of the study. Participants received course credit at the completion of their participation.

Stimuli

Encoding phase stimuli consisted of 75 sentences ranging from 5-10 words in length. Sixty sentences featured 30 pairs of related stem facts that could be combined to generate 30 novel integration facts. The remaining 15 sentences featured unrelated distracter facts that were of equivalent perceived difficulty and were drawn from similar subject domains.

The test phase stimuli consisted of 30 sentences ranging from 4-10 words in length, none of which had been previously presented. Each sentence featured a novel integration fact that could be derived through integration of stem facts previously presented during the encoding phase. For instance, two stem facts were about art history (*A popular sculpture made from a urinal is called Fountain; Duchamp's most well-known work is named Fountain)*. Integration of separate but related stem facts could lend itself to self-generation of a novel integration fact (*Duchamp's most popular work consisted of a urinal*). The test sentences were presented in the form of questions by omitting the final word of each fact. Participants were asked to generate a one-word answer that could accurately fill in the blank space. Importantly, unlike the stimuli employed in Experiment 1, the integration facts in the present study ended in a sentence-final word that should have been familiar to participants (e.g., *urinal*). This additional constraint on the stimulus set allowed for assessment of *self-generation* of new knowledge through integration (as opposed to only forced-choice selection of self-derived knowledge). On the other hand, this stimulus design also increased the likelihood of accurate responding in forced-choice testing based on familiarity to the sentence-final word of each integration fact. As such, forced-choice analyses are presented only to provide a fuller picture of participants' performance. However, because participants were expected to have prior familiarity of the sentence-final words, forced-choice performance was not examined as a function of self-reported prior knowledge.

Procedure

Encoding phase. At the start of the session, participants were instructed that we were interested in whether memory for newly learned factual information differs as a function of its subject domain. To ensure that knowledge extension through integration could only be accomplished through integration of the target stem facts, participants read a total of 60 sentences: 30 2-stem facts (15 complete pairs of stem facts), 15 1-stem facts (15 individual stem facts without a corresponding paired fact), and 15 distracter facts. As depicted in Figures 3A and 3B, sentences were presented one word at a time for 400ms. Each sentence ended in a target word, which served as the relational link between to-be-integrated stem facts in the 2-stem condition. At the end of each sentence, participants were shown a decision screen and asked to indicate, via a button-press response, whether the information conveyed was novel or known. The distracter task was designed to ensure that participants were attending to the facts while also corroborating the pretext of the study purpose (i.e., learning of *novel* information). At no time were participants informed that any of the sentences were related.

Across the encoding presentation, to-be-integrated stem facts were separated by a lag of 2, 3, or 4 intervening sentences. Variability in lag created temporal distance between to-be-integrated information and also prevented participants from anticipating the content of the next fact. Moreover, the short lag was intended to promote recognition of the relatedness between paired facts, thereby allowing for examination of knowledge integration under supportive learning conditions. Lag assignment was equated such that each fact was tested in each lag condition an equal number of times across the sample. Additionally, fact assignment was counterbalanced such that pairs of stem facts were tested in a 1-stem and a 2-stem condition an approximately equal number of times, with stem facts from a target pair appearing as the 1-stem control sentence equally often. Fact order within the 2-stem condition was also counterbalanced. That is, each stem fact from a complete pair was presented in the first or second serial position an approximately equal number of times across the sample.

Test phase. After a short break lasting 5-10 minutes, participants were presented with 30 facts derived through integration of the previously presented stem facts using PowerPoint® software. As depicted in Figure 3C, the sentences were presented in the form of questions by omitting the final word of each fact. Participants were asked to generate a one-word answer that could accurately fill in the blank space. Participants were given an unlimited amount of time and provided their answers aloud. Specifically, when an answer was generated, participants made a button-press response which was followed by an “Answer” screen cueing them to speak the answer aloud. Once an answer was provided, the experimenter initiated presentation of the next integration question. Following the open-ended questions, participants received forced-choice questions for any integration facts that were not successfully self-generated. Specifically, they were shown the same incomplete integration fact as in open-ended testing while the experimenter read four answer choices aloud. Participants were instructed to select the answer that accurately completed the fact, one of which was correct (the other three choices served as conceptual distracters).

Immediately following forced-choice testing, participants heard the 30 integration facts on which they were previously tested (e.g., *Apple seeds contain* cyanide) in addition to 35 distracter facts not previously presented. Of the new facts, 20 were expected to be familiar (e.g., *a ruler measures the length of objects*) whereas 15 were expected to be novel (e.g., *the most pungent fruit in Asia is the Durian*). After each fact, participants were asked to indicate whether they knew the fact prior to participating in the study, thereby providing a subjective measure of participant's pre-existing knowledge of the integration facts on which they were tested.

Results

Performance across Fact Conditions

The main purpose of the present experiment was to test whether the newly designed integration facts could be self-generated when participants were exposed to both stem facts from a pair (2-stem condition), but not when only one of the two facts was presented (1-stem condition). As shown in Figure 4 (Panel A), participants self-generated significantly more integration facts in the 2-stem as compared to the 1-stem condition, $t(30) = 7.55$, $p < 0.001$. A parallel pattern of results was observed in forced-choice testing, $t(28) = 8.15$, $p < 0.001$. Because one of the purposes of the present research was to characterize knowledge extension through integration under more demanding test conditions at both the group and individual levels, we also examined variability in performance. As shown in Figure 4 (Panel B), successful self-generation performance in the 2-stem condition ranged from 7% to 100%.

Prior Knowledge of the Stem Facts and Integration Facts

We first examined participants' self-reported prior knowledge of the individual stem facts. For each of the 60 facts, participants received a score of 1 or 0 (known or unknown). When ratings were collapsed across facts presented in the 1-stem and 2-stem conditions, 19% (255 out of 1333 possible trials) of the total facts were reported as previously known (*M proportion* = 0.19; *SD* = 0.11). Ratings of prior knowledge did not differ as a function of whether the fact appeared in a 1-stem (*M proportion* = 0.23; *SD* = 0.20; 99 out of 439 possible trials) or a 2-stem condition (*M proportion* = 0.17; *SD* = 0.11; 156 out of 894 possible trials), *t*(30) = 1.45, *p* = 0.16. Importantly, for facts presented in the 1-stem condition, there was no significant relation between prior knowledge and subsequent self-generation performance, *r*(29) = .19, *p* = 0.31. Thus, participants' reported prior knowledge of a single stem fact did not predict subsequent production of the novel integration facts. There was a significant correlation between reported prior knowledge and subsequent self-generation performance when both members of the stem fact pair were presented (i.e., the 2-stem condition) *r*(29) = .38, *p* = 0.04. This relation implies that prior knowledge of one or the other of the stem facts facilitated self-generation performance.

When self-reported knowledge of the integration facts was collapsed across facts presented in the 1-stem and 2-stem conditions, 11% (99 out of 886 possible trials) of the total facts were reported as previously known (*M proportion* = 0.11; *SD* = 0.09). Ratings of prior knowledge did not differ as a function of whether the fact appeared in a 1-stem (*M proportion* = 0.10; *SD* = 0.09; 43 out of 439 possible trials) or a 2-stem condition (*M proportion* = 0.13; *SD* = 0.12; 56 out of 447 possible trials), *t*(30) = 1.24, *p* = 0.23. There was no significant correlation between self-generation performance and self-reported knowledge of the integration facts, $r(29) = .24$, $p = 0.20$, indicating that participants' ratings were largely unrelated to whether they had produced the integration fact during the open-ended test.

Discussion

A major challenge encountered with designing any ecologically valid paradigm is the need to ensure that performance cannot be accounted for by prior knowledge of the materials employed. Significantly lower 1-stem performance relative to 2-stem performance in both open-ended and forced-choice testing in the present experiment indicates that, with the exception of a select number of trials, integration of the related stem facts was necessary for successful knowledge extension. Although participants claimed to know 19% of the individual stem facts and 11% of the target integration facts, prior knowledge of the 1-stem facts and the integration facts was not reliably associated with self-generation performance. Logically, however, a significant relation was evident between prior knowledge of the individual 2-stem facts and self-generation performance. That is to say, if participants knew one stem fact, it could reasonably be expected to facilitate processing of the related stem fact and thus self-generation of integrated knowledge. This raises the interesting possibility and need for future research examining self-generation through integration when prior knowledge of the stem facts is directly manipulated. The important point with respect to the present research, however, is that knowledge of the 1-stem facts was not associated with heightened self-generation performance. In other words, prior knowledge of only one of the stem facts did not support production of the novel integration fact.

The results from the present experiment also indicated that there are extensive individual differences in the capacity to *self-generate* novel integration facts. Given the sensitivity of the paradigm to individual differences, in Experiment 3, we extended the paradigm to examine factors that might be associated with successful knowledge extension, including explicit awareness of the task structure and response times at test. In addition, in Experiment 3, we conducted the first empirical test of whether knowledge newly self-generated through integration persists in semantic memory over time in adults.

Experiment 3

Method

Participants

Participants were 117 adults between 18-24 years (*M* = 19.76 years, *SD* = 1.15; 63 females) drawn from the same population as in the previous experiments. None of the participants had taken part in Experiments 1 or 2. Based on self-report, the sample was 9% African American, 25% Asian, 59% Caucasian, and 4% mixed racial descent. Eight percent of the participants were of Hispanic descent. Three participants did not report racial or ethnic information An additional 3 participants took part in the study but were excluded due to failure to comply with task instructions (*N* = 1) and self-reported diagnosis of Dyslexia which may have negatively impacted task performance (*N* = 2). Retention data were missing for two participants due to failure to return for the second visit (*N* =1) and failure to return within the specified delay interval (*N* = 1). The protocol and procedures were approved by the university Institutional Review Board. Written informed consent was obtained prior to the start of the study session. Participants received course credit upon completion of the second visit.

Stimuli

The encoding phase stimuli consisted of the same 30 pairs of related stem facts employed in Experiment 2. Based on significantly poorer performance in the one-stem control condition reported in the previous experiment, all stem facts were presented in a two-stem condition in the present research. The test phase materials were identical to those employed in Experiment 2.

Procedure

Participants completed two sessions separated by a delay of approximately 1 week (*M* delay = 6.91, *SD* = 0.54, *Range* = 6-8 days). Participants were tested individually by two female experimenters, each of whom tested an approximately equal number of participants from each gender. With the exception of six individuals, participants were tested by the same experimenter at each session. The experimenters followed the same detailed written protocol and regularly reviewed audio-recorded sessions with each other to ensure protocol fidelity.

Session 1. The procedure was the same as in Experiment 2 with some exceptions. First, sentences were presented in the same manner as in the previous experiment (see Figure 3), but in order to collect precise reaction time measures, facts were presented using E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). As depicted in Figure 3 Panel C, reaction time measures were time-locked to the presentation of the question mark which first cued participants to generate an answer. Second, although the test phase procedures were identical to those employed in Experiment 2, forced-choice performance was not tested until Session 2. This prevented successful self-derivation via forced-choice from inflating open-ended performance at the second time point, thereby ensuring an uncontaminated measure of long-term retention. Lastly, following the open-ended questions at Session 1, participants completed a debriefing survey inquiring about their explicit perceptions of the task, which was designed to assess whether individuals recognized the opportunity to integrate separate yet related facts (see Appendix A). This survey was added to the protocol midway through data collection, after several participants spontaneously commented on the relational structure of the task. As a consequence, self-reported awareness was only assessed for the final 78 participants tested.

To reiterate, though participants were told that their memory would be tested, at no time were they informed that any of the sentences were related. Moreover, the same counterbalancing scheme from Experiment 2 was employed. Thus, each fact was presented in a lag of 2, 3, and 4 approximately equally often. Similarly, fact order was also counterbalanced with each fact from a target pair appearing equally often in the first or second serial position (i.e., Stem 1 vs. Stem 2).

Session 2. Participants returned to the laboratory approximately one week after their initial visit. After completion of a number of unrelated tasks, memory for the integration facts was assessed. First, participants were tested for recall of the integration facts via the same open-ended questions asked during Session 1, though in a different order than the questions appeared at the first visit as to reduce practice effects. The facts were presented on a laptop using PowerPoint® software; reaction time was not recorded. Unlike at Session 1, the facts appeared as full sentences rather than one word at a time. Following the open-ended portion, participants were asked forced-choice questions for any integration questions that were answered incorrectly in the same manner as in Experiment 2.

Results

Knowledge Extension and Retention

At Session 1, participants self-generated the novel integration facts on 50% of the trials (*M* Proportion score = 0.50; *SD* = 0.21). As depicted in Figure 5 (Panel A), substantial individual differences were observed with performance ranging from 3% to 93% correct across the sample. Moreover, as reflected in Figure 6 (Panel A), successful performance did not differ as a function of lag, $F(2, 232) = 0.66$, $p = .52$. Mean response time necessary to self-generate the integration facts also did not differ by lag, $F(2, 212) = 0.09$, $p = 0.92$ (Figure 6, Panel B). These results represent a replication of Experiment 2, with a substantially larger sample.

The primary question of interest was whether young adults successfully retained the newly self-generated knowledge over the 1-week delay. At Session 2, when tested in an open-ended format, participants recalled 42% of the integration facts (*M* Proportion score = 0.42; *SD* = 0.21). Despite high levels of recall and similar patterns of variability to that observed at Session 1 (Figure 5, Panel B), a significant loss of information was observed between sessions, $t(114) = 10.04$, $p < .001$, $d = 0.53$. Nevertheless, for the facts that were not successfully recalled, participants selected the correct answer on 51% of the forced-choice trials (*M* Proportion score = 0.51; *SD* = 0.15), which significantly differed from chance (25%), $t(114) = 18.98$, $p < .001$.

Individual Differences in Initial Self-generation Performance

To further investigate variability in initial self-generation performance, knowledge extension through integration was additionally examined as a function of the amount of time participants took to generate the novel integration fact. As depicted in Table 1, a dependent samples *t*-test indicated that participants were significantly faster to respond on successful versus unsuccessful trials, $t(106) = 12.89$, $p < .001$, $d = 1.48$. Moreover, self-generation performance and reaction time were significantly negatively correlated, $r(105) = -.24$, $p = .01$. That is, high performers generated the novel integration fact more quickly than low performers. An interesting pattern emerged when this relation was examined separately for successful versus unsuccessful trials. On successful trials, a similar negative correlation between reaction time and self-generation performance was found, $r(105) = -.23$, $p = .02$. Interestingly, however, examination of unsuccessful trials revealed a marginally significant positive correlation between reaction time and self-generation performance, $r(105) = .18$, $p = .06$. Thus, high-performing participants not only responded faster on successful trials, but also were marginally slower on unsuccessful trials.

We next addressed whether awareness of the opportunity to integrate was related to variability in self-generation performance. Of the 78 participants who completed the debriefing survey assessing participants' perceptions of and strategic approaches to the study, 62% ($N = 48$) reported explicit awareness that some of the facts were related. Moreover, explicit awareness was significantly correlated with the proportion of integration facts successfully self-generated at Session 1, $r_s(71) = .44$, $p < .001$. Despite the relation between explicit awareness and successful performance, explicit awareness was not correlated with the amount of time it took participants to generate a response to the total corpus of open-ended integration questions, r_s (71) = .16, p = .17, nor with the mean reaction time on successful trials, r_s (71) = .03, p = .84. Interestingly, however, a significant positive correlation was found between explicit awareness and mean reaction time on unsuccessful trials, r_s (71) = .33, p = .005, such that participants who were aware of the opportunity to integrate spent longer on trials in which they were unsuccessful.

Lastly, it is possible that the presence of lag effects would be contingent upon whether participants were explicitly aware of the ability to integrate. Nevertheless, when ANOVAs were conducted to examine self-generation by lag separately for participants who were and were not aware of the relational structure of the task, no significant main effects of lag were observed, F(2, 94) = 0.63, p = 0.54.

Prior Knowledge of the Stem Facts

Although we could not assess participants' prior knowledge of the integration facts at the initial time of learning (due to the subsequent test for retention), we were able to examine participants' self-reported prior knowledge of the individual stem facts. On average, 21% of the facts were identified as known (M = 12.04; SD = 7.74). The total number of previously known facts did not significantly correlate with self-generation of the novel integration facts, r(111) = .15, p = .11. The same pattern was observed when the relation was examined separately for Stem 1 facts, r(111) = .13, p = .19, and Stem 2 facts, r(111) = .14, p = .14.

Discussion

The present experiment replicated and extended the findings reported in Experiment 2. That is, young adults extended new knowledge under more challenging testing conditions (open-ended versus forced-choice) and exhibited striking variability in self-generation of new knowledge through memory integration. Although newly self-generated knowledge was significantly less accessible following a one-week delay, participants still recalled 42% of the novel integration facts. Moreover, of the facts that participants failed to recall in an open-ended form, 51% were successfully identified under more supportive forced-choice testing conditions. Therefore, information self-generated through integration persisted in memory over time. Consistent with results from Experiment 1, the degree of lag between to-be-integrated stem facts had no impact on subsequent knowledge extension, even when controlling for the potential effects of explicit awareness.

Examination of relations between initial knowledge extension and additional individual difference measures is potentially revealing with respect to the underlying processes involved. High-performing individuals were faster on correct trials and marginally slower on incorrect trials, suggesting that they may have attempted to execute a strategy during the test phase, and might have persisted in instances in which that strategy returned no results, namely, on unsuccessful trials. Support for this conclusion comes from the finding that explicit awareness of the opportunity to integrate was strongly correlated with the amount of time spent on unsuccessful trials, but not on successful trials. Therefore, it is possible that participants who were more aware of the task structure spent significantly longer on trials in which they were ultimately unsuccessful as they persisted with attempts to identify relevant related material, whereas, more careful encoding during the learning phase might have led to quicker access of knowledge needed for self-generation on successful trials.

General Discussion

The present research was an investigation of self-generation and retention of new factual knowledge through integration of separate yet related episodes of new learning. We extended prior research by elucidating the factors and conditions that impact knowledge extension through memory integration at both the group and individual levels in young adults. The findings indicated that temporal spacing between related episodes had no impact on knowledge extension through integration, at least within a single learning session. We also examined the extent to which adults were successful at extending knowledge through more challenging means than have previously been examined, namely, in open-ended as opposed to forced-choice testing. Adults successfully self-generated integrated knowledge under these more challenging conditions. Moreover, although some loss of information was observed over time, integrated knowledge remained highly accessible after a 1-week delay, thereby providing direct evidence for the role of memory integration in the long-term accumulation of knowledge. Finally, striking individual differences were evident in the extent to which adults successfully self-generated new knowledge through integration, which was strongly related to whether individuals spontaneously identified the relational structure of the learning task. The theoretical implications of these findings are discussed below.

The current research sheds light on our theoretical understanding of the nature of memory integration, at least as it has been investigated in the adult literature. Prior research with adults makes clear that knowledge extension through integration is accomplished through retrieval-mediated learning, whereby prior event details are reinstated during encoding of related experiences (Zeithamova et al., 2012a). Moreover, findings from the developmental literature indicate that knowledge extension through memory integration is more prevalent under conditions that promote simultaneous activation of separate but related episodes to one another, such as when to-be-integrated episodes share many features in common (Bauer et al., 2012) or when separate episodes are distributed closely in time rather than across 1-week delays (Varga & Bauer, 2013). Interestingly, however, temporal spacing between related episodes had no impact on knowledge extension through memory integration in the present research. That is, in Experiment 1, there were virtually no differences in either accuracy or response speed between knowledge derived through integration of stem facts separated by an average of 5 intervening facts as compared to stem facts separated by an average of 60 intervening facts. The finding that the degree of temporal distance between separate yet related episodes had no impact on the capacity to integrate across them is particularly interesting in light of evidence that temporal spacing produces differential performance in other forms of knowledge extension. For instance, inductive generalization is facilitated when instances of a category are interleaved (as opposed to massed) within a single study session (Kornell & Bjork, 2008; Kornell, Castel, Eich, & Bjork, 2010; Vlach, Sandhofer, & Kornell, 2008). Thus, future research should examine whether knowledge extension through integration is sensitive to other traditional episodic manipulations, such as surface similarity between related stem facts (e.g., Bauer et al., 2012) or recency judgments of the temporal order of studied items (e.g., Milner, 1971; Petrides, 1991).

The present experiments also take an important step toward furthering our understanding of the long-term retention of self-derived knowledge. Recent investigations with adults have made great strides in characterizing knowledge extension through integration within a *single* experimental session. Yet in everyday learning contexts, delays between initial learning and later use are commonly encountered. Although many researchers have acknowledged the growing need to examine self-derivative processes under conditions that better mirror everyday learning conditions (e.g., Gentner & Smith, 2012; Jee et al., 2010), the long-term accessibility of self-derived knowledge has received little attention. Nevertheless, if memory integration serves as a pervasive process underlying knowledge development, it is important to test whether the products of knowledge extension persist in the knowledge base over time. Consistent with findings from 4- and 6-year-olds (Varga et al., 2016; Varga & Bauer, 2013), Experiment 3 of the present research indicated that young adults retained knowledge newly derived through memory integration. That is, at Session 1 individuals self-generated 50% of the integration facts. When tested for retention one week later, 42% of the facts were recalled in an open-ended form; an additional 51% of the remaining facts were successfully accessed when individuals were provided with additional support in the form of forced-choice cues. Relatively high accessibility following a considerable delay suggests that knowledge newly self-generated through integration had been incorporated into the knowledge base. This conclusion is further substantiated by results from the previously discussed study by Bauer and Jackson (2015). When event-related potentials (ERPs) were recorded while participants read well-known facts, novel facts, and facts derived through integration of the previously encoded stem facts, neural responses to well-known and integration facts differed from novel facts, but well-known and integration facts did not differ from each other, indicating that integrated knowledge is rapidly incorporated into semantic memory. In light of comparable patterns of incorporation in young adults and retention in young children, the results from the present research provide direct evidence that memory integration serves as a key mechanism underlying the long-term accumulation of semantic knowledge in adults (see Bauer & Varga, 2016, for further discussion).

Despite high levels of retention, it is also worth noting that there was significant loss between sessions in Experiment 3. Nevertheless, this finding is consistent with prior research. In the only analogous investigation of long-term retention of integrated knowledge conducted with adults to date, Sweegers and colleagues (2014) found that information exhibiting stable, cross-episode relations was preferentially consolidated in memory as compared to isolated memories, yet even these integrated representations exhibited significant degradation over the course of 48 hours. Although long-term retention was assessed only for directly learned associations that contained integrated regularities, rather than knowledge that was newly self-generated through memory integration, significant loss was similarly demonstrated albeit over a much shorter delay. Indeed, diminished retention in the face of a 1-week delay in the present research is not altogether surprising. In fact, it is rather remarkable that undergraduate students retained as much knowledge as they did, especially given that they acquired the novel integration facts through single-trial procedures and likely encountered a wealth of additional information in the classes they attended over the course of the retention interval. Notwithstanding, because significant loss was still apparent, additional research aimed at promoting the long-term accessibility of factual knowledge derived through memory integration is warranted.

The current research also advances our understanding of potential factors underlying individual differences in the capacity to extend new knowledge through integration. Indeed, the present experiments constitute the first demonstration of striking individual differences in successful *self-generation* of new knowledge in adults. As suggested by Barsalou and Prinz (1997), individuals who exhibit exceptional self-generative learning abilities (i.e., “exceptional creativity”) may perceive subtle structural relations in the world that others do not, allowing them to integrate information and generate novel combinations that others would not naturally consider. In accord with this predication, 62% of the individuals in Experiment 3 possessed explicit awareness of the opportunity to integrate; whereas, 38% of the sample made no mention of the possibility to do so. What is more, perception of the structural relations between to-be-integrated facts was strongly associated with self-generation. This finding is particularly interesting in light of conflicting findings from prior research. For instance, in one of the first demonstrations of extensive variability in knowledge extension through integration, Shohamy and Wagner (2008) found that explicit awareness of the task structure was virtually non-existent. A precise explanation for why differential effects of explicit awareness are observed across different paradigms is beyond the scope of the present research. However, as discussed previously, existing paradigms employed in the adult literature test memory integration in one of two ways. Whereas Shohamy and Wagner (2008) elicited integration through trial-and-error, reinforcement training, participants in the present research were required to integrate via single-trial observation. In addition to differences in learning procedures, the nature of the stimuli also might have accounted for differential findings. Whereas participants integrated arbitrary stimulus pairs in Shohamy and Wagner (2008), participants in the present research integrated real-world factual knowledge. Although the current research cannot directly speak to these alternative accounts, the present results provide initial support for the conclusion that striking variability in the capacity to derive new factual knowledge through memory integration is associated with the ability to perceive subtle structural relations present in the environment. Thus, in an effort to better explain variability in what is assumed to be a fundamental learning mechanism, future research is needed to delineate the role of domain-general cognitive processes in supporting initial identification of the opportunity to integrate and further extend that information.

The current research also provides some clues into the time-course of knowledge extension through integration. An ongoing debate in the literature concerns whether integrated representations are directly encoded into memory at the time of initial learning (i.e., integrative encoding; see Shohamy & Wagner, 2008) or whether these relations are only established when discrete memories are retrieved and recombined in response to a demand to do so at test (i.e., retrieval-based generalization; see Kumaran, 2012; Kumaran & McClelland, 2012). Conflict arises from the fact that a hippocampally-mediated integrative encoding signature has been observed in several studies in which no retrieval-based processing was evidenced while participants later made inferential judgments (e.g., Shohamy & Wagner, 2008; Zeithamova et al., 2012a). However, the opposite pattern has also been observed in a separate line of research (e.g., Heckers et al., 2004; Preston et al., 2004; Zeithamova & Preston, 2010). As proposed by Zeithamova and colleagues (2012b), it is possible that the relative contribution of encoding and retrieval-based mechanisms is determined by the demands of a particular task. That is to say, integrative encoding is likely sufficient in situations in which individuals are repeatedly exposed to associations, thereby allowing for on-line generalization at the time of learning. On the other hand, in cases in which knowledge extension is contingent on single-trial learning procedures, integrative encoding is probably necessary but not sufficient for explicit knowledge extension. Consistent with this notion, in the child analogue of the paradigm employed in the current research, Varga and Bauer (2013) inserted delays throughout the learning process and found that while integrative encoding mechanisms certainly affect the capacity for subsequent knowledge extension, explicit self-generation of the novel integration fact occurs at the time of test. Similarly, eye-tracking data from 8-year-olds indicates that cognitive effort while reading Stem 2 of a pair of related facts predicts subsequent recall of integration facts derived from them (Bauer et al., in press). It is therefore reasonable to suggest that related stem facts may be integrated in advance of a test probe prompting explicit knowledge extension.

Although the present research cannot directly address the mechanisms involved during encoding and test, differential response latencies during the test for knowledge extension are potentially revealing with respect to the time-course of knowledge extension through integration. First, high-performing individuals were faster on trials in which they subsequently self-generated the novel integration fact. One possible explanation for this pattern of results is that high-performing individuals engaged in integrative encoding during initial learning, thereby leading to faster responses when presented with a demand to further extend that knowledge at test. In contrast, poor performers might have engaged in integrative encoding to a lesser extent, thus requiring retrieval, integration, and further extension of discretely stored stem facts at the time of test, which could therefore explain both slower response times and less success overall. Second, high performers were nominally slower on trials in which they failed to self-generate, which was also related to heightened awareness of the task structure. Building on the previous explanation, it is plausible that highly successful participants spent significantly longer on unsuccessful trials because they were aware of the possibility to link facts in order to generate new understandings. However, if they failed to integrate during initial encoding, self-generation was still likely to be unsuccessful. Notwithstanding, in light of knowledge about the strategy to employ, high-performers appeared to persevere for longer in comparison to low-performers.

In conclusion, the present experiments took important steps toward understanding the long-term retention of self-derived knowledge, as well as some of the factors that promote the likelihood of successfully self-generating new knowledge through integration. In addition to contributing to our theoretical understanding of knowledge extension and retention, the findings also have implications for the promotion of knowledge development and for educational practice more broadly (see Bauer, 2012; Bauer & Varga, 2016, for discussion). The results reported here demonstrate that even among young adults, there is substantial variability in performance that depends on integrating and further extending new knowledge. Moreover, based on the beneficial effects of explicit awareness, the findings highlight the potential effectiveness of cues that enable students to see the connection between related information and which encourage individuals to flexibly extend their knowledge. In an effort to design interventions aimed at promoting this fundamental learning ability, it will be necessary to elucidate the factors that contribute to the ability to detect relational similarities spontaneously and to scaffold such skills. Moreover, because temporal spacing had no impact on knowledge extension in the present research, to further understand the nature of this learning phenomenon, future research must determine how knowledge extension through integration operates under other traditional episodic manipulations.

References

Barsalou, L. W., & Prinz, J. J. (1997). Mundane creativity in perceptual symbol systems. In T. B. Ward, S. M. Smith, & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 267-307). Washington, DC: American Psychological Association.

Bauer, P. J. (2012). The life I once remembered: The waxing and waning of early memories. In D. Bensten & D. C. Rubin (Eds.), *Understanding autobiographical memory: theories and approaches* (pp. 205-225). Cambridge, UK: Cambridge University Press.

Bauer, P. J., Blue, S. N., Xu, A., Esposito, A. G. (in press). Productive extension of semantic memory in school-aged children: Relations with reading comprehension and deployment of cognitive resources. *Developmental Psychology.*

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), 271-282. doi: http://dx.doi.org/10.1037/a0037405

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., & Larkina, M. (in press). Realizing relevance: The influence of domain-specific information on generation of new knowledge through integration in 4- to 8-year-old children. *Child Development.*

Bauer, P. J., & San Souci, P. (2010). Going beyond the facts: Young children extend knowledge by integrating episodes. *Journal of Experimental Psychology*, 107(4), 452-465.

Bauer, P. J., Varga, N. L. (2016). The developmental cognitive neuroscience of memory: Implications for education. In É. Tardif & P. Doudin (Eds.), Collective work on the topics of neuroscience, cognition and education. Oxford, UK: De Boeck.

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. doi: 10.1080/15248372.2013.849707

Bunsey, M., & Eichenbaum, H. (1996). Conservation of hippocampal memory function in rats and humans. *Nature*, *379*(6562), 255-257.

Dusek, J. A., & Eichenbaum, H. (1997). The hippocampus and memory for orderly stimulus relations. *Proceedings of the National Academy of Sciences*, *94*(13), 7109-7114.

Gentner, D., & Smith, L. (2012). Analogical reasoning. In V. S. Ramachandran (Ed.), *Encyclopedia of Human Behavior: 2nd Ed.* (pp. 130-136). Oxford, UK: Elsevier.

Haist, F., shimamura, A. P., & Squire, L. R. (1992). On the relationship between recall and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(4), 691-702. doi: http://dx.doi.org/10.1037/0278-7393.18.4.691

Heckers, S., Zalesak, M., Weiss, A. P., Ditman, T., & Titone, D. (2004). Hippocampal activation during transitive inference in humans. *Hippocampus*, *14*(2), 153-162.

Jee, B. D., Uttal, D. H., Gentner, D., Manduca, C., Shipley, T. F., Tikoff, B., … Sageman, B. (2010). Commentary: Analogical thinking in geoscience education. *Journal of Geoscience Education, 58*(1). 2-13. doi: http://dx.doi.org/10.5408/1.3544291

Kahana, M. J., & Howard, M. W. (2005). Spacing and lag effects in free recall of pure lists. *Psychonomic Bulletin & Review*, *12*(1), 159-164.

Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories is spacing the “enemy of induction”?. *Psychological science*, *19*(6), 585-592.

Kornell, N., Castel, A. D., Eich, T. S., & Bjork, R. A. (2010). Spacing as the friend of both memory and induction in young and older adults. *Psychology and Aging*, *25*(2), 498-503.

Kumaran, D. (2012). What representations and computations underpin the contribution of the hippocampus to generalization and inference. *Frontiers in Human Neuroscience, 6*, 1-11.

Kumaran, D., & McClelland, J. L. (2012). Generalization through the recurrent interaction of episodic memories: a model of the hippocampal system. *Psychological review, 119*(3), 573-616.

Kumaran, D., Summerfield, J. J., Hassabis, D., & Maguire, E. A. (2009). Tracking the emergence of conceptual knowledge during human decision making. *Neuron, 63*, 889-901. DOI: 10.1016/j.neuron.2009.07.030

Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. *British Medical Bulletin, 27*(3), 272-277.

Petrides, M. (1991). Functional specialization within the dorsolateral frontal cortex for serial order memory. *Proceedings of the Royal Society of London B: Biological Sciences*, *246*(1317), 299-306.

Preston, A. R., & Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. *Current Biology*, *23*(17), R764-R773.

Preston, A. R., Shrager, Y., Dudukovic, N. M., & Gabrieli, J. D. (2004). Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus, 14*(2), 148-152. DOI 10.1002/hipo.20009

Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N., & Szpunar, K. K. (2012). The future of memory: remembering, imagining, and the brain. *Neuron*, *76*(4), 677-694.

Schlichting, M. L., & Preston, A. R. (2015). Memory integration: Neural mechanisms and implications for behavior. *Current opinion in behavioral sciences*, *1*, 1-8.

Schlichting, M. L., Zeithamova, D., & Preston, A. R. (2014). CA1 subfield contributions to memory integration and inference. *Hippocampus*, *24*(10), 1248-1260.

Shohamy, D., & Wagner, A. D. (2008). Integrating memories in the human brain: Hippocampal-midbrain encoding of overlapping events. *Neuron, 60*(2), 378-389. doi:10.1016/j.neuron.2008.09.023

Siegler, R. S. (1989). Mechanisms of cognitive development. *Annual Review of Psychology*, *40*(1), 353-379.

Squire, L. R., Wixted, J. T., & Clark, R. E. (2007). Recognition memory and the medial temporal lobe: A new perspective. *Nature Reviews Neuroscience, 8*(11), 872-883. doi: 10.1038/nrn2154

Sweegers, C. C. G., Takashima, A., Fernández, G., Talamini, L. (2014). Neural mechanisms supporting the extraction of general knowledge across episodic memory. *NeuroImage, 87*, 138-146.

Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., ... & Morris, R. G. (2007). Schemas and memory consolidation. *Science*, *316*(5821), 76-82.

Tse, D., Takeuchi, T., Kakeyama, M., Kajii, Y., Okuno, H., Tohyama, C., ... & Morris, R. G. (2011). Schema-dependent gene activation and memory encoding in neocortex. *Science*, *333*(6044), 891-895.

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. (2014). Conceptual knowledge extension: An examination of its development and the underlying cognitive processes involved. In R. Chen (Ed.), *Cognitive development: Theories, stages & processes and challenges* (pp. 1-16). Huntington, NY: Nova Science Publishers.

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. doi: 10.1016/j.jecp.2015.11.015

Vlach, H. A., Sandhofer, C. M., & Kornell, N. (2008). The spacing effect in children's memory and category induction. *Cognition*, *109*(1), 163-167.

Zeithamova, D., Dominick, A. L., & Preston, A. R. (2012a). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron, 75*(1), 168-179.

Zeithamova, D., & Preston, A. R. (2010). Flexible memories: Differential roles for medial temporal lobe and prefrontal cortex in cross-episode binding. *The Journal of Neuroscience, 30*(44), 14676-14684.

Zeithamova, D., Schlichting, M. L., & Preston, A. R. (2012b). The hippocampus and inferential reasoning: building memories to navigate future decisions. *Frontiers in Human Neuroscience*, *6*, 1-14.

Appendix A

Impressions of the Fact-Learning Task

1. Did you notice anything interesting or odd about the facts that you learned in the first portion?

2. What (if any) strategies did you use to learn the facts in the first portion?

3. What (if any) strategies did you use to answer the questions in the second portion?

Thank you for your time!

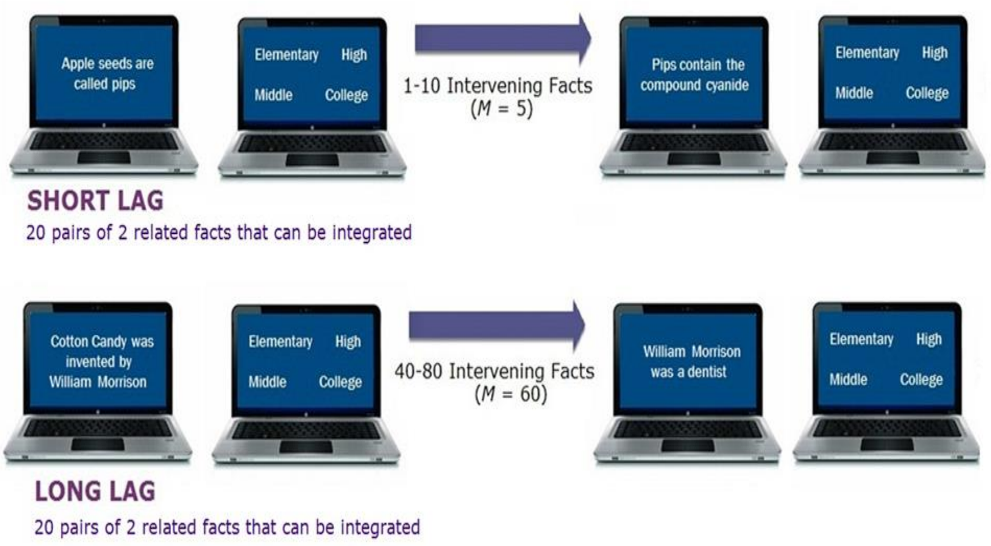
Table 1.

Mean Response Time (ms) on Successful and Unsuccessful Trials in Experiment 3

	M	*SD*
Successful	4540.91	2184.98
Unsuccessful	13456.74	8211.76
Total	8609.40	4738.03

Figure 1

Panel A: Encoding Phase



Panel B: Test Phase

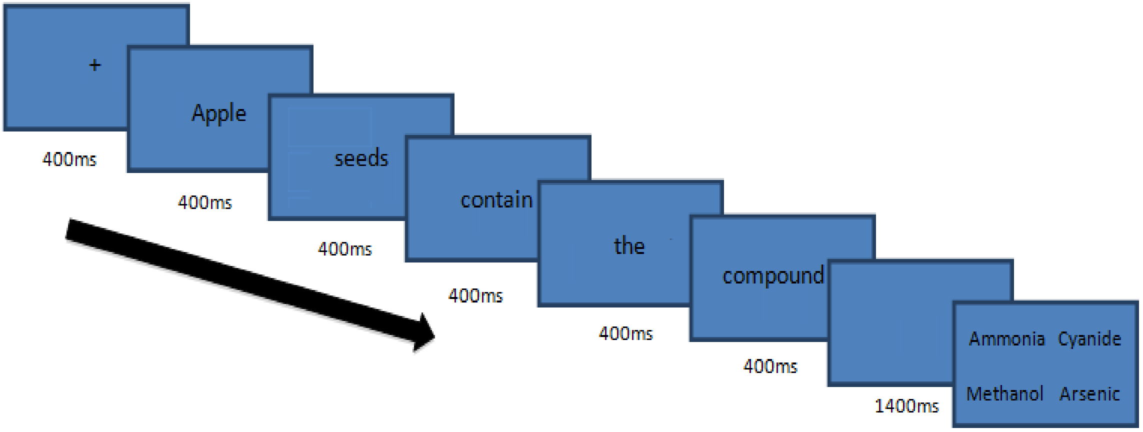
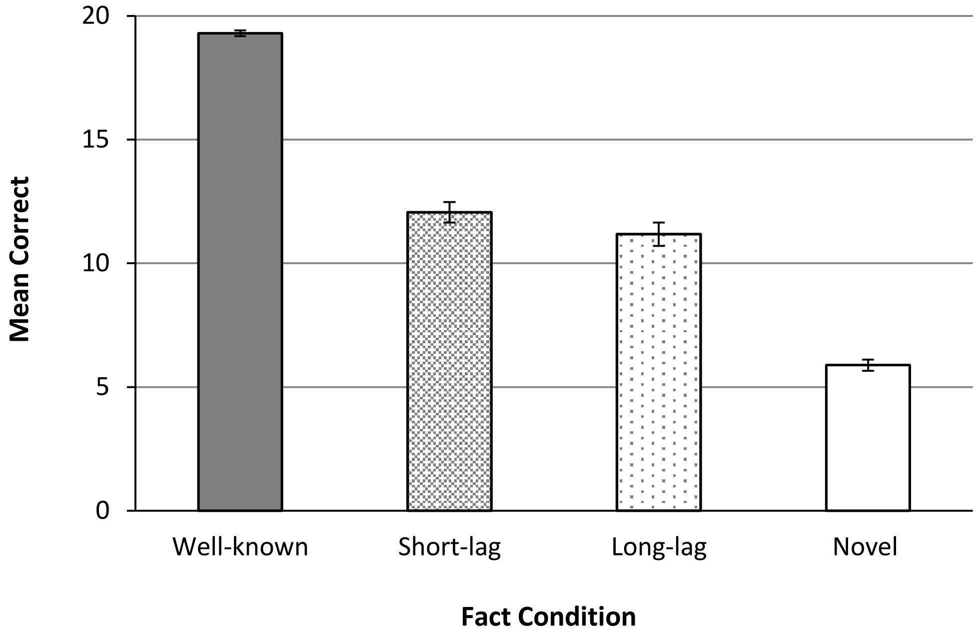


Figure 1. Schematic of encoding phase (Panel A) and test phase (Panel B) procedures employed in Experiment 1. Response time was recorded from the onset of the decision screen during test.

Figure 2

Panel A: Accuracy



Panel B: Reaction Time

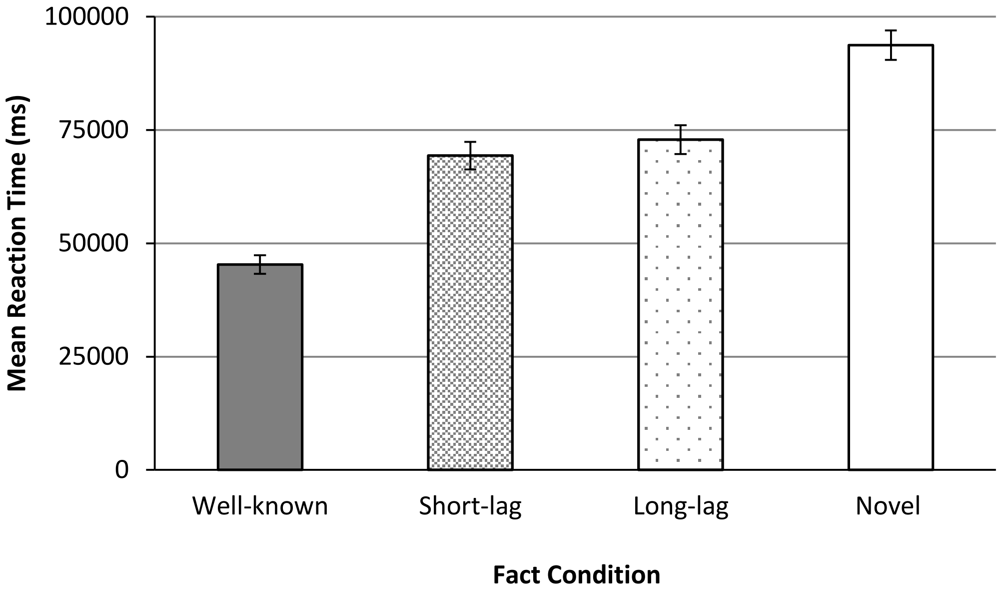
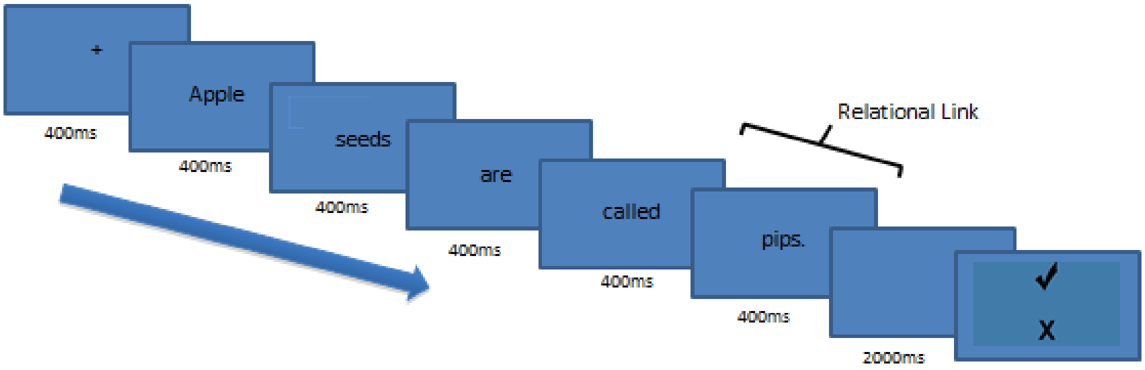


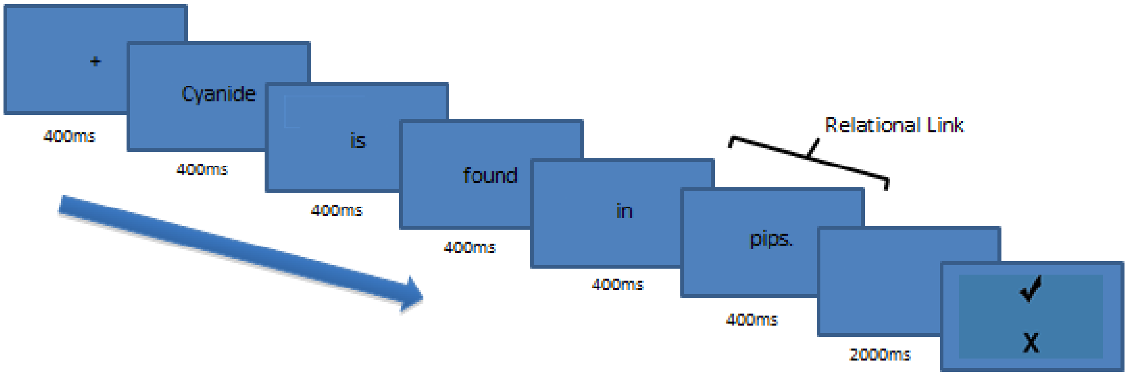
Figure 2. Mean accuracy (Panel A) and mean reaction time (Panel B) across conditions in Experiment 1. Error bars represent standard error of the mean.

Figure 3

Panel A: Encoding of Stem Fact 1



Panel B: Encoding of Stem Fact 2



Panel C: Test for Self-generation

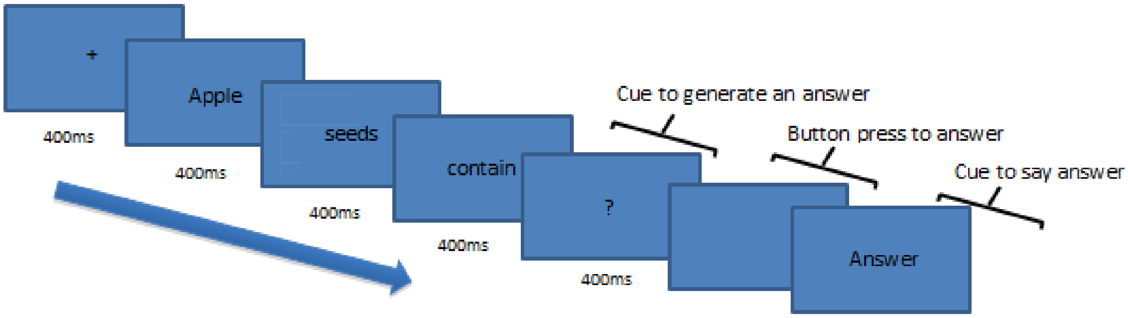
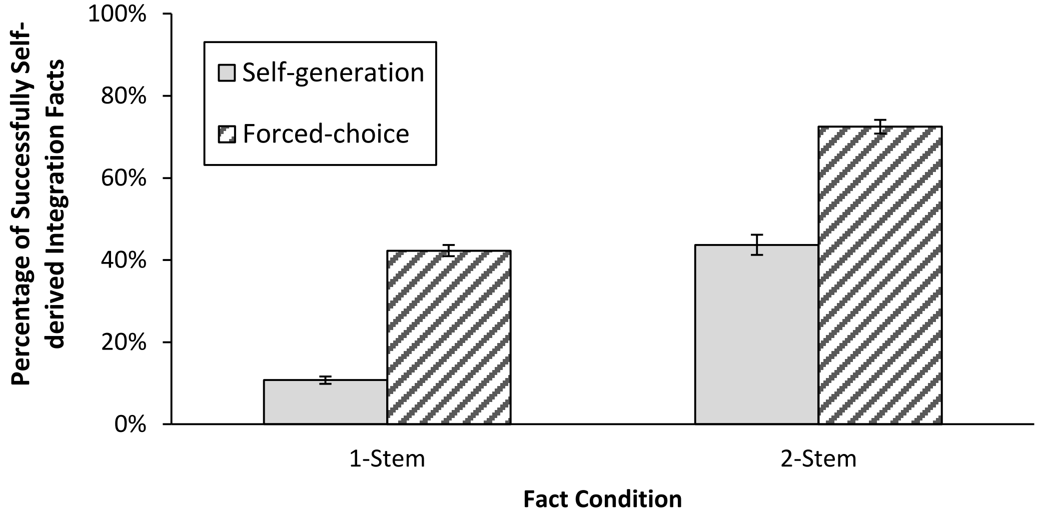


Figure 3. Schematic of encoding (Panels A and B) and test phase (Panel C) procedures in Experiments 2 and 3. Reaction time was recorded from the onset of “?” in Experiment 3.

Figure 4

Panel A: Performance across Conditions



Panel B: Variability across Participants

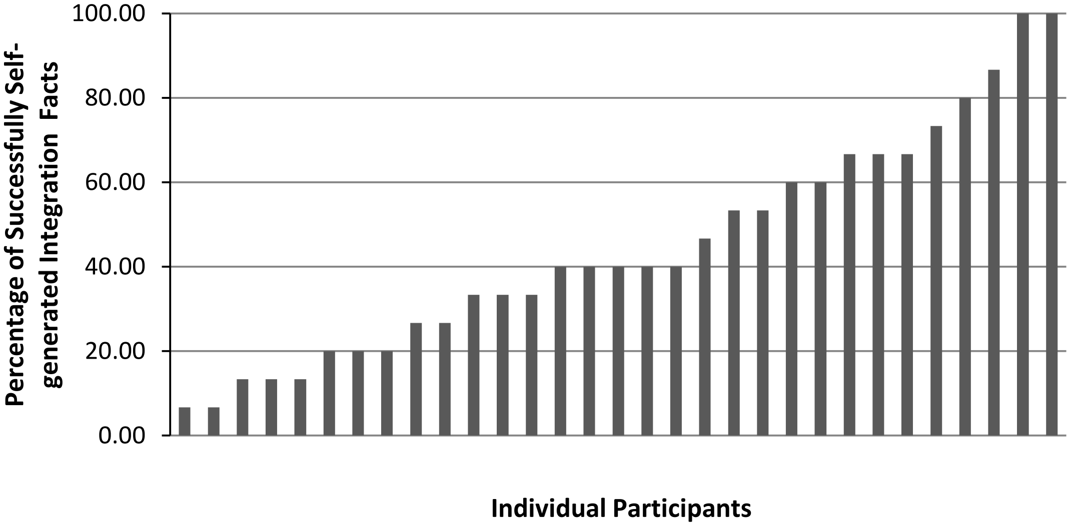
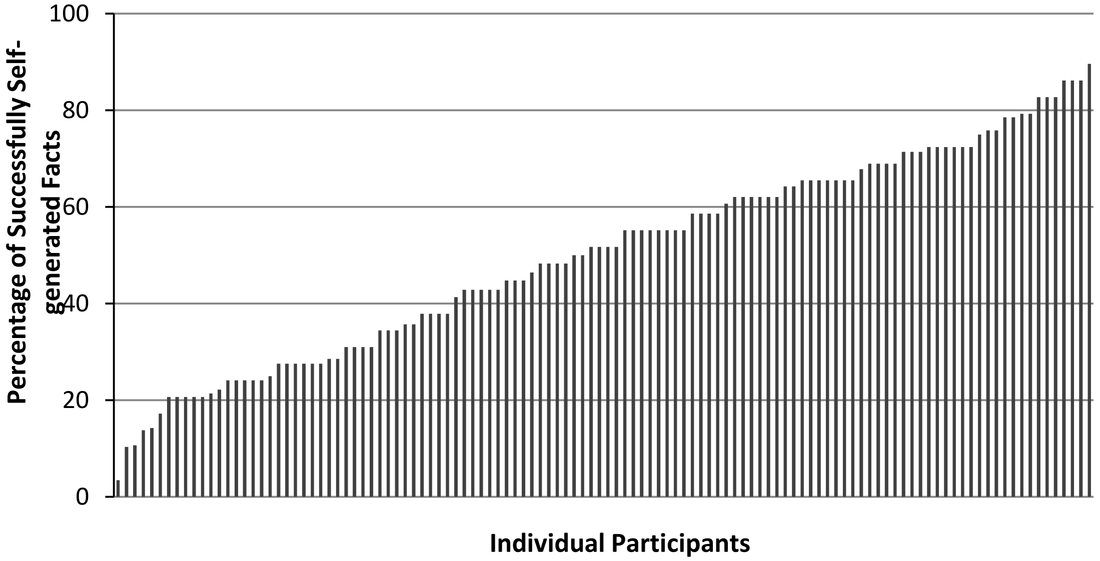


Figure 4. Percentage of trials on which the novel fact was successfully derived in a 1-stem versus a 2-stem condition (Panel A) as well as self-generated by individual participants in the 2-stem condition (Panel B) in Experiment 2. Error bars represent standard error of the mean.

Figure 5

Panel A: Session 1



Panel B: Session 2

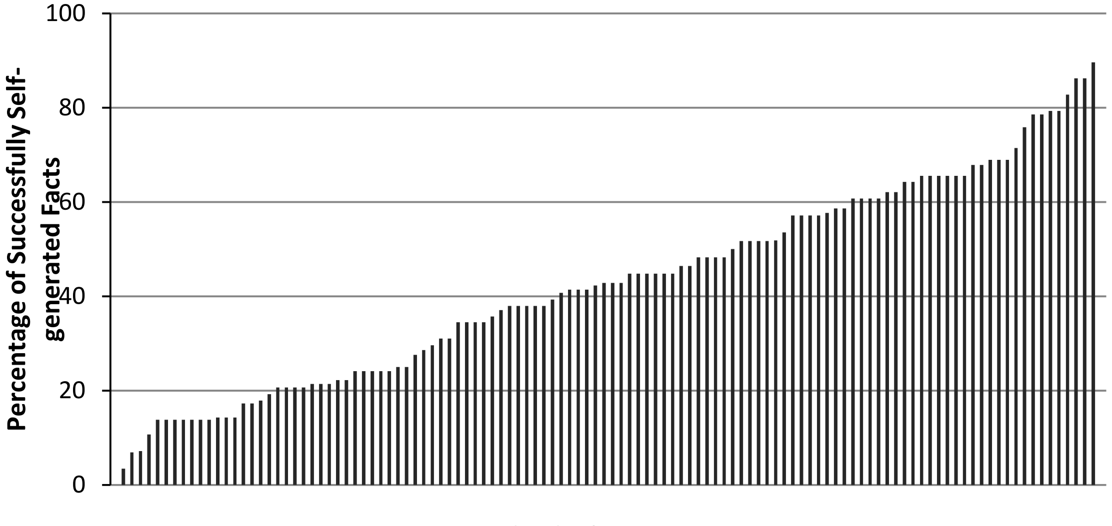
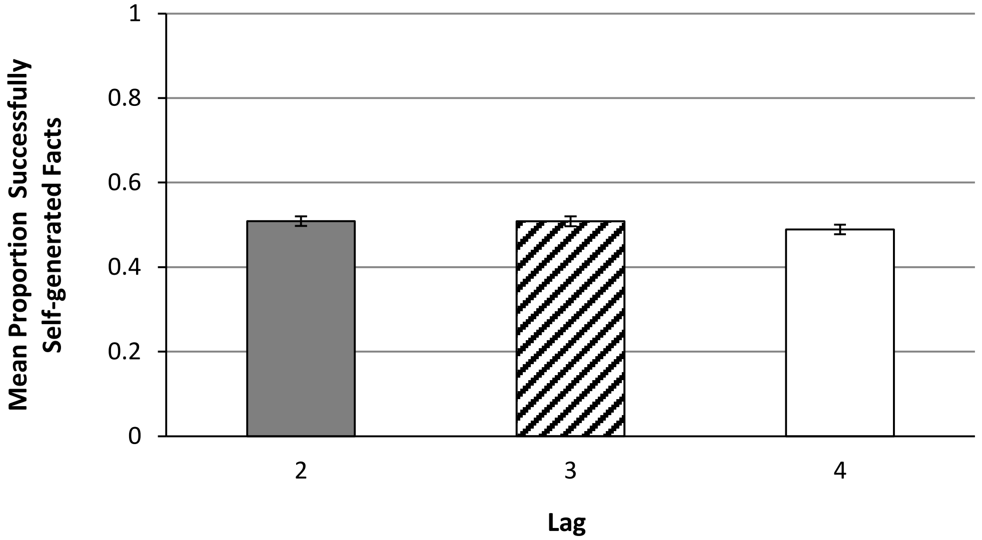


Figure 5. Mean percentage of successfully self-generated integration facts among participants in Experiment 3 at Session 1 (Panel A) and Session 2 (Panel B). The X-axis shows individual participant performance arranged from lowest to highest.

Figure 6

Panel A: Self-generation Performance



Panel B: Reaction Time

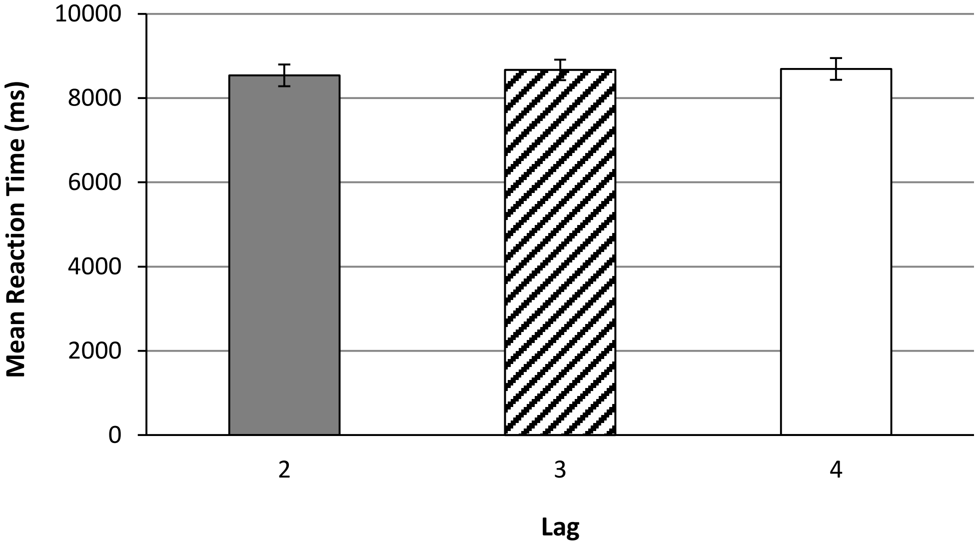


Figure 6. Mean proportion successfully self-generated integration facts (Panel A) and mean response time (Panel B) by lag in Experiment 3. Error bars represent standard error of the mean.

Study 2

Using ERPs to dissociate the neurocognitive processes underlying knowledge extension through memory integration

Nicole L. Varga & Patricia J. Bauer

Emory University

Using ERPs to dissociate the neurocognitive processes underlying knowledge extension through memory integration

The semantic memory system represents the knowledge we accrue across experience, organizes this information categorically, and provides accessible representational support for all our cognitive functions. Depending on the field of study, the terms “semantic memory,” “conceptual system,” and “knowledge representation” are used interchangeably. Despite differences in vocabulary, researchers across psychology, cognitive science, and cognitive neuroscience use these terms to refer to the same interconnected network of factual information that constitutes our knowledge of the world, thereby signifying the widespread view that our knowledge is represented as organized networks. A key mechanism through which an organized knowledge base is formed, and by which the further extension of existing knowledge is accomplished, is memory integration—the combination of separate yet related traces of information (Bauer & San Souci, 2010). There has been substantial investigation of the brain areas underlying memory integration in recent years, resulting in gains in our understanding of the processes that might support this behavior in adults. More specifically, studies employing functional magnetic resonance imaging (fMRI) have shown that a hippocampal-medial prefrontal (mPFC) circuit is involved (e.g., Zeithamova, Dominick, & Preston, 2012a). Nevertheless, because neuroimaging methods cannot easily distinguish separate stages of processing, the precise mechanisms by which these regions contribute to memory integration are not fully understood. In the current research we capitalized on the temporal resolution of event-related potentials (ERPs) to investigate the time-course of neurocognitive processing responsible for successful knowledge extension through memory integration in young adults.

Recent findings from behavioral (Bauer, Varga, King, Nolen, & White, 2015; Varga & Bauer, 2013) and neuroimaging research (Zeithamova, et al., 2012a; Zeithamova & Preston, 2010) provide converging evidence that self-generation of new knowledge through memory integration is not a unitary mechanism that is set into motion as a result of experience of related learning episodes. Instead, it is better characterized by separate encoding and retrieval processes: integration of related episodes during encoding (O'Reilly & Rudy, 2001; Shohamy & Wagner, 2008) followed by interrogation of those episodes of experience for their relevance to a specific demand to generate new knowledge (Eichenbaum & Cohen, 2001). In adults, one commonly employed approach to assessing knowledge extension through integration involves training individuals to associate overlapping stimulus pairs (e.g., *AB: Chair & Zucchini; BC: Zucchini & Blender*) in order to infer a novel, indirect relation (e.g., *AC: Chair & Blender*) (Preston, Shrager, Dudukovic, & Gabrieli, 2004; Schlichting, Zeithamova, & Preston, 2014, Zeithamova et al., 2012a; Zeithamova & Preston, 2010). Critically, arbitrary stimulus pairs are temporally distributed throughout encoding, thereby necessitating integration of newly and previously learned information in memory. Success is then measured through a forced-choice transfer test, requiring manipulation of previously learned information in response to a specific retrieval cue (e.g., *AC*: *Chair = Blender or House?*). The use of arbitrary stimulus pairs provides strong experimental control for the effects of prior knowledge. Yet it is unlikely that the newly derived associations are incorporated into one's long-term knowledge base, thus limiting the light the paradigm can shed on the accrual of real-world factual knowledge.

The method adopted in the present research was developed in response to the growing need to examine the processes supporting knowledge acquisition under more ecologically valid conditions (e.g., Gentner & Smith, 2012; Jee et al., 2010). Inspired by a paradigm recently employed by Bauer and Jackson (2015), individuals were taught true but novel "stem" facts that were related to one another (*Apple seeds are called pips; Cyanide is found in pips*). They later were tested for self-derivation of new knowledge through integration of the target information (*Apple seeds contain cyanide*). Compared to other inference paradigms, this approach has the added value of being about real-world information, and thus is directly relevant to the issue of how a semantic knowledge base is built over time. Moreover, consistent with the proposal that memory integration supports the *organization* of knowledge within semantic memory (e.g., Preston & Eichenbaum, 2013), prior research employing this paradigm provides evidence that integrated knowledge is retained in memory, at least over a period of 1 week (Varga and Bauer, under review). Based on this finding, in the present investigation, we examined ERPs while individuals encoded related information at the time of initial learning and while they were later tested for self-generation of new knowledge via an explicit test. The ERP data were collected in conjunction with behavioral measures reported in Varga and Bauer (under review); the ERP findings have not been previously reported. Findings from the ERP data stand to inform our understanding of the distinct neurocognitive processes involved in the formation and further extension of an integrated knowledge base, and their time-course.

The specific processes involved in knowledge extension through memory integration are under active debate. However, traditional explanations share one feature in common: they agree that it is invoked under conditions that permit pattern completion, in which exposure to a novel item causes reactivation of memories for prior episodes (Eichenbaum & Cohen, 2001; O'Reilly & Rudy, 2001). A framework that has emerged to account for *integration* of separate yet related information in memory, in particular, is retrieval-mediated learning (Hall, 1996; Holland, 1981), also known as integrative encoding (Shohamy & Wagner, 2008). This theory proposes that integration consists of (1) reactivation of prior, related episodes followed by (2) binding of current and previously acquired information in memory. To continue with the previous example, to successfully integrate two novel facts (AB: *Apple seeds are called pips;* BC: *Cyanide is found in pips*), the initially learned factual information (AB) must be reactivated while processing the second, related fact (BC), and the shared meaning amongst them must be bound in memory (ABC). Yet this leaves open the questions of how information becomes combined in memory and what initiates integration mechanisms in the first place. These issues are the subject of two additional theoretical accounts which posit that retrieval-mediated, integrative encoding is invoked for the explicit purpose of resolving conflicts between pre-existing knowledge and newly learned information (Preston & Eichenbaum, 2013; Van Kesteren, Ruiter, Fernández, & Henson, 2012). By this logic, one's current knowledge about pips (that they are equivalent to apple seeds) is challenged when an individual newly learns that pips are also associated with cyanide, thus violating memory-based predictions about the semantic meaning one expects to be associated with pips. Therefore, some form of comparison is hypothesized to trigger memory integration, in which individuals recognize that incoming information deviates from previously learned information in some way. To resolve the conceptual mismatch, individuals must then comprehend and integrate the conceptual relation, which therefore leads to an update of one's existing knowledge.

Support for the subprocesses of productive knowledge extension comes from a growing body of fMRI research. In this literature, the approach has tended to be that researchers examine patterns of neuroanatomical activation and then draw backward inferences about the process(es) being observed. For instance, using the paired associate inference task (and other paradigms) fMRI studies have converged on the finding that the medial temporal lobe (MTL) and the medial PFC are critical for knowledge extension through memory integration (see Schlichting & Preston, 2015; Zeithamova, Schlichting, & Preston, 2012b for reviews). Based on the hypothesized functions of these regions, such as in binding and retrieval, we can therefore use activation in neural structures to begin to inform the underlying processes involved. However, as the brief review of this work will make clear, with the exception of one noteworthy study (Zeithamova et al., 2012a), the low temporal resolution of fMRI prevents demarcation of the precise mechanisms supporting knowledge extension through integration or of the time-course of these processes. Thus, the present ERP investigation provides a complementary level of analysis that affords high temporal resolution, thereby offering clarity on the current debate.

Signatures of integrative encoding have consistently been demonstrated across studies, with unique patterns of activation suggesting the involvement of several subprocesses, including reactivation, novelty detection, and comparison. Consistent with the argument that memory integration occurs at the time of learning, activity in the hippocampus, parahippocampal cortex (PHC), and fusiform gyrus is associated with encoding of a related paired associate (BC), and is also predictive of subsequent transfer success (Zeithamova & Preston, 2010). To further examine the specific mechanisms supported by these regions, Zeithamova and colleagues (2012a) directly tested the assumption that memory integration is accomplished through retrieval-mediated learning. They presented participants with object-scene paired associates and used multivoxel pattern analysis (MVPA) to determine if simultaneous processing of prior objects (AB) occurred upon exposure to novel, related scenes (BC), and vice versa. They found that reactivation of prior, overlapping episodes occurs upon experience of new information, providing conclusive evidence that integration is accomplished through retrieval of prior memory representations. It has additionally been proposed that integration requires, "the detection of differences between reactivated memories and current events" (Schlichting et al., 2014; pg. 1249). This argument arises from the finding that during encoding of separate yet related paired associates (BC) only activity in hippocampal subfield CA_1 is predictive of subsequent success on inferential judgments. Based on the hypothesized role of CA_1 in detecting the existence of novelty (e.g., Larkin, Lykken, Tye, Wickelgren, & Frank, 2014), the proposal is that CA_1 serves as a comparator, and that this hippocampal region triggers a host of subsequent encoding processes that then enable one to abstract and represent the relation between newly and previously learned content. Thus, in summary, these studies suggest that memory integration occurs during encoding and might be supported by some combination of reactivation, novelty detection, and comparison. Yet the abstraction and binding of relational understandings, as well as the timing of these processes, have thus far only been assumed, not directly observed.

The precise mechanisms supporting knowledge extension at the time of explicit testing are similarly under debate, though it has been proposed that reactivation, recombination, organization, and/or conflict resolution of newly and previously learned information might be engaged. Broadly speaking, inferential reasoning is supported by activation in both the MTL and PFC (Kumaran, Summerfield, Hassabis, & Maguire, 2009; Sweegers, Takashima, Fernández, & Talamini, 2014; Zeithamova et al., 2012a; Zeithamova & Preston, 2010). Thus, knowledge extension is thought to rely on retrieval of previously integrated knowledge (MTL) followed by flexible recombination of that information to infer the correct answer (PFC) (Ranganath, 2010; Zeithamova & Preston, 2010). Consistent with this argument, whereas greater activation in PFC regions tracks with trial-by-trial transfer accuracy, individual differences between good and poor performers are predicted by greater PHC and hippocampal activity, with greater coupling between the PFC and the PHC, hippocampus, and posterior visual cortex observed during successful inferential judgments (Zeithamova & Preston, 2010). Further support for the idea that integrated representations are retrieved and flexibly recombined during knowledge extension comes from the finding that the same pattern of hippocampal activity shown at encoding is reinstated when a successful inference is made at test (Schlichting et al., 2014).

Beyond flexible recombination, activity invoked by the PFC after initial integration has also been interpreted as reflecting the processes of long-term storage (e.g., Takashima et al., 2006, 2009) and of organizing and/or resolving the conflict between newly and previously learned information (e.g., Hasselmo & Eichenbaum, 2005; Ross, Sherrill, & Stern, 2011). Support for these arguments comes from the finding that the PFC is not exclusively recruited when participants are faced with explicit demands to use prior knowledge. That is, when participants are exposed to the same paired associates repeatedly at encoding, decreased hippocampal engagement and increased VMPFC engagement is observed across repetitions, which relates to the ability to successfully infer the relation during a subsequent test (Zeithamova et al., 2012a). The gradual shift from hippocampal to VMPFC activity across learning, even in the absence of an explicit test, suggests that the VMPFC is recruited after initial integration, and has been proposed to reflect transfer of integrated representations for future use (e.g., Takashima et al., 2006, 2009) as well as the process of organizing and/or resolving the conflict between newly and previously learned information (e.g., Hasselmo & Eichenbaum, 2005; Ross et al., 2011). Conversely, this same pattern of results has also been interpreted to reflect the role of the medial PFC in biasing reactivation toward behaviorally relevant memories needed *prior* to integration (Schlichting & Preston, 2015). Thus, additional research is needed to elucidate the differential processes invoked at the time of encoding, as opposed to at the time of test, and how these subprocesses unfold over time.

In summary, the neuroimaging literature suggests that successful memory integration depends on several distinct cognitive processes, including reactivation, novelty detection, comparison, meaning abstraction, and binding. On the other hand, explicit extension of previously learned information also seemingly involves multiple processes, including PFC-guided retrieval, recombination, conflict resolution, and organization. Yet as this literature makes clear, these processes have not been isolated in a single investigation, thereby preventing a clear consensus on the temporal unfolding of the processes involved as well as how they jointly contribute to behavior. To address this issue, in the current study we employed ERPs, a methodology with millisecond-level temporal resolution thus permitting observation of different stages of processing as they unfold over time (Luck, 2014). Moreover, as recently demonstrated by Bauer and Jackson (2015), ERPs are a powerful tool for tracking the rapid transition of information from newly integrated to well-known within the knowledge base. Specifically, college students read long lists of facts, some of which could be integrated to generate new knowledge and ERPs were recorded at the test phase as subjects read: (a) well-known facts, (b) novel facts, and (c) facts derived through integration of information presented at encoding. The main finding was that neural responses to well-known and integration facts differed from novel between 500-900msec, but well-known and integration facts did not differ from each other. Although the study was not designed for examination of self-generation of new knowledge or of integration at the time of encoding, it shows the sensitivity of ERPs to changes in the status of information within the knowledge base and the potential to detect temporally-staged component processes.

In the present experiment we recorded ERPs during encoding of separate yet related facts as well as during self-generation of new knowledge at the time of test. We subsequently sorted the trials based on whether knowledge extension through integration was successful or not. This allowed for direct observation of the time-points at which the ERP waveforms diverge on correct and incorrect trials, thus shedding light on the distinct mechanisms involved. Critically, the approach does not rely on eliciting specific ERP components, per se. Nevertheless, guided by the extant fMRI literature, we expected differences in ERPs reflective of novelty detection, retrieval, and binding (see Wilding & Ranganath, 2012 for review) and relational reasoning ERPs indicative of logical manipulation (e.g., Qiu, Li, Chen, & Zhang, 2008). In light of the component independent approach, we also sought to use standardized cognitive measures to constrain our interpretation of the processes underlying the neural components elicited. Specifically, we examined relations between the ERPs and six distinct cognitive processes: long-term retrieval, logical reasoning, working memory, short-term memory, verbal comprehension, and reading comprehension. Together, the approach provides a framework from which to better understand the distinct neurocognitive mechanisms underlying the formation of an organized semantic knowledge base.

Method

Participants

A total of 120 adults participated in the study. The sample consisted of individuals whose behavioral self-generation data was collected as part of a larger investigation (Varga & Bauer, under review). In the course of the study, we also obtained electrophysiological measures, the results of which were not included in any prior published reports. Here we feature analyses of the electrophysiological assessments during encoding of the related information and the test for knowledge extension, and their relation with performance on standardized cognitive measures.

All participants were native-English speakers enrolled in psychology courses at a select private university. Because the goal of the present investigation was to dissociate the neural mechanisms involved in successful (as opposed to unsuccessful) knowledge extension through memory integration, data are reported for participants in the middle of the performance distribution, who contributed comparable levels of successful and unsuccessful trials. Of the 120 individuals tested, 58 satisfied our inclusion criteria of a minimum of 10 successful and 10 unsuccessful trials. The data from five of these participants were excluded because of: technical malfunction during recording ($N = 3$), self-reported diagnosis of Dyslexia which may have altered the electrophysiological indices ($N = 1$), and excessive movement resulting in an insufficient number of artifact-free trials following data reduction ($N = 1$). Thus, in total data are reported for 53 adults between 18-24 years ($M = 19.91$; $SD = 1.36$; 30 females). The final sample was 9% African American, 21% Asian, 64% Caucasian, and 2% mixed racial descent, based on self-report. Nine percent of the participants were of Hispanic descent. Two participants did not report racial or ethnic information. The protocol and procedures were approved by the university Institutional Review Board. Written informed consent was obtained prior to the start of the study. Participants were compensated with course credit for their time and participation.

Stimuli

The stimuli were 60 stem facts and 30 novel integration facts. Facts ranged from 4-10 words and were intended to be educationally meaningful. For instance, two stem facts were about art history (*A popular sculpture made from a urinal is called Fountain; Duchamp's most well-known work is named Fountain)*. Integration of separate but related stem facts could lend itself to self-generation of a novel integration fact (*Duchamp's most popular work consisted of a urinal*). The stimuli captured a variety of logical relations regularly encountered in everyday learning conditions and which have previously been shown to invoke integration mechanisms (see Zeithamova et al., 2012b for review).

Importantly, prior research employing these stimuli has demonstrated that the facts are novel to young adults and that both facts from a given pair are necessary for generation of the target integration facts (see Varga & Bauer, under review; Experiment 2). Specifically, when participants were exposed to only one of the two stem facts from a pair, they generated the novel integration facts only 11% of the time (which significantly differed from the 44% demonstrated when both stem facts were provided). Thus, exposure to the information presented in both stem facts is necessary to reliably derive the corresponding integration fact.

Procedure

Encoding phase. At the start of the session, participants were advised that we were interested in whether neural activity for newly learned factual information differs as a function of its subject domain. They were fitted with a Brain Products high precision ActiCap (Brain Products Inc, GmbH, Munich, Germany) with 32 electrodes positioned according to an adaptation of International 10-20 system (Jasper, 1958). Once the cap was fitted and electrode impedances were lowered, participants read a total of 60 sentences presented in rapid serial visual presentation on an LCD computer monitor. As depicted in Figures 1A and 1B, words of the sentences were presented one at a time for 400msec each. ERPS were time-locked to a target word at the end of each sentence, which consisted of a repeated word linking to-be-integrated stem facts to one another (e.g., *Pips*, *Fountain*, etc.). Following each sentence, participants were shown a decision screen and asked to indicate, via a button-press response, whether the information conveyed was novel or known. The incidental task was designed to ensure that participants were attending to the facts while also corroborating the pretext of the study purpose (i.e., learning of *novel* information). At no time were participants informed that any of the sentences were related. Across the encoding phase, to-be-integrated stem facts were separated by a lag of 2 to 4 intervening sentences. Variability in lag created temporal distance between to-be-integrated information and also prevented participants from anticipating the content of the next fact. Moreover, the short lag served to promote recognition of the relatedness between paired facts, thereby allowing for examination of memory integration under optimal learning conditions. Fact order was counterbalanced such that each stem fact from a given pair was presented in the first or second serial position an approximately equal number of times across the sample.

Test Phase. After a short rest lasting 5-10 minutes, participants were presented with 30 facts derived through integration of the previously presented stem facts. As depicted in Figure 1C, the sentences were presented in the form of questions by replacing the final word of each fact with "?". ERPs were time-locked to the onset of the "?", which cued participants to generate a one-word answer that could accurately complete the sentence. Participants were given an unlimited amount of time to generate a response. When an answer was generated, participants made a button-press response which was followed by an "Answer" screen signaling them to speak the answer aloud. Once an answer was provided, the participant was instructed to remain as still and alert as possible. When the speaking-related muscle activity returned to a baseline level (after about 10-20 seconds), the experimenter initiated presentation of the next question.

Cognitive Assessments. Participants returned to the laboratory approximately one week after their initial visit (*M* delay = 7.08, *SD* = 0.76, *Range* = 6-11 days) and were instructed that we were interested in whether performance on a number of cognitive tasks related to performance at the initial session. The *Woodcock Johnson Test of Cognitive*

Abilities, Third Edition (*WJ-III COG*) (Woodcock, McGrew, & Mather, 2001), the *Test of Memory and Learning, Second Edition* (*TOMAL-2*) (Reynolds & Voress, 2004), and the *Woodcock Language Proficiency Battery—Revised* (*WLPB-R*) (Woodcock, 1991) were used to assess six standardized cognitive domains (see Table 1 for thorough description of each task). Cognitive assessments were then administered in the following fixed order: (1) short-term memory (Digits Forward), (2) fluid reasoning (Concept Formation), (3) working memory (Digits Backward), (4) long-term retrieval (VAL), (5) reading comprehension, and (6) verbal comprehension. To avoid the potential for unique order effects across the sample, we did not counterbalance the sequence of cognitive assessments, thereby reducing error and increasing power.

ERP Recording

Brain electrical activity measured at Session 1 was recorded at 32 scalp sites using active Ag/AgCl electrodes mounted in an elastic cap (Brain Products). Application and recording were conducted while participants were seated approximately 90cm in front of the LCD monitor. Reference electrodes were mounted to the left and right mastoid bones via double adhesives. Impedances were consistently below 35 kΩ, and generally below 15 kΩ. The EEG was continuously sampled at 500Hz for off-line analysis using open-source Pycorder software (Brain Products). Data were referenced online to a virtual ground. No band-pass filters were applied during recording.

ERP Data Reduction & Analysis

Offline data processing was completed using EEGLAB 13.4.3b (Delorme & Makeig, 2004) and ERPLAB 4.0.3.1 (www.erplab.org) operating in Matlab R2014a (MathWorks, Natick, MA, USA). A high-pass filter with a half-amplitude cut-off of 0.1

Hz and a roll-off of 12 decibels/octave was first applied to the raw EEG data. Data then were processed with independent component analysis (ICA) to identify and remove artifacts caused by eye-blinks, saccades, line noise, and bursts of electromyography (EMG). Following ICA, a low-pass filter with a half-amplitude cut-off of 30 Hz and a roll-off of 12 decibels/octave was applied, and data were subsequently re-referenced to mathematically linked mastoids. The continuous EEG data for each participant was segmented into 2600msec epochs beginning 200msec before stimulus onset and ending 2400msec after stimulus onset. The 200msec pre-stimulus window was used to correct for baseline activity on each individual trial, and trials containing amplitudes that exceeded ±150 μv were rejected from the averaged epochs. Across participants, we created separate grand averages for trials in the Stem 1 (Encoding), Stem 2 (Encoding), and Self-generation (Test) conditions, which were each further sorted trial-by-trial according to self-generation performance (successful versus unsuccessful). Final trial counts for each condition of interest following data reduction are provided in Table 2.

Several key assumptions guided our analysis of the ERP data. First and foremost, if new knowledge is integrated with existing knowledge at the time of encoding, it is necessarily the case that memory integration occurs upon encoding of the second, related stem fact (as opposed to encoding of the first stem fact from a target pair). Therefore, we hypothesized that integration would be set into motion upon exposure to the Stem 2 fact, and that differences in neural processing of Stem 2 as a function of subsequent performance would reflect the mechanisms involved in memory integration. Second, because each Stem 2 fact is time-locked to a repeated word (e.g., *Pips*, *Fountain,* etc.) which serves as a relational link to previously learned factual content (i.e., first presentation of *Pips*, *Fountain,* etc. in Stem 1), the Stem 1 and Stem 2 conditions should naturally give rise to different levels of familiarity which will contribute to differential neural responses. As a result, the primary encoding analyses concerned time points in which processing within Stem 1 and within Stem 2 differed as a function of subsequent performance. Although direct comparison between Stem 1 and Stem 2 was not the main focus, when of potential theoretical relevance, supplemental analyses exploring differential processing between these conditions was provided. Third, we also anticipated that the processes required to self-generate the integration facts would differ from encoding of the individual stem facts. Thus, similar to the encoding analyses, we examined the test phase data separately from that of Stem 1 and Stem 2 processing, according to whether participants were subsequently successful.

Based on previous research and visual inspection of the ERP waveforms, we identified time windows in which successful and unsuccessful neural processing diverged across the three conditions (Stem 1, Stem 2, Self-generation). For the encoding analyses, mean amplitude of the ERP response was examined over four time windows (400-600msec, 600-1100msec, 1100-1350msec, 1350-1700msec) at the frontal and central electrode sites (see Figure 2; circle). As discussed below, in the latest encoding window (1350-1700msec) we also examined ERP activity in an additional parietal cluster (see Figure 2, ellipse). For the test phase analyses, ERP responses were examined in a single time window (1000 – 1500msec) at the same parietal electrode cluster examined during the latest encoding window (Figure 2; ellipse). For each condition (Stem 1, Stem 2, Self-generation), two-way repeated-measures analyses of variance (ANOVA) were conducted with factors of performance (Correct or Incorrect) and electrode site. Greenhous-Geisser corrections were applied in cases of violation of sphericity. Because they do not inform the research question, effects of electrode site were not reported unless they interacted with performance.

Results

On average, participants self-generated the novel integration fact on 51% (*SD* = 9.69%) of the possible trials, with performance ranging from 34% to 66% across the sample. Figure 3 depicts grand averaged waveforms for all electrode channels during the encoding phase. Figure 4 depicts grand averaged waveforms for all electrode channels during the test phase. Results are divided into six sections: responses to the stem facts across the four different encoding time windows, responses during self-generation of the novel integration facts at test, and associations amongst encoding and test phase responses as well as with standardized cognitive measures.

Encoding responses at Frontal-Central Sites: 400-600msec

Grand averaged waveforms for all frontal-central sites included in the encoding analyses are plotted in Figure 5a. Means and standard errors of neural responses to stem facts 1 and 2, by subsequent performance, are presented in Table 3. No main effect of performance was observed for processing of the Stem 1 facts in this early time window, $F(1, 50) = 0.89$, $p = 0.35$, $p\eta^2 = 0.89$. In contrast, a main effect of performance was observed for the Stem 2 facts, $F(1, 50) = 4.53$, $p = 0.04$, $p\eta^2 = 0.83$, such that responses were more positive on incorrect trials (Table 3). Given the more positive response to Stem 2 on incorrect trials, we further examined whether processing in this condition differed from that of Stem 1 correct and incorrect trials. The results of these ANOVAs are summarized in Table 4, along with other contrasts of potential interest. As is indicated in Table 4, processing of Stem 2 on incorrect trials differed from all other conditions, whereas, processing of Stem 2 on correct trials did not differ from Stem 1 processing.

Encoding responses at Frontal-Central Sites: 600-1100msec

In contrast to the differential activity observed in the earliest time window, neural responses did not differ by performance for either the Stem 1 facts, $F(1, 50) = 0.38$, $p = 0.54$, $p\eta^2 = 0.01$, or for the Stem 2 facts, $F(1, 50) = 0.59$, $p = 0.45$, $p\eta^2 = 0.01$, in this middle time window. Thus, as reflected by the averaged waveforms in Figure 5a, the differential processing of Stem 2 previously evidenced on incorrect trials was resolved in the window between 600 and 1100msec.

Encoding responses at Frontal-Central Sites: 1100-1350msec

Differential neural responses at the frontal-central sites were once again observed at the frontal-central sites from 1100-1350msec (Figure 5a). Statistical analysis of mean amplitude revealed a main effect of performance for the Stem 1 facts, $F(1, 50) = 8.64$, $p = 0.01$, $p\eta^2 = 0.15$, such that mean responses were more negative on incorrect trials (Table 3). Responses to the Stem 2 facts also differed as a function of subsequent performance, $F(1, 50) = 4.08$, $p = 0.049$, $p\eta^2 = 0.08$. Consistent with the pattern observed for Stem 1, mean responses were more negative on incorrect trials (Table 3). We also examined potentially theoretically relevant contrasts between processing of Stem 1 and Stem 2. As indicated in Table 5, processing differed across all possible stem fact x performance comparisons. Thus, Stem 1 responses were more negative than Stem 2 responses, and incorrect performance was more negative within each stem condition.

Encoding responses at Frontal-Central and Parietal Sites: 1350-1700msec

Processing of Stem 1 at the frontal-central sites did not differ by performance, $F(1, 50) = 1.51$, $p = .23$, $p\eta^2 = 0.03$. Although processing of Stem 2 at the frontal-central sites also did not differ by performance, $F(1, 50) = 0.21$, $p = .65$, $p\eta^2 = 0.004$, there was a significant interaction between performance and electrode site, $F(3.20, 159.97) = 9.30$, $p <.001$, $p\eta^2 = 0.12$. Follow-up post-hoc pairwise comparisons indicated that processing of correct and incorrect Stem 2 facts differed significantly at the central-parietal electrode sites (i.e., CP1 and CP2), such that responses were more positive on incorrect trials as compared to correct trials (ps < .05; contrasts became non-significant when Bonferroni corrections were applied). Guided by the finding that a Stem 2 performance effect was evident at the only parietal sites included in the frontal-central cluster, coupled with the pronounced separation observed between correct and incorrect Stem 2 trials at other posterior sites (Figure 3), we additionally examined stem fact processing during the same time window within a parietal cluster (Figure 2). As reflected in the averaged parietal waveform in Figure 5b, Stem 1 processing did not differ by performance, $F(1, 50) = 1.51$, $p = .06$, $p\eta^2 = 0.80$. However, processing of Stem 2 significantly differed as a function of subsequent performance, $F(1, 50) = 9.22$, $p = .004$, $p\eta^2 = 0.16$. More specifically, responses were more positive on incorrect trials ($M = 5.89$; $SE = 0.46$) as compared to correct trials ($M = 4.78$; $SE = 0.39$). Consistent with the supplemental analyses provided for the other time windows, we further explored whether processing of Stem 1 and Stem 2 differed. As summarized in Table 6, Stem 2 responses on both correct and incorrect trials differed from Stem 1, regardless of performance. Thus, responses to Stem 2 were generally more positive than responses to Stem 1, and responses to Stem 2 were further differentiated as a function of subsequent success.

Test phase responses at Parietal Sites: 1000-1500msec

As depicted in the grand averaged waveforms plotted in Figure 4, prominent differences were observed across the scalp between trials in which participants were and were not successful at self-generating the novel integration facts. Differential processing was particularly striking when activity at the parietal sites was averaged, as plotted in Figure 6. Statistical analysis of mean amplitude across the parietal electrode cluster revealed a main effect of performance, $F(1, 48) = 8.02$, $p = 0.007$, $p\eta^2 = 0.14$, such that responses were more positive on incorrect trials ($M = 3.86$, $SE = 0.40$) as compared to correct trials ($M = 2.88$, $SE = 0.35$). There was also a performance by electrode site interaction, $F(3.90, 187.41) = 2.76$, $p = 0.03$, $p\eta^2 = 0.05$. Follow-up pairwise comparisons revealed that responses on correct and incorrect trials significantly differed at all electrode sties with the exception of channel P8 ($p = 0.23$).

Associations amongst Self-generation, Neural Responses, and Cognitive Processes

In light of differential neural responses to Stem 2 as a function of whether participants were successful or unsuccessful at extending knowledge through integration, we also examined whether neural responses were associated with performance on the behavioral self-generation task. Next, we investigated whether neural responses during the encoding phase were correlated with neural responses during the test phase. Finally, in an effort to elucidate the cognitive indices of successful memory integration and self-generation, we assessed associations amongst neural responses and six standardized cognitive abilities (see Table 1 for description). In each set of analyses we examined neural responses shown to differ according to performance, Specifically, we analyzed correct Stem 2 responses at encoding (Frontal-Central sites from 400-600msec and 1100-1350msec; Parietal sites from 1350-1700msec) and correct self-generation responses at test (Parietal sites from 1000-1500msec). Given the distinct pattern of responses to Stem 2 on incorrect trials in the 400-600msec window, we also included Stem 2 incorrect responses in this time window in the correlation analyses.

No significant correlations were observed between behavioral self-generation performance and neural responses. Nevertheless, it is worth noting that the association between behavioral self-generation and mean amplitude on incorrect Stem 2 trials from 400-600msec approached significance, $r(49) = .24$, $p = 0.087$. Examination of neural responses during encoding and at the test for self-generation revealed significant correlations. Specifically, mean amplitude of neural responses during explicit self-generation was positively associated with Stem 2 processing between 1100-1350msec (frontal-central sites), $r(44) = .37$, $p = 0.01$**,** as well as between 1350-1700msec (parietal sites), $r(44) = .35$, $p = 0.02$. Lastly, standardized cognitive processes were correlated with neural activity during the early encoding window (frontal-central sites, 400-600msec) and during the test for self-generation (parietal sites, 1000-1500msec). Specifically, neural responses to Stem 2 on successful trials (400-600msec) were positively associated with measures of Concept Formation, $r(49) = .49$, $p < 0.001$, and Reading Comprehension, $r(51) = .35$, $p = 0.01$. Furthermore, neural responses during successful self-generation (1000-1500msec) were positively associated with Concept Formation, $r(45) = .35$, $p = 0.02$, and marginally associated with Verbal Comprehension, $r(44) = .28$, $p = 0.06$. This pattern of results corroborates the prior observation of multiple temporally staged electrophysiological indices during encoding and self-generation, and provides further insight into the cognitive processes involved.

General Discussion

The present research afforded observation of the distinct, temporally-staged processes required to successfully extend knowledge through memory integration. Processing that indexed memory integration was evident by as early as 400-600 msec after the onset of a second, related stem fact. This early effect served as a precursor to subsequent encoding processes recruited between 1100-1350 msec and 1350-1700 msec. Unlike the multiple processes observed at encoding, only one component was sensitive to successful self-derivation of new knowledge at the time of test (1000-1500 msec). Interestingly, this effect also was significantly associated with the neural activity elicited during successful memory integration during the latest encoding windows (spanning from 1100-1700 msec). Importantly, several of the components shown to distinguish successful from unsuccessful knowledge extension through integration in the present research exhibit the same timing and topography as well-established ERPs documented in the memory and reasoning literatures. Thus, in the discussion that follows we provide an interpretation of the processes likely reflected by each component elicited and discuss how this pattern of results contributes to our theoretical understanding of semantic knowledge extension.

Encoding Process 1: Detection of a Semantic Deviation

When participants were presented with a novel fact (i.e., Stem 2) that was related to a previously encountered fact (i.e., Stem 1), information that individuals failed to integrate (as suggested by unsuccessful self-generation) elicited a more positive ERP deflection than successfully integrated information between 400-600 msec at frontal-central sites. This difference is both temporally and topographically similar to the *midfrontal old-new effect*—an effect typically evident between 300-500 msec at frontal midline sites in studies of recognition memory (Curran, Tepe, & Piatt, 2006; Rugg et al., 1998; Wilding & Ranganath, 2014 for review). As the name implies, this component is elicited when participants accurately judge previously studied items to be "old" and unstudied items as "new," with a more positive-going ERP deflection elicited in response to old items. In a common manipulation employed to examine this effect, ERPs are recorded while subjects make recognition judgments in response to studied items (i.e., old), unstudied items (i.e., new), and unstudied items that are highly similar to studied items (i.e., lures). For instance, Curran and Cleary (2003) showed that ERPs elicited when participants false alarmed to picture lures (i.e., judged mirror reversals of studied pictures as old) were more positive than those elicited when subjects correctly judged a picture as new. This same pattern has also been shown using semantically similar words (Nessler, Mecklinger, & Penney, 2001) and plurality-reversed words (e.g., cat vs. cats) (Curran, 2000), with many arguing that the midfrontal old-new effect is a neural correlate of familiarity, rather than explicit recollection (Curran et al., 2006; Rugg et al., 1998). In the present research the midfrontal old-new effect was apparent on incorrect, but not correct, Stem 2 trials (e.g., at the second mention of information about *pips*). It is therefore plausible that the failure to integrate was linked with the tendency to treat novel but related information as "old" (i.e., to be "lured" into treating new information as if it were repeated or old information).

If a reliance on familiarity-based processing leads one to mistakenly identify novel yet related information as old, then we must characterize the alternative mechanism that enables participants to recognize successfully integrated information as new. Although the midfrontal old-new effect is most commonly argued to reflect familiarity (Curran et al., 2006; Rugg et al., 1998), others have claimed that it reflects conceptual priming and should be interpreted with respect to its precursor, the N400 (Paller, Voss, & Boehm, 2007; Voss & Paller, 2006). Indeed, due to the similarity between the midfrontal old-new effect and the semantically related N400, some have even coined it the FN400 (familiarity-related N400) (Curran, 2000), with others going so far as to suggest that semantic priming (i.e., the N400) actually drives familiarity-based recognition judgments (Wilding & Ranganath, 2012; Yonelinas, 2002). The N400 is a negative-going fluctuation that reaches maximum amplitude between 350 to 500 msec over the centro-parietal region of the scalp and is functionally specific to conceptual processing (Federmeier & Laszlo, 2009; Kutas & Federmeier, 2011; Kutas & Hillyard, 1980). A large body of literature suggests that the N400 indexes the extent to which information is expected in context, with more negative amplitudes indicative of greater incongruence between incoming information and information that is either stored in memory or presented in the immediate sentence context that precedes it (e.g., Barrett & Rugg, 1990; Bentin, McCarthy, & Wood, 1985; Federmeier & Kutas, 1999; Holcomb & Neville, 1990; Kellenbach, Wijers, & Mulder, 2000; Kutas & Federmeier, 2000; Zhang, Guo, Lawson & Jiang, 2006). In the present research, a more negative-going response was apparent on correct than incorrect Stem 2 trials (e.g., at the second mention of information about *pips*). It is thus reasonable to conclude that when individuals identify a

Stem 2 fact as novel, a deviation is detected between the current stimulus and prior, related knowledge.

The finding that ERPs to correct (but not incorrect) Stem 2 facts were positively associated with performance on standardized measures of concept formation and passage comprehension further supports our interpretation of comparison-based novelty detection between 400-600 msec. Whereas concept formation assessed the ability to identify relations between concepts, passage comprehension tapped the ability to comprehend meaning in context (i.e., according to cloze probability). Indeed, the N400 amplitude is modulated as a function of cloze probability (Kutas & Hillyard, 1984), thereby supporting the interpretation that differential processing between 400-600 msec likely reflects N400 activity, and whether participants successfully identified the Stem 2 fact as novel as compared with prior, related information.

This comparison-based interpretation is also highly consistent with that proffered by current theories of memory integration (Preston & Eichenbaum, 2013). As discussed previously, CA_1 is thought to trigger memory integration by acting as a comparator, and it is this process that is hypothesized to enable novelty detection. Although we must be cautious in inferring underlying cortical sources on the basis of scalp-based ERP recordings, the midfrontal old-new effect has been interpreted as, “a downstream index of familiarity signals that are generated in the medial-temporal lobe,” (Rugg & Curran, 2007 as cited in Wilding & Ranganath, 2012, pg. 382), consistent with the proposed role of this region in identifying familiar, overlapping content (e.g., second mention of *pips*). It is thus possible that activity originating in the MTL initiates the comparison-based process that either results or does not result in novelty detection. Consistent with this proposal, studies using magnetoencephologram (MEG) and the event-related optical signal (EROS) have converged on the finding that the source of the N400 resides in the MTL, showing that activity originating in the posterior temporal lobe spreads to the anterior temporal lobe and the frontal lobe by its peak and then travels back again (e.g., Federmeier & Laszlo, 2009; Halgren et al., 2002; Tse et al., 2007). This evidence supports the hypothesis that the PFC influences memory integration by biasing reactivation toward relevant memories (Schlichting & Preston, 2015). That is to say, when presented with a second, related stem fact, participants must first recognize that the content is familiar (e.g., they have prior knowledge about pips), a process likely initiated by the MTL. Through interaction with the PFC, it is possible that this region guides reactivation of the appropriate memory trace (i.e., Stem 1 about pips). If the pre-existing memory trace is weak, it will be insufficient to guide detection of a deviation between novel and known information, resulting in only familiarity-based processing on unsuccessful trials. Conversely, if the pre-existing memory trace is intact, participants’ command of the prior knowledge will be sufficient to detect the deviation, which will then initiate subsequent encoding processes aimed at resolving and integrating the mismatch.

Encoding Process 2: Interpretation of the Semantic Deviation

Because successful integration is first distinguished by what appears to be an N400, it is worth noting that meaningful stimuli elicit an N400 with little or no conscious awareness, leading many to suggest that this early processing is fairly “automatic” (e.g., Brualla, Romero, Serrano, & Valdizan, 1998; Rolke, Heil, Streb, & Henninghausen, 2001). Further, unlike other components, the latency of the N400 is highly consistent across paradigms and stimulus types, suggesting that the semantic information extracted in this time window is constrained, thereby necessitating additional processing for other aspects of meaning (see Federmeier & Laszlo, 2009 for discussion). As such, additional effects related to semantic processing often follow the N400 as the semantic memory system "adds to, subtracts from, or otherwise modifies the activation that was established in the initial 'sweep' of semantic memory during the N400" (Federmeier & Laszlo, 2009, Pg. 32). Consistent with this notion, in the present research, pronounced differences between successfully and unsuccessfully integrated information were also observed in a later encoding window that occurred between 1100-1350 msec at the same frontal-central sites.

One commonly observed post-N400 ERP response that bares similarity to the effect observed between 1100-1350 msec consists of a frontal negativity, with greater amplitudes indexing processes associated with active meaning selection in response to perceived ambiguity (e.g., Lee & Federmeier, 2006). In the present research, ERP responses to facts that were unsuccessfully integrated elicited more negative responses than those that were integrated. Therefore, in this second time window, it is possible that participants recruited processes in a more strategic manner in an attempt to interpret the new meaning of the presented information (e.g., *cyanide is found in pips*). If participants previously activated an intact representation of relevant information (e.g., *apple seeds are called pips*), resolution of the meaning should be facilitated through preferential access to related information. Conversely, if participants failed to activate related knowledge previously, or if the activated representation was insufficient, significantly more effort would be required to select relevant content in order to comprehend the meaning in this time window. Interestingly, this frontal negativity was even more pronounced for the

Stem 1 facts, of which participants had no prior knowledge from which to interpret the meaning, which further supports the interpretation that this processing window reflects comprehension of the semantic meaning conveyed.

Support for this account also comes from the reasoning literature. For instance, when subjects need to retrieve information from long-term memory in order to abstract the meaning conveyed by a pair of letter strings (e.g., abc : abe), a post-N400 ERP is observed at frontal-central sites (Qiu et al., 2008). Although the effect is observed in an earlier time window (likely owing to the more constrained nature of the task as compared with the present research), larger negative amplitudes are similarly observed on trials in which induction of the analogical meaning is more difficult. Of interest, source analysis mainly linked this component to the medial PFC, consistent with its role in biasing access to relevant information. In light of complementary interpretations of the late frontal post-N400 negativity in the semantic memory and reasoning literatures, it is reasonable to conclude that participants engaged in strategic meaning interpretation from 1100-1300 msec, with greater ease associated with processing of correct Stem 2 facts, possibly due to PFC-driven preferential access to the relevant Stem 1 facts.

Encoding Process 3: Integration of the Semantic Relation

Immediately following the post-N400 frontal negativity, a posterior positivity was observed between 1350-1700 msec. Post-N400 posterior positivity, often referred to as the late positive complex (LPC) or P600, has been linked to the process of building or extracting relational understandings, such as during language comprehension and production (e.g., Osterhout & Holcomb, 1992), music reading (e.g., Patel, Gibson, Ratner, Besson, & Holcomb, 1998), and abstract pattern completion (e.g., Lelkov-

Boissard & Dominey, 2002). Importantly, the LPC is also reflective of the status of relational understandings in long-term memory (Bauer & Jackson, 2015; Donaldson & Rugg, 1998; 1999). For instance, increased LPC amplitude is evidenced when participants view a face that matches a previously paired scene compared with a nonmatching scene, which has been interpreted to reflect processing of the relation between a current stimulus (e.g., a face) and a reactivated memory trace (e.g., a face-scene representation) (Hannula, Federmeier, & Cohen, 2006). In the present research, greater posterior positivity was evident on trials in which participants failed to integrate, suggesting that more processing resources were needed to represent the relation between overlapping stem facts. Interestingly, processing of Stem 2 facts was more positive than that of Stem 1 facts, regardless of whether information was successfully integrated. This finding suggests that participants might not have engaged in the same relational processing in the absence of knowledge to integrate and/or to resolve in memory (i.e., when presented with Stem 1 facts). It is worth noting that the typical LPC is observed around 500-800 msec post-stimulus onset. However, given that in the present research, the relation between to-be-integrated stem facts must be self-generated through reactivation and evaluation of prior knowledge, and is not directly presented to participants, it is not surprising that this component would be elicited in a later time window. Support for this conclusion comes from the finding that this late posterior activity was also significantly correlated with explicit self-generation at test.

Test Phase Process: Self-generation of New Knowledge

The same pattern of posterior positivity observed during the 1350-1700 msec encoding window was again observed between 1000-1500 msec when participants were prompted to self-generate the novel integration fact. Again, larger amplitudes were elicited on unsuccessful trials, likely reflective of the greater effort associated with extracting the necessary relation. In support of this interpretation, ERPs on successful trials were associated with measures of concept formation (relational reasoning) and moderately associated with verbal comprehension (the ability to reason based on previously learned information). Given that neural activity on successful self-generation trials was significantly correlated with frontal-central activity from 1100-1350 msec (interpretation of semantic meaning) as well as posterior activity from 1350-1700 msec (integration of stem facts in memory), we further propose that these encoding-related processes were reinstated at the time of test in order to extend previously stored knowledge in response to an explicit demand. When participants successfully integrated this information in memory at encoding, self-generative processing was facilitated and subsequently successful. In accord with this account, we know that hippocampal activity at the time of integrative encoding is reinstated during inferential reasoning, with such reinstatement predictive of transfer success (Schilichting et al., 2014). Indeed, the LPC has been linked to the MTL structures (Düzel, Vargha-Khadem, Heinze, & Mishkin, 2001; Olichney et al., 2000), and to the hippocampus in particular (Dietl et al., 2005; Trautner et al., 2004).

Limitations

The present results indicate that several temporally distinct processes support the extension of semantic knowledge through memory integration. However, it is important to note that interpretation of scalp-recorded ERP components is susceptible to the superposition problem. That is, multiple components are superimposed onto any one ERP waveform, making it difficult to decompose the mixture of underlying components (see Luck, 2014 for discussion). Thus, it is possible that the four ERPs elicited in the present research reflect a combination of additional processes. Despite this limitation, in the present research we used difference waves (successful versus unsuccessful performance) to isolate the critical ERP components, which is the most widely accepted approach to dealing with this issue. Nevertheless, further attempts must be made to more completely pull apart these processes through additional experimental manipulations and to test the interpretations offered here. It should also be stressed that without neuroanatomical measures, the brain areas implicated in each stage of processing are only speculative. Thus, future research is needed to determine the temporal dynamics of cortical coupling that gives rise to self-generation through integration.

Conclusion

In summary, the present study employed ERPs to investigate the distinct neurocognitive indices underlying knowledge extension through memory integration. The results indicated that three ERP components were involved in successful memory integration at encoding, and only one component distinguished successful self-generation at the time of test. Guided by the ERP and fMRI literatures, the present results implicate at least three distinct stages of processing during encoding of a second, related stem fact:
(1) detection of a semantic deviation;
(2) semantic interpretation; and (3) semantic integration. With respect to the explicit demand to self-generate the novel integration fact, processing involved in semantic interpretation and integration appeared to be reinstated, thereby leading to greater success and less effortful processing at the time of test. Together, these findings provide support

for many of the processes previously hypothesized to support this critically important form of knowledge extension.

References

Barrett, S. E., & Rugg, M. D. (1990). Event-related potentials and the semantic matching of pictures. Brain and Cognition, 14(2), 201–212.

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), 271-282. doi: http://dx.doi.org/10.1037/a0037405

Bauer, P. J., & San Souci, P. (2010). Going beyond the facts: Young children extend knowledge by integrating episodes. *Journal of Experimental 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. doi: 10.1080/15248372.2013.849707

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials associated with semantic priming. *Electroencephalography and Clinical Neurophysiology*, 60, 343–355.

Brualla, J., Romero, M. A. F., Serrano, M.A, & Valdizan, J. R. (1998). Auditory event-related potentials to semantic priming during sleep. *Electroencephalography and Clinical Neurophysiology: Evoked Potentials*, *108*(3), 283–290.

Curran, T. (2000). Brain potentials of recollection and familiarity. *Cognitive Brain Research, 28,* 923-938.

Curran, T., & Cleary, A. M. (2003). Using ERPs to dissociate familiarity from recollection in picture recognition. *Cognitive Brain Research, 15,* 191-205.

Curran, T., Tepe, K. L., & Piatt, C. (2006). ERP explorations of dual processes in recognition memory. In H. D. Zimmer, A. Mecklinger, & U. Lindenberger (Eds.), *Binding in human memory: A neurocognitive approach* (pp. 139-179). Oxford: Oxford University Press.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9-21.

Dietl, T., Trautner, P., Staedtgen, M., Vannuchi, M., Mecklinger, A., Grunwald, T., et al. (2005). Processing of famous faces and medial temporal lobe event-related potentials: A depth electrode study. *Neuroimage*, *25*(2), 401-407.

Donaldson, D. I., & Rugg, M. D. (1998). Recognition memory for new associations: Electrophysiological evidence for the role of recollection. *Neuropsychologia*, *36*(5), 377-395.

Donaldson, D. I., & Rugg, M. D. (1999). Event-related potential studies of associative recognition and recall: electrophysiological evidence for context dependent retrieval processes. *Cognitive Brain Research*, *8*(1), 1-16.

Düzel, E., Vargha-Khadem, F., Heinze, H. J., & Mishkin, M. (2001). Brain activity evidence for recognition without recollection after early hippocampal damage. *Proceedings of the National Academy of Sciences*, *98*(14), 8101-8106.

Eichenbaum, H., & Cohen, N. J. (2001). *From conditioning to conscious recollection: Memory systems of the brain*. New York: Oxford University Press.

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Languag*e, *41*(4), 469–495.

Federmeier, K. D., & Laszlo, S. (2009). Time for meaning: electrophysiology provides insights into the dynamics of representation and processing in semantic memory. In BH Ross (Ed.), *The Psychology of Learning and Motivation* (pp. 1-44). San Diego, CA: Elsevier.

Gentner, D., & Smith, L. (2012). Analogical reasoning. In V. S. Ramachandran (Ed.), *Encyclopedia of Human Behavior: 2nd Ed.* (pp. 130-136). Oxford, UK: Elsevier.

Halgren, E., Dhond, R. P., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J. D., et al. (2002). N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. *Neuroimage*, *17*(3), 1101–1116.

Hall, G. (1996). Learning bout associatively activated stimulus representations: Implications for acquired equivalence and perceptual learning. *Animal Learning & Behavior*, *24*(3), 233-255.

Hannula, D. E., Federmeier, K. D., & Cohen, N. J. (2006). Event-related potential signatures of relational memory. *The Journal of Cognitive Neuroscience*, *18*(11), 1863-1876.

Hasselmo, M. E., & Eichenbaum, H. (2005). Hippocampal mechanisms for the context-dependent retrieval of episodes. *Neural networks*, *18*(9), 1172-1190.

Holcomb, P. J., & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes, 5*(4), 281–312

Holland, P. C. (1981). Acquisition of representation-mediated conditioned food aversions. *Learning and Motivation, 12*(1), 1-18.

Jasper, H. H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography & Clinical Neurophysiology, 10*, 371-375.

Jee, B. D., Uttal, D. H., Gentner, D., Manduca, C., Shipley, T. F., Tikoff, B., … Sageman, B. (2010). Commentary: Analogical thinking in geoscience education. *Journal of Geoscience Education, 58*(1). 2-13. doi: http://dx.doi.org/10.5408/1.3544291

Kellenbach, M. L., Wijers, A. A., & Mulder, G. (2000). Visual semantic features are activated during the processing of concrete words: Event-related potential evidence for perceptual semantic priming. Cognitive Brain Research, 10(1–2), 67–75.

Kumaran, D., Summerfield, J. J., Hassabis, D., & Maguire, E. A. (2009). Tracking the emergence of conceptual knowledge during human decision making. *Neuron, 63*, 889-901. DOI: 10.1016/j.neuron.2009.07.030

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4*(12), 463–470.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621-647.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203-205.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*(5947), 161–163.

Larkin, M. C., Lykken, C., Tye, L. D., Wickelgren, J. G., & Frank, L. M. (2014). Hippocampal output area CA1 broadcasts a generalized novelty signal during an object-place recognition task. *Hippocampus*, *24*(7), 773-783.

Lee, C. L., & Federmeier, K. D. (2006). To mind the mind: An event-related potential study of word class and semantic ambiguity. *Brain Research, 1081*(1), 191–202.

Lelekov-Boissard, T., & Dominey, P. F. (2002). Human brain potentials reveal similar processing of non-linguistic abstract structure and linguistic syntactic structure. *Clinical Neurophysiology, 32*(1), 72-84.

Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed.). Cambridge, MA: MIT Press.

Nessler, D., Mecklinger, A., & Penney, T. B. (2001). Event related brain potentials and illusory memories: The effects of differential encoding. *Cognitive Brain Research, 10*, 283-301.

Olichney, J. M., Van Petten, C., Paller, K., Salmon, D., Iragui, V., Kutas, M. (2000). Word repetition in amnesia: Electrophysiological evidence of spared and impaired memory. *Brain*, *123,* 1948-1963.

O'Reilly, R. C., & Rudy, J. W. (2000). Computational principles of learning in the neocortex and hippocampus. *Hippocampus*, *10*(4), 389-397.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*(6), 785-806.

Paller, K. A., Voss, J. L., & Boehm, S. G. (2007). Validating neural correlates of familiarity. *Trends in Cognitive Sciences*, *11,* 243-250.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience, 10*(6), 717-733.

Preston, A. R., & Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. *Current Biology*, *23*(17), R764-R773.

Preston, A. R., Shrager, Y., Dudukovic, N. M., & Gabrieli, J. D. (2004). Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus, 14*(2), 148-152. DOI 10.1002/hipo.20009

Qiu, J., Li, H., Chen, A., Zhang, Q. (2008). The neural basis of analogical reasoning: An event-related potential study. *Neuropsychologia, 46*(2), 3006-3013.

Ranganath, C. (2010). Binding items and contexts: The cognitive neuroscience of episodic memory. *Current Directions in Psychological Science*, *19*(3), 131-137.

Reynolds, C. R., & Voress, J. (2007). Test of Memory and Learning (TOMAL-2). *TX*: PRO-ED.

Rolke, B., Heil, M., Streb, J., & Hennighausen, E. (2001). Missed prime words within the attentional blink evoke an N400 semantic priming effect. *Psychophysiology, 38*(2), 165–174.

Ross, R. S., Sherrill, K. R., & Stern, C. E. (2011). The hippocampus is functionally connected to the striatum and orbitofrontal cortex during context dependent decision making. *Brain Research, 1423*, 53-66.

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences, 11*, 251-257.

Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature, 392*, 595-598.

Schlichting, M. L., & Preston, A. R. (2015). Memory integration: Neural mechanisms and implications for behavior. *Current opinion in behavioral sciences, 1*, 1-8.

Schlichting, M. L., Zeithamova, D., & Preston, A. R. (2014). CA1 subfield contributions to memory integration and inference. *Hippocampus*, *24*(10), 1248-1260.

Shohamy, D., & Wagner, A. D. (2008). Integrating memories in the human brain: Hippocampal-midbrain encoding of overlapping events. *Neuron, 60*(2), 378-389. doi:10.1016/j.neuron.2008.09.023

Sweegers, C. C. G., Takashima, A., Fernández, G., Talamini, L. (2014). Neural mechanisms supporting the extraction of general knowledge across episodic memory. *NeuroImage, 87*, 138-146.

Takashima, A., Nieuwenhuis, I. L., Jensen, O., Talamini, L. M., Rijpkema, M., & Fernández, G. (2009). Shift from hippocampal to neocortical centered retrieval network with consolidation. *The Journal of Neuroscience*, *29*(32), 10087-10093.

Takashima, A., Petersson, K. M., Rutters, F., Tendolkar, I., Jensen, O., Zwarts, M. J., ... & Fernández, G. (2006). Declarative memory consolidation in humans: a prospective functional magnetic resonance imaging study. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(3), 756-761.

Trautner, P., Dietl, T., Staedtgen, M., Mecklinger, A., Grunwald, T., Elger, C. E., & Kurthen, M. (2004). Recognition of famous faces in the medial temporal lobe An invasive ERP study. *Neurology, 63*(7), 1203-1208.

Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., ... & Morris, R. G. (2007). Schemas and memory consolidation. *Science, 316*(5821), 76-82.

van Kesteren, M. T., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences, 35*(4), 211-219.

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. (under review). Knowledge extension through memory integration: Factors and conditions that promote the long-term accumulation of knowledge.

Voss, J. L., & Paller, K. A. (2006). Fluent conceptual processing and explicit memory for faces are electrophysiologically distinct. *Journal of Neuroscience, 26,* 926-933.

Wilding, E. L., & Ranganath, C. (2012). Electrophysiological correlates of episodic memory processes. In S. J. Luck & E. S. Kappenman (Eds.), *The oxford handbook of event-related potential components* (pp. 373-395). New York: Oxford University Press.

Woodcock, R. W. (1991). *Woodcock Language Proficiency Battery-Revised: English and Spanish Forms.* Itasca, IL: Riverside Publishing.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Cognitive Abilities*. Itasca, IL: Riverside Publishing.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of the 30 years of research. *Journal of Memory and Language, 46,* 441-517.

Zhang, Q., Lawson, A., Guo, C., & Jiang, Y. (2006). Electrophysiological correlates of visual affective priming. *Brain Research Bulletin, 71*(1–3), 316–323.

Zeithamova, D., Dominick, A. L., & Preston, A. R. (2012a). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron, 75*(1), 168-179.

Zeithamova, D., & Preston, A. R. (2010). Flexible memories: Differential roles for medial temporal lobe and prefrontal cortex in cross-episode binding. *The Journal of Neuroscience, 30*(44), 14676-14684.

Zeithamova, D., Schlichting, M. L., & Preston, A. R. (2012b). The hippocampus and inferential reasoning: building memories to navigate future decisions. *Frontiers in Human Neuroscience, 6,* 1-14.

Table 1.

Description of the standardized measures used to assess distinct cognitive processes

Measure	Cognitive Domain
Verbal Comprehension (WJ-III COG: Test 1)	Extent of one's knowledge, the ability to verbally communicate knowledge, and the ability to reason based on previously learned information
Passage Comprehension (WLPB-R: Test 7)	Ability to use syntactic and semantic clues to complete a missing word in a passage based on cloze probability
Concept Formation (WJ-III COG: Test 5)	Ability to identify relations, form concepts, and solve novel problems in the absence of prior knowledge
Visual-Auditory Learning (WJ-III COG: Test 2)	General processing abilities of storing and recalling information from long-term memory
Digits Forward (TOMAL-2)	Short-term memory span
Digits Backward (TOMAL-2)	Working memory span

Note: WJ-III COG = Woodcock Johnson Test of Cognitive Abilities, Third Edition, TOMAL-2 = Test of Memory and Learning, Second Edition, and *WLPB-R = Woodcock Language Proficiency Battery—Revised.*

Table 2.

Trial counts for each condition sorted according to subsequent performance

Condition	Correct	Incorrect	Total
Stem 1	753	706	1459
Stem 2	754	704	1458
Self-generation	734	668	1402

Table 3.

Means and standard errors of ERP responses in the frontal-central cluster during encoding

			Time Window (msec)			
Condition	Trials	Measure	400-600	600-1100	1100-1350	1350-1700
Stem 1	Correct	Mean	-2.11	-1.65	-1.10	2.45
		(S.E.)	0.35	0.33	0.39	0.43
	Incorrect	Mean	-1.77	-1.44	-2.24	1.90
		(S.E.)	0.34	0.37	0.48	0.48
Stem 2	Correct	Mean	-1.67	-1.06	1.03	3.79
		(S.E.)	0.23	0.30	0.50	0.40
	Incorrect	Mean	-0.65	-0.73	0.08	3.99
		(S.E.)	0.37	0.32	0.45	0.55

Table 4.

ANOVAs across conditions in the early encoding window (400-600 msec) at frontal-central sites

	Conditions compared	df	*F*	*p*
1	Stem 2 [Incorrect] / Stem 1 [Correct]	1, 50	8.08	0.01
2	Stem 2 [Incorrect] / Stem 1 [Incorrect]	1, 50	4.87	0.03
3	Stem 2 [Correct] / Stem 1 [Correct]	1, 50	1.19	0.28
4	Stem 2 [Correct] / Stem 1 [Incorrect]	1, 50	0.05	0.82

Table 5.

ANOVAs across conditions in a later encoding window (1100-1350 msec) at frontal-central sites

	Conditions compared	df	*F*	*p*
1	Stem 1 [Correct] / Stem 2 [Correct]	1, 50	25.00	<.001
2	Stem 1 [Correct] / Stem 2 [Incorrect]	1, 50	6.89	0.01
3	Stem 1 [Incorrect] / Stem 2 [Correct]	1, 50	42.84	<.001
4	Stem 1 [Incorrect] / Stem 2 [Incorrect]	1, 50	21.24	<.001

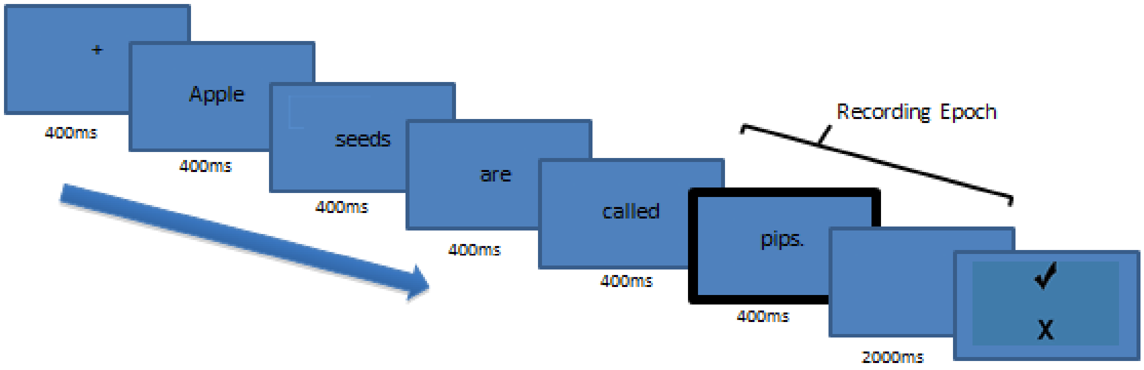
Table 6.

ANOVAs across conditions in the latest encoding window (1350-1700 msec) at parietal sites

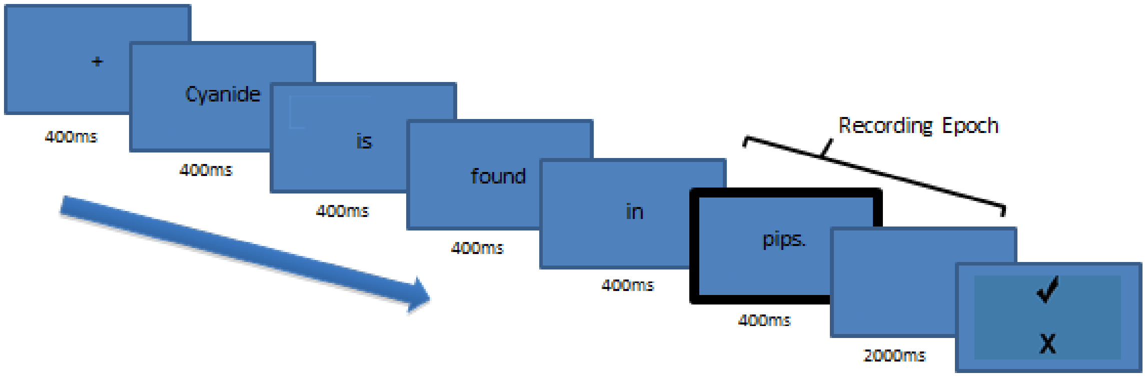
	Conditions compared	df	*F*	*p*
1	Stem 1 [Correct] / Stem 2 [Correct]	1, 50	16.88	<.001
2	Stem 1 [Correct] / Stem 2 [Incorrect]	1, 50	30.73	<.001
3	Stem 1 [Incorrect] / Stem 2 [Correct]	1, 50	10.09	0.003
4	Stem 1 [Incorrect] / Stem 2 [Incorrect]	1, 50	23.15	<.001

Figure 1

Panel A: Encoding of Stem Fact 1



Panel B: Encoding of Stem Fact 2



Panel C: Test for Self-Generation

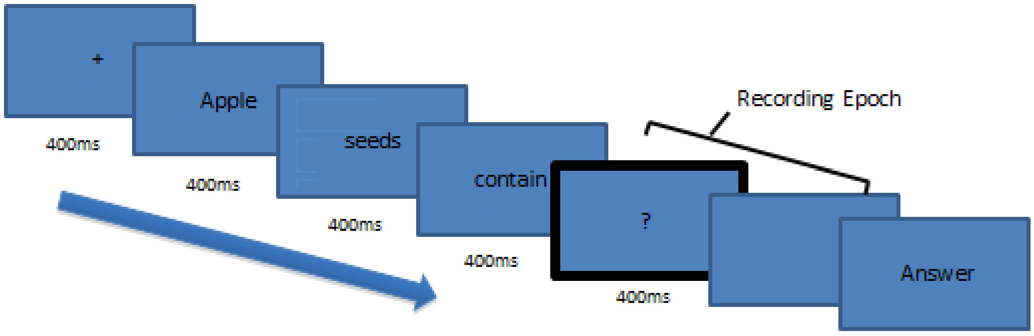


Figure 1. Schematic of encoding (Panels A and B) and test phase (Panel C) procedures. ERPs were time-locked to the onset of the sentence-final word and to the “?” (illustrated in black bolded target stimuli).

Figure 2

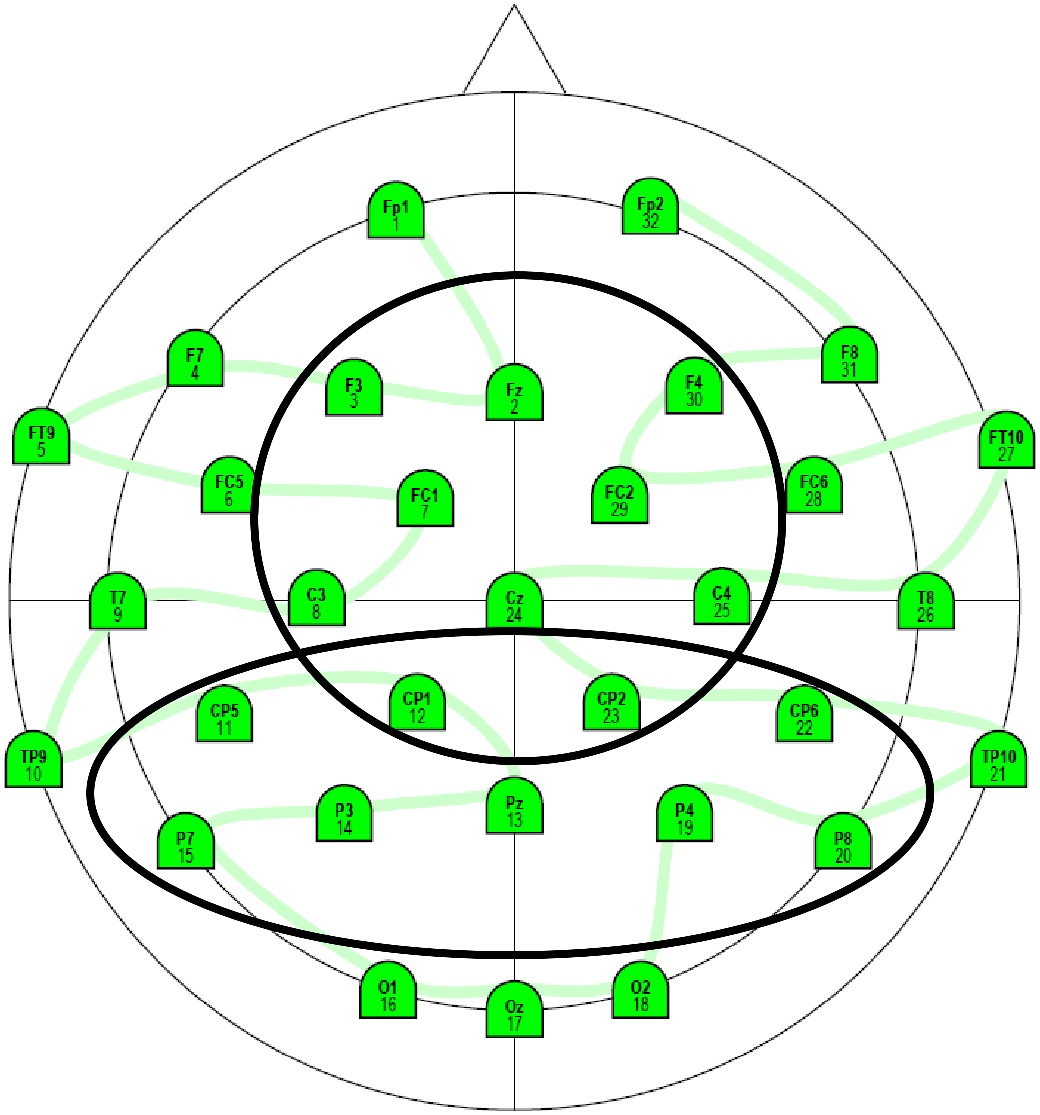


Figure 2. Electrode montage, using the 10-20 system. Frontal-central electrode sites (circle) and parietal electrode sites (ellipse) examined in analyses are circled. Data were re-referenced to the average of the left and right mastoids, TP9 and TP10, respectively.

Figure 3

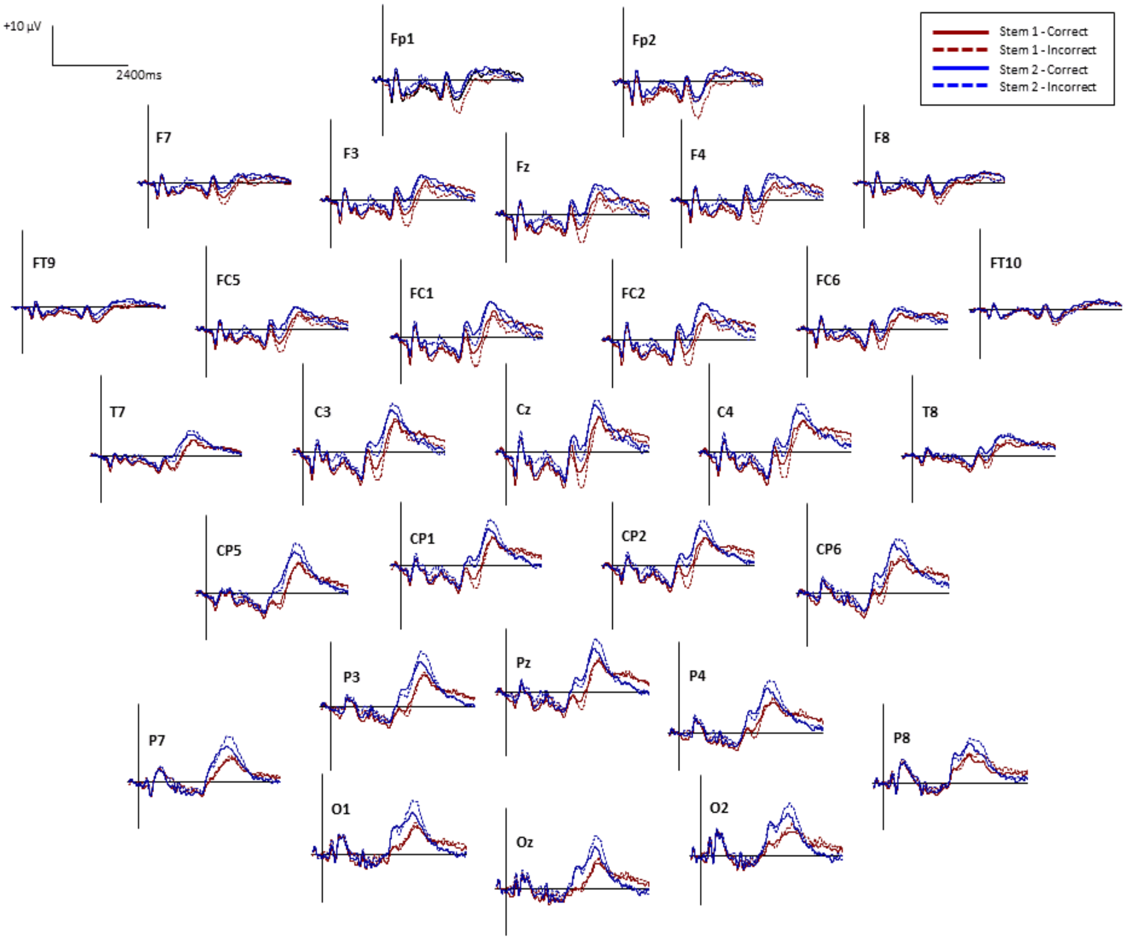


Figure 3. Grand averaged ERP waveforms across all scalp sites during encoding of the Stem 1 facts (red) and Stem 2 facts (blue). For presentation purposes, the data plotted were down-sampled to smooth the waveforms.

Figure 4

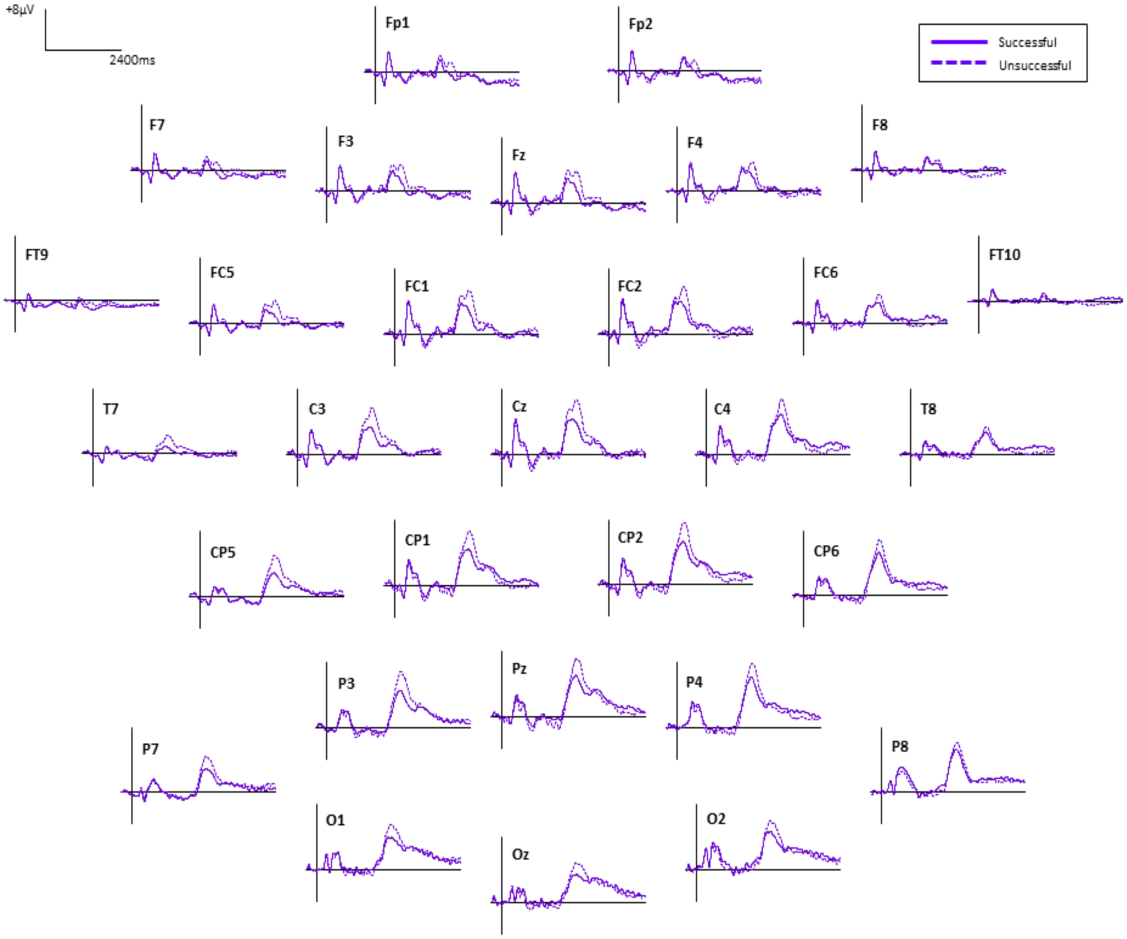
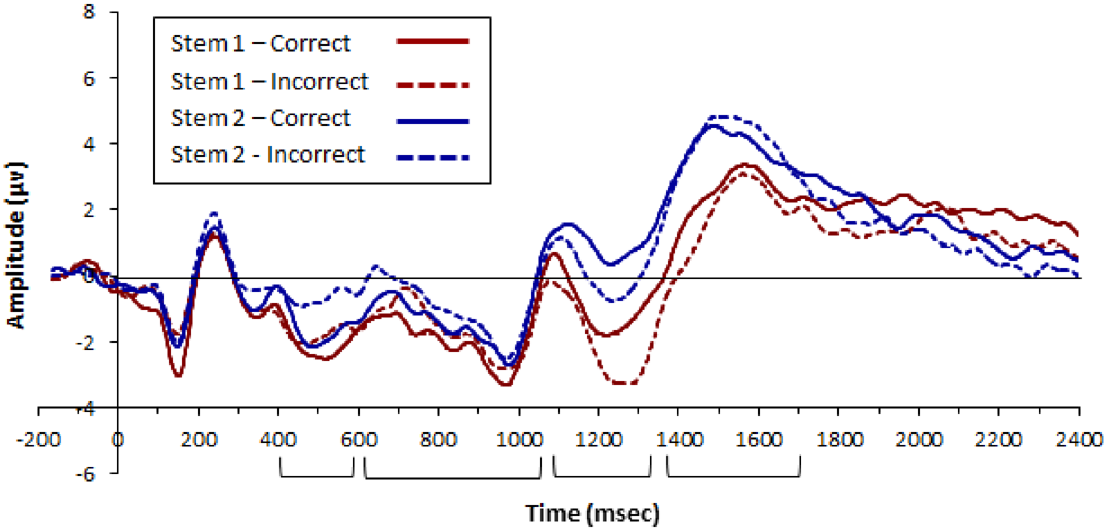


Figure 4. Grand averaged ERP waveforms across all scalp sites during self-generation of the integration facts at the test phase. For presentation purposes, the data plotted were down-sampled to smooth the waveforms.

Figure 5

Panel A: Frontal-Central Cluster



Panel B: Parietal Electrode Cluster

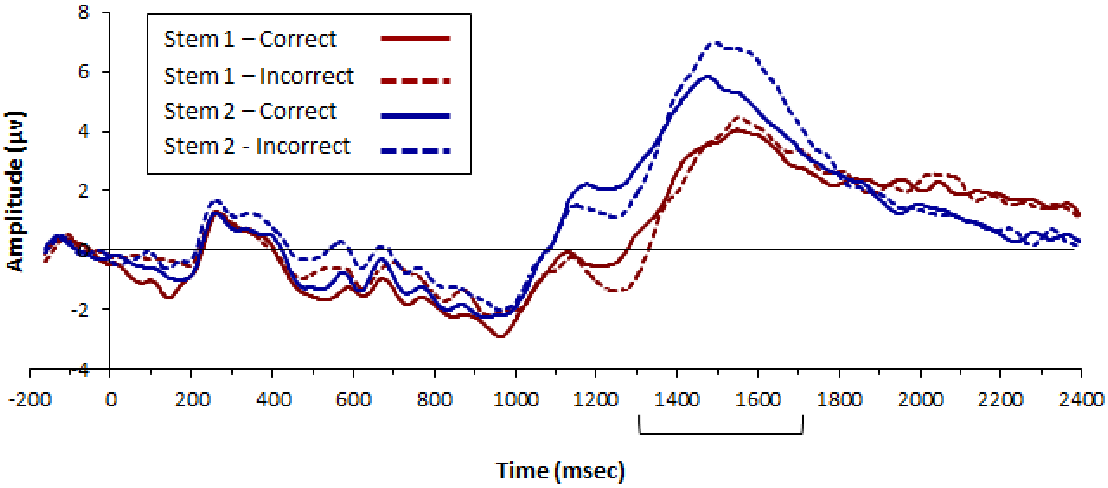


Figure 5. Grand averaged ERP waveforms across the frontal-central sites (Panel A: Fz, F3/F4, FC1/FC2, Cz, C3/C4, and CP1/CP2) and parietal sites (Panel B: CP1/CP2, CP5/CP6, Pz, P3/P4, and P7/P8) at encoding. For presentation purposes, the waveforms were smoothed by down-sampling the time points plotted. Brackets represent the time windows examined in analyses.

Figure 6

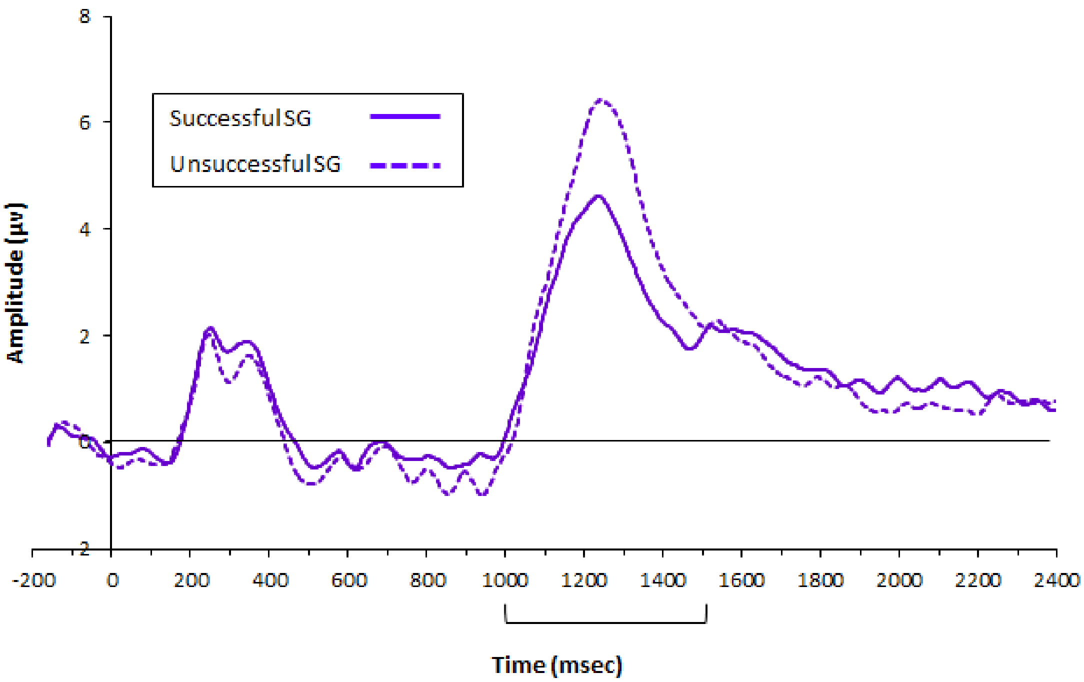


Figure 6. Grand averaged ERP waveforms across the parietal electrode sites (CP1/CP2, CP5/CP6, Pz, P3/P4, and P7/P8) at the test for self-generation. For presentation purposes, the waveforms were smoothed by down-sampling the total number of time points plotted. Brackets represent the time window examined in analyses.

Study 3

Knowledge extension through memory integration:

An explanation of variability in an educationally relevant phenomenon

Nicole L. Varga & Patricia J. Bauer

Emory University

Knowledge extension through memory integration: An explanation of variability in an educationally relevant phenomenon

Building a knowledge base is a central task of development. Much of the content of the knowledge base—so-called *semantic* memory—is directly learned through first-hand experiences. Semantic memory also is productive, permitting extension of knowledge beyond direct experience, through logical processes such as deduction, induction, and analogy (see Goswami, 2011 for review). Without these productive processes, learning would be significantly hindered, as each bit of information is acquired one piece at a time. Moreover, to be maximally effective, productive processes must operate over information learned at different times and in different places, such as over a single semester or an entire course of study. As an example, consider that at some time, you learned that *George Washington was the first president*. Later, you may have learned that *George Washington led the Continental Army during the American Revolution*. Integration of the new information with information already stored in semantic memory permits self-generation of the novel understanding that *the leader of the American Revolution was also the first president*, information never directly specified. Self-generation of new knowledge through integration of separate yet related episodes has been shown to support the long-term accumulation of knowledge in young adults (Bauer & Jackson, 2015; Varga & Bauer, under review) and in preschool and school-age children (Varga & Bauer, 2013; Varga, Stewart, & Bauer, 2016). Yet although memory integration supports expansion of the knowledge base, substantial individual differences are observed in this ability. As such, in the present research we advanced our understanding of the cognitive abilities underlying variability in this productive means of knowledge extension in young adults, as well as relations with educational success.

Self-generation of new factual knowledge through memory integration has logical implications for educational success. As students progress through formal schooling, they are increasingly faced with the challenge of establishing connections between related concepts in order to demonstrate and extend integrative understandings, such as on projects, essays, or exams. Moreover, educators look to psychology to ascertain ways in which our empirical understanding of learning can be applied to enhance academic achievement (see Meltzoff, Kuhl, Movellan, & Sejnowski, 2009 for discussion). In fact, interventions have been employed in science classrooms for the explicit purpose of scaffolding knowledge extension through integration. For example, Davis (2010) demonstrated the efficacy of self-monitoring prompts in helping middle-school students to recognize the opportunity for integration, thereby leading to the formation of new ideas and understandings that were never directly specified, which in turn contributed to higher grades on assignments. If memory integration predicts variability in the capacity to attain and extend school-based knowledge, it is reasonable to conclude that it would also affect one’s potential to excel in educational endeavors more broadly. Furthermore, because educational attainment impacts access to higher education and is a strong predictor of occupational outcome as well as socioeconomic status (e.g., Ceci & Williams, 1997), it is imperative that we identify real-world learning mechanisms that affect academic success.

In the laboratory, progress toward understanding the mechanisms involved in successful knowledge extension through memory integration has been made. A common approach to assessing this learning process in adults involves training individuals to associate overlapping stimulus pairs (e.g., *Chair & Basketball; Basketball & Blender*) in order to infer a novel, indirect relation (e.g., *Chair & Blender*) (Schlichting, Zeithamova, & Preston, 2014; Shohamy & Wagner, 2008; Zeithamova & Preston, 2010; see Varga & Bauer, under review, for a different approach). Critically, arbitrary stimulus pairs are temporally distributed throughout the learning phase, thereby necessitating integration of newly and previously learned information in memory. Successful knowledge extension through integration is measured through a forced-choice test (e.g., *Chair = Blender or House?*). Importantly, vast individual differences in performance are observed in adults, typically ranging from 47% to 90% correct (e.g., Schlichting et al., 2014; Zeithamova & Preston, 2010).

Knowledge extension paradigms such as those just described are invoked as models for the means by which integrated semantic networks are formed (e.g., Preston & Eichenbaum, 2013). Yet training of arbitrary paired associates may not capture the mechanisms involved in the acquisition of real-world factual knowledge that is often the target of classroom learning. Moreover, for memory integration to support the development of a semantic knowledge base, the products must be retained over time. Nevertheless, empirical research employing arbitrary paired associates has focused on the products of memory integration immediately following initial learning, and thus has not provided compelling evidence for the role of memory integration in the long-term accumulation of knowledge, per se.

In contrast to the method of training of arbitrary stimulus pairs, the method adopted in the present research was inspired by a paradigm employed by Bauer and Jackson (2015) in which individuals were taught true but novel, related “stem” facts

(*Apple seeds are called pips; Cyanide is found in pips*) and tested for self-derivation of new factual knowledge through integration of the target information (*Apple seeds contain ___?*). Relative to other inference paradigms, this approach has the advantage of being about real-world, factual information, and thus is directly relevant to the issue of how a semantic knowledge base is built over time. Under the incidental learning conditions of Bauer and Jackson (2015), adults selected the correct answer on 56% of the forced-choice trials (*Range* = 10% to 100%). Performance was not mediated by prior knowledge, as evidenced by performance in a single-stem control condition. When only a single stem fact was provided, performance did not exceed chance levels (27% with 4 choice alternatives). Therefore, individuals successfully derived new knowledge but only through integration of information acquired in the context of the experiment, thus supporting the utility of this novel paradigm.

In an adaptation and extension of the ecological approach, Varga and Bauer (under review) replicated findings of extensive individual differences. That is, successful self-generation of new factual knowledge through memory integration ranged from 3% to 93% correct. Moreover, when adults were tested for retention after 1 week, some loss of information was apparent, yet high levels of recall were still observed. These patterns further substantiate the conclusions that the paradigm captures integrative mechanisms involved in the construction of a semantic knowledge base and is sensitive to individual differences. Thus, in light of the functional consequences that individual differences in knowledge extension through integration likely have for academic outcome, the major purpose of the present research was to examine additional data from Varga and Bauer (under review) to shed light on the association between behavioral task performance and educational success, namely, educational achievement (GPA) and standardized scholastic aptitude (SAT).

Examination of the relation between knowledge extension through integration and educational success has limited usefulness unless we can determine the cognitive abilities associated with variability in this fundamental form of learning. A more detailed understanding of which cognitive factors are important for self-generation of new factual knowledge through integration could enable novel educational interventions. As proposed by Goswami and Szűcs (2011), to gain a fuller understanding of individual differences in educational attainment, investigations of educationally relevant phenomena must prioritize basic research into the underlying factors supporting more complex learning behaviors. In accord with this recommendation, in the current research we examined cognitive factors derived from the two-component theory of intellectual development, specifically, fluid and crystallized cognition (Cattell, 1971; 1987; Horn, 1965, 1968). Fluid abilities support the capacity to, “solve problems, think and act quickly, and encode new episodic memories” and rely heavily on basic information-processing skills. In contrast, crystallized abilities underlie the capacity to accumulate, “verbal knowledge and skills,” and depend heavily on semantic knowledge and verbal skills acquired throughout the course of formal education (Akshoomoff et al., 2013, p. 120). Importantly, although these are clearly separable constructs, fluid abilities have been argued to support the acquisition of crystallized knowledge (Akshoomoff et al., 2013; Horn & Noll, 1997). Moreover, both types of cognition are strong predictors of educational and occupational attainment, such that individuals higher on crystallized and fluid abilities tend to go further in school and to report higher employment status in adulthood (Heaton et al., 2014). Based on these findings, we reasoned that both types of cognition would impact the capacity to accumulate and extend real-world knowledge through memory integration in young adults.

In the present research, we assessed the relation between knowledge extension through integration and several standardized measures of fluid and crystallized cognition (Reynolds & Voress, 2007; Woodcock, 1991; Woodcock, McGrew, & Mather, 2001). Fluid measures of working memory, long-term retrieval, and logical reasoning were chosen due to their hypothesized importance in other forms of flexible knowledge extension (Goswami, 2011; Kyllonen & Christal, 1990). Hypotheses regarding the role of crystallized abilities were less straightforward. The heavily verbal nature of the task, and the demands placed on the abilities to read and comprehend verbal knowledge could lead to important effects of crystallized factors. To examine this possibility, we employed two measures of crystallized intelligence, the first of which reflected reading skills (i.e., reading comprehension) and the second of which reflected the ability to comprehend both individual concepts and the relation between concepts (i.e., verbal comprehension). Finally, we also chose a measure of short-term memory to assess a cognitive ability that serves relatively transient processing needs, and thus should not predict self-generation through integration.

Based on the logical arguments just outlined, we hypothesized that self-generative learning might indirectly relate to measures of academic success. Thus, we explored the relation between self-generation, fluid cognition, crystallized cognition, and academic success. Specifically, we asked whether self-generation and the cognitive constructs identified made independent contributions to predictions of academic success, or whether self-generation was related to the acquisition of knowledge (i.e., crystallized cognition) and the ability to transfer knowledge to new objectives (i.e., fluid reasoning), which in turn were directly associated with educational attainment. Through conducting a more fine-grained analysis of the cognitive factors responsible for individual differences in knowledge extension through integration, as well as the specific cognitive constructs that drive academic success, the present research sheds light on why particular individuals do well or poorly in school, and the specific abilities that must be targeted to improve educational attainment.

Method

Participants

Participants were 117 adults between 18-24 years (*M* = 19.76 years, *SD* = 1.15; 63 females). The sample consisted of individuals whose self-generation data was collected as part of a prior investigation (Varga & Bauer, under review). In the course of the investigation, we also obtained cognitive assessments and academic outcome measures, the results of which were not included in any prior published reports. Here we feature analyses of the cognitive and academic assessments, and their relation with performance on the factual knowledge extension task.

All participants were recruited from a volunteer pool consisting of undergraduate students enrolled in psychology courses at a competitive, private institution. Based on participant self-report, the sample was 9% African American, 25% Asian, 59% Caucasian, and 4% mixed racial descent. Eight percent of the participants were of Hispanic descent. Three participants did not report racial or ethnic information. An additional 3 participants took part in the study but were excluded due to failure to comply with task instructions (*N* = 1) and self-reported diagnosis of Dyslexia which may have negatively impacted task performance (*N* = 2). Cognitive assessments were missing for one participant due to failure to return for the second session in which the measures were collected. Official records of scholastic aptitude (SAT scores) and academic achievement (college GPA) were obtained from the university registrar for participants who authorized release of that information (*N* = 107, 91% of the sample). The protocol and procedures were approved by the university Institutional Review Board. Written informed consent was obtained prior to the start of the study, which also included permission to obtain academic outcome measures. Participants were compensated with course credit for their participation.

Stimuli

The stimuli were 60 stem facts and 30 novel integration facts. Facts ranged from 4-10 words. The facts conveyed true information that was intended to be educationally meaningful and to be unknown to participants prior to the study. For instance, two stem facts were about art history (*A popular sculpture made from a urinal is called Fountain; Duchamp's most well-known work is named Fountain)*. Integration of separate but related stem facts could lend itself to self-generation of a novel integration fact (*Duchamp's most popular work consisted of a urinal*). The stimuli captured a variety of logical relations regularly encountered in everyday learning conditions and which have previously been shown to invoke integration mechanisms (see Zeithamova et al., 2012a for review). Importantly, prior research employing these stimuli has demonstrated that the facts are novel to young adults and that both facts from a given pair are necessary for generation of the target integration facts (see Varga & Bauer, under review; Experiment 2). Specifically, when participants were exposed to only one of the two stem facts from a pair, they produced the novel integration facts only 11% of the time (which significantly differed from the 44% demonstrated when both stem facts were provided). Thus, exposure to the information presented in both stem facts is necessary to reliably derive the corresponding integration fact.

Cognitive Measures

The *Woodcock Johnson Test of Cognitive Abilities, Third Edition* (*WJ-III COG*) (Woodcock, McGrew, & Mather, 2001), the *Test of Memory and Learning, Second Edition* (*TOMAL-2*) (Reynolds & Voress, 2004), and the *Woodcock Language Proficiency Battery—Revised* (*WLPB-R*) (Woodcock, 1991) were used to assess six standardized cognitive domains.

Verbal comprehension. The Verbal Comprehension subtest of the *WJ-III* (Test 1) served as a measure of verbal ability (*Median* reliability = 0.90 from 5-19 years; 0.95 from 20-90 years) (Woodcock et al., 2001). This task consisted of four subtests: Picture Vocabulary, Synonyms, Antonyms, and Analogies, which assessed comprehension of individual words as well as the relationship among words. According to the Cattell-Horn-Carroll (CHC) theory of cognitive abilities (Carroll, 1993; Cattell, 1941; Horn, 1965), scores on Test 1 of the *WJ-III* contribute to the broad factor of crystallized intelligence, reflecting the extent of one's knowledge, the ability to verbally communicate knowledge, and the ability to reason based on previously learned information and experiences. Moreover, this store of linguistically-based declarative knowledge (conceptual understandings) and procedural knowledge (learning skills) is thought to accrue across both educational and general life experiences (Horn & Noll, 1997). Participants received one point for each correctly answered item and the test was discontinued when three items on a page were answered incorrectly. A total score was derived by summing scores across the four individual subtests.

Reading comprehension. The Passage Comprehension subtest of the *WLPB-R* (Test 7) was used to assess reading comprehension (Woodcock, 1991). In this task participants must use syntactic and semantic clues to identify a missing word within a short passage. Using a modified cloze procedure, Test 7 of the *WLPB-R* assesses how well an individual comprehends written discourse as it is being read, requiring both basic reading skills and inferential abilities. Critically, the words in each sentence are designed to be familiar, leading to the reasonable expectation that participants have prior knowledge of the content conveyed in each passage. As such, the task is intended to assess reading ability, independent of verbal comprehension. Participants received one point for each correctly answered item and the test was discontinued when incorrect answers were provided for six consecutive items. A total score was derived by summing the number of correct items.

Reasoning ability. The Concept Formation subtest of the *WJ-III COG* (Test 5) was used as a measure of categorical reasoning based on inductive logic (*Median* reliability = 0.94 from 5-19 years and 0.96 from 20-90 years) (Woodcock et al., 2001). In this task individuals are shown a stimulus set (i.e., a series of shape and color patterns) and are required to derive the rule that governs each sequence. Because there is no memory component, this task contributes to the broad CHC factor of fluid reasoning. That is, this test reflects a mixture of mental operations including the general abilities to reason, generate concepts, and solve novel problems in the absence of prior knowledge. Participants received one point for each correctly answered item and were provided with corrective feedback throughout task administration. A total score was derived by summing the number of correct items.

Long-term retrieval. The Visual-Auditory Learning (VAL) subtest of the *WJ-III COG* (Test 2) served as a measure of associative memory (*Median* reliability = 0.86 from 5-19 years; 0.91 from 20-90 years) (Woodcock et al., 2001). In this task participants are shown a series of rebuses (pictographic symbols of words) and later asked to recall the visual-auditory associations from long-term memory. Specifically, once the rebuses are learned, participants are presented with several rebuses forming a sentence and are asked to speak the associated words aloud. Because this task requires the ability to store information and to retrieve it later in the process of thinking, Test 2 is argued to reflect the broad CHC factor of long-term retrieval. Importantly, this factor is not to be confused with one's long-term memory store which constitutes the *contents* of knowledge. Instead, this task taps the general processing abilities of storing and retrieving information from long-term memory. Participants received one point for each incorrectly answered item, defined as a failure to identify the correct word or to do so within 5 seconds of viewing a rebus. The correct word was provided if participants failed to state it within the 5-second time limit. Because scoring was conducted on-line, an independent coder listened to all audio recordings to ensure that the 5-second pause was reliably scored. If participants were allotted more than 5 seconds, the item was subsequently scored as incorrect. If participants were corrected too soon, the item was counted as a missing trial (0.59% of trials). A proportion score was then derived by dividing the total number of errors by the number of valid trials. Due to reliance on the number of errors, this score should negatively correlate with other variables.

Short-term memory. The Digits Forward subtest of the *TOMAL-2* was used as a traditional measure of short-term memory span. This task requires the individual to hold a sequence of numbers in immediate awareness before repeating them back to the experimenter, with the sequence length increasing throughout task administration. One point was awarded for the correct recall of each digit within the serial position in which it was presented. The task was discontinued when participants recalled three or fewer digits on two consecutive sequences. A total score was derived by summing the number of correctly recalled digits.

Working memory. The Digits Backward subtest of the *TOMAL-2* was used as a traditional measure of working memory span. In this task the individual must hold a sequence of numbers in immediate awareness while performing a mental operation on it (i.e., reversing the sequence). Task administration and scoring was conducted in the same manner as in the Digits Forward subtest with the exception that the digits must be correctly placed in the reverse order.

Academic Measures

SAT. The SAT is a standardized college admissions test which assesses academic readiness for college. The test primarily measures knowledge and skills learned in school, however, some items also assess aspects of fluid intelligence. The verbal section requires reading of passages and sentence completions, and taps skills such as determining the meaning of words, text comprehension, inferential reasoning, organization of ideas, and understanding literary elements (*Max Score* = 800). The quantitative section includes questions on arithmetic operations (e.g., fractions), algebra, geometry, statistics, problem solving, and reasoning (*Max Score* = 800). In the event that participants took the exam multiple times, the highest scores on each section were obtained and summed to derive a total SAT score (*Max Score* = 1600).

GPA. Grades from all college courses were averaged to derive a measure of academic achievement. Because grades are assumed to be based on some criterion level of performance, college GPA reflects the degree to which participants mastered specific course content. Average GPAs were obtained at the end of the semester in which individuals participated (*Max* = 4.00).

Procedure

Participants completed two sessions spaced 1 week apart (*M* delay = 6.95, *SD* = 0.66, *Range* = 6-11 days). Participants were tested individually by one of two female experimenters (including the first author), each of whom tested an approximately equal number of participants from each gender. With the exception of six individuals, participants were tested by the same experimenter at both sessions. The experimenters followed the same detailed written protocol and regularly reviewed audio-recorded sessions to ensure protocol fidelity.

Session 1: Initial learning and extension of knowledge. At the start of the session, participants were instructed that we were interested in whether memory for newly learned factual information differs as a function of its subject domain. Participants read a total of 60 sentences (i.e., the 30 pairs of related stem facts). To equate total reading time across participants, sentences were presented one word at a time for 400ms. Each sentence ended in a target word, which served as the relational link between to-be-integrated stem facts. At the end of each sentence, participants were shown a decision screen and asked to indicate, via a button-press response, whether the information conveyed was novel or known. The incidental task was designed to ensure that participants were attending to the facts while also corroborating the pretext of the study purpose (i.e., learning of *novel* information). At no time were participants informed that any of the sentences were related. Moreover, because prior knowledge of the stem facts was not correlated with behavioral self-generation performance, we did not exclude from analysis facts of which participants reported prior knowledge (see Varga & Bauer, under review; Experiment 3).

Across the encoding presentation, to-be-integrated stem facts were separated by a lag of 2 to 4 intervening sentences. Variability in lag created temporal distance between to-be-integrated information and also prevented participants from anticipating the content of the next fact. Moreover, the short lag served to promote recognition of the relatedness between paired facts, thereby allowing for examination of knowledge integration under optimal learning conditions. Additionally, fact order was counterbalanced such that each stem fact from a pair was presented in the first or second serial position an approximately equal number of times across the sample.

After a short break lasting 5-10 minutes in which participants completed several surveys, participants were presented with 30 facts derived through integration of the previously presented stem facts. The sentences were presented in the form of questions by omitting the final word of each fact. Participants were asked to generate a one-word answer that could accurately fill in the blank space. When an answer was generated, participants made a button-press response which was followed by an “Answer” screen cueing them to speak the answer aloud.

Session 2: Standardized cognitive assessments. Participants returned to the laboratory approximately one week after their initial visit. Participants were instructed that we were interested in whether performance on a number of cognitive tasks related to performance at the initial session. Standardized cognitive assessments were then administered in the following fixed order: (1) short-term memory (Digits Forward), (2) fluid reasoning (Concept Formation), (3) working memory (Digits Backward), (4) long-term retrieval (VAL), (5) reading comprehension, and (6) verbal comprehension. To avoid the potential for unique order effects across the sample, we did not counterbalance the sequence of cognitive assessments, thereby reducing error and increasing power. Following completion of these tasks which lasted approximately 40 minutes, memory for the integration facts was assessed (see Varga & Bauer, under review; Experiment 3).

Results

We explored associations between initial knowledge extension, standardized cognitive abilities, and academic success. As a first step in the process, we described variability in self-generation through integration performance. Second, we assessed whether standardized cognitive abilities differentially related to self-generation performance by conducting a multiple regression analysis. Third, we examined the relative contributions of self-generation and cognitive factors on academic success by conducting a series of linear regression and mediation analyses. To ensure that the assumptions of linear regression were met for each model, casewise diagnostics were examined to identify outliers with undue influence on the model (i.e., standardized residuals, DFBeta statistics, Cook's distance, leverage statistics, and Mahalanobis distance). We also verified that the residuals were independent (i.e., Durban-Watson), homoscedastic, and normally distributed. In cases in which the assumptions underlying certain statistical tests were violated, steps taken to further explore and/or correct for these issues are described.

Description of Self-generation Behavior

Participants received a score of 1 or 0 for each trial, indicating whether they were successful or unsuccessful at self-generating the novel integration fact, respectively. A proportion score was derived by dividing the total number of successfully self-generated facts by the number of possible trials. Descriptive statistics, including measures of central tendency and variability for self-generation performance and for all other target variables, are reported in Table 1. On average, participants self-generated the integration facts on 50% of the trials (*M* Proportion score = 0.50; *SD* = 0.21). Substantial variability was observed with performance ranging from 3% to 93% correct across the sample. As reflected in Figure 1, self-generation scores were approximately normally distributed.

Relations between Self-generation and Standardized Cognitive Abilities

A primary aim of the present research was to determine the cognitive abilities that are associated with self-generation of new factual knowledge through memory integration. Raw Pearson correlation coefficients among all standardized cognitive abilities are reported in Table 2. As hypothesized, with the exception of short-term memory, significant intercorrelation was evident across all cognitive factors. It is important to highlight that the measure of short-term memory was significantly skewed, whereas measures of fluid reasoning and long-term retrieval exhibited both skew and kurtosis (see Table 1). Outliers were removed in the case of short-term memory, whereas logarithmic transformations were applied to the reasoning and long-term retrieval variables to correct for non-normally distributed data. Comparison of raw and transformed data did not change the overall pattern of results, with the exception that the significant correlation between fluid reasoning and working memory fell below the level of statistical significance when the fluid reasoning variable was transformed (p = .07). Based on largely similar results, all subsequent analyses were conducted using the raw data.

Individually, all measures of standardized cognitive abilities were associated with self-generation (see Table 2). To directly test whether the standardized cognitive abilities predicted unique statistical variance in knowledge extension through integration, we conducted a multiple regression analysis. Based on the high intercorrelation amongst variables, we examined collinearity statistics (i.e., VIF and tolerance) and the eigenvalues of the scaled, uncentred cross-product matrix to ensure that the variance of each predictor loaded to a different dimension (i.e., a different eigenvalue). Consistent with the assumption that the standardized cognitive factors constituted identifiably distinct constructs, collinearity amongst variables was not an issue. When all six cognitive abilities were entered into the model via the forced entry method, the full model explained a significant proportion of variance in self-generation, R^2 = .367, F(6, 100) = 9.68, p < .001. Regression coefficients and standard errors are presented in Table 3. Only verbal comprehension, t(100) = 3.59, p < .001, and working memory, t(100) = 2.27, p = .03, were significantly associated with successful self-generation through integration, with verbal comprehension explaining more variance than working memory (R^2= .205 and .137, respectively). Thus, in our final set of analyses we explored the differential contribution of these cognitive factors and self-generation in predicting statistical variability in academic success.

Relations amongst Self-generation, Cognitive Abilities, and Academic Success

Given that participants were sampled from a selective private institution, it is not surprising that average SAT scores (*M* = 1340.82; *SD* = 114.29) and GPAs (*M* = 3.46; *SD* = 0.45) were high, though importantly, variability was still observed (Table 1). Pearson correlations among self-generation performance and academic success are reported in Table 2. As predicted, the capacity to extend knowledge through memory integration was significantly positively associated with standardized measures of scholastic aptitude (total SAT) and academic achievement (college GPA). Whereas the relation between self-generation and SAT scores was of moderate strength, the relation between self-generation and GPA was small. Despite this difference in effect size, the correlation between self-generation performance and academic outcome was significant for both measures ($p < .01$). There also was significant skew and kurtosis evident in the univariate distribution of GPAs (Table 1). To address the violations of the assumption of normality, we applied an inverse transformation to the raw GPA data which resulted in non-significant skewness (Skew = -0.06; $SE = 0.23$; $Z = -0.24$; $\alpha = .05$) and kurtosis (Kurtosis = -0.64; $SE = 0.46$; $Z = -1.39$; $\alpha = .05$). When we compared raw and transformed GPA scores, relations to self-generation were significant in both cases, though the relation was less striking using the transformed distribution ($p < .01$ vs. $p < .05$). Because mediation analyses assume normally distributed residuals, inverse-transformed GPA scores were used to correct for this violation when examining the GPA mediating frameworks in subsequent analyses.

We next conducted a series of empirically-driven regression analyses to determine whether a mediating variable (e.g., standardized cognitive factors) carried the influence of another variable (e.g., self-generation) on academic outcome (i.e., SAT and GPA) (see

MacKinnon, Fairchild, & Fritz, 2007 for discussion of the utility of mediating frameworks in psychological research). A variable can only serve as a mediator if it has a significant unique effect on the dependent variable in the presence of the independent variable. Thus, we first conducted four separate linear regression analyses to determine: (1) whether self-generation and verbal comprehension uniquely predicted variance in SAT; (2) whether self-generation and verbal comprehension uniquely predicted variance in GPA; (3) whether self-generation and working memory uniquely predicted variance in SAT; and (4) whether self-generation and working memory uniquely predicted variance in GPA. Results from these preliminary analyses indicated that only the standardized cognitive constructs uniquely predicted variance in academic outcome when self-generation was included, thus indicating that these factors could serve as mediating variables (and that self-generation could not serve as a mediator). As such, in subsequent analyses we tested whether verbal comprehension and working memory statistically mediated the relation between self-generation through integration and academic success.

We first tested whether verbal comprehension mediated the relation between knowledge extension and SAT. As Figure 2 (Panel A) indicates, self-generation and verbal comprehension were significantly associated, $t(80) = 5.49$, $p < .001$, $R^2 = .274$, as were verbal comprehension and SAT while the effect of self-generation was controlled, $t(80) = 2.88$, $p =.005$. Together, verbal comprehension and self-generation accounted for approximately 19% of the variance in scholastic aptitude ($R^2 = 0.19$). Further, consistent with mediation, the standardized total effect of self-generation on SAT [$\beta = 0.33$, $t(80) = 9.76$, $p = .002$] was reduced upon the addition of verbal comprehension to the model [$\beta = 0.15$, $t(79) = 1.27$, $p = 0.21$]. However, the standardized direct effect of self-generation on

SAT remained a nonzero coefficient ($\beta = 0.15$), indicating that verbal comprehension accounted for some, but not all, of the relation. Indeed, the product of the standardized coefficients indicated that the amount of mediation (i.e., standardized indirect effect) was (0.52)(0.34) = 0.18 (out of 0.33 total). Results of the Sobel test confirmed that the partial mediation model was statistically significant, $Z' = 2.55$, $p = 0.01$. Thus, self-generation was significantly associated with scholastic performance, but some of the relation was accounted for by the influence of self-generation on verbal comprehension.

We next tested whether working memory mediated the relation between knowledge extension and SAT. As depicted in Figure 2 (Panel B), self-generation and working memory were significantly associated, $t(82) = 4.44$, $p < .001$, $R^2 = .194$, as were working memory and SAT when the effect of self-generation was controlled, $t(81) = 2.53$, $p = 0.01$. Approximately 18% of the variance in SAT score was accounted for by self-generation and working memory ($R^2 = 0.178$). Moreover, consistent with partial mediation, the effect of self-generation on SAT [$\beta = 0.33$, $t(82) = 3.23$, $p = .002$] was reduced (though to a nonzero coefficient) upon the addition of working memory [$\beta = 0.21$, $t(81) = 1.88$, $p = 0.06$]. The product of the standardized coefficients indicated that the amount of mediation (i.e., standardized indirect effect) was (0.44)(0.28) = 0.12 (out of 0.33 total). Results of the Sobel test confirmed that working memory partially mediated the relation between knowledge extension and SAT, $Z' = 2.20$, $p = 0.03$. Thus, a portion of the association between self-generation and GPA was accounted for by working memory.

Third, we tested whether verbal comprehension similarly mediated the relation between knowledge extension and GPA. As reflected in Figure 3 (Panel A), self-generation and verbal comprehension were associated, $t(102) = 6.13$, $p < .001$, as were verbal comprehension and GPA while controlling for self-generation, $t(101) = 2.95$, $p = 0.004$. Together, approximately 15% of the variance in GPA was accounted for by self-generation and verbal comprehension ($R^2 = 0.149$). Additionally, the effect of self-generation on GPA [$\beta = 0.28$, $t(102) = 2.89$, $p = 0.005$] was reduced (though to a nonzero coefficient) upon the addition of verbal comprehension [$\beta = 0.11$, $t(101) = 1.03$, $p = 0.31$], thus consistent with partial mediation. The product of the standardized coefficients indicated that the amount of mediation was (0.52)(0.32) = 0.16 (out of a total effect of 0.27). Results of the Sobel test confirmed that verbal comprehension mediated the effect of knowledge extension on GPA, $Z' = 2.55$, $p = 0.01$, thus indicating that a portion of the association between self-generation and GPA was accounted for by verbal comprehension.

Although working memory mediated the relation between self-generation and SAT, it did not mediate the relation between self-generation and GPA. Specifically, as depicted in Figure 3 (Panel B), working memory was not significantly association with GPA, $t(103) = -0.13$, $p = 0.89$. Therefore, working memory could not mediate the effect between knowledge extension performance and participants' academic achievement in college.

General Discussion

The goals of the current research were three-fold. First, we identified the cognitive factors that contribute to individual differences in extension of knowledge through integration. Second, we tested whether the ability to accrue knowledge through memory integration is associated with educational attainment, a relation which has previously been assumed but not directly examined. Third, we assessed whether self-generation and the fluid and crystallized constructs identified made independent contributions to statistical predictions of academic success, or whether a mediational framework was evidenced.

The findings of the research were clear. Individual differences in this fundamental form of knowledge extension were jointly explained by verbal comprehension and working memory. Moreover, as anticipated, a bidirectional relation between these cognitive factors and successful knowledge extension was observed. That is to say, the capacity to self-generate new knowledge through memory integration also statistically predicted variance in standardized measures of verbal comprehension and working memory. Finally, self-generative learning was associated with variability in scholastic aptitude (SAT) as well as educational achievement (GPA). Mediational analyses revealed that the association between self-generation and SAT was partially attributed to the association between self-generation and verbal comprehension (i.e., crystallized knowledge) and working memory (i.e., fluid cognition), respectively. Much like SAT, the relation between self-generation and college GPA was also partially mediated by the relation between self-generation and verbal comprehension. However, unlike SAT, college GPA was not associated with working memory. Interpretations and implications of these findings are discussed in turn.

The current research advances our theoretical understanding of the cognitive abilities that contribute to variability in the capacity to extend knowledge through memory integration. More than a third of the variability in self-generation performance was accounted for by standardized measures of crystallized and fluid cognition.

Specifically, whereas verbal comprehension accounted for approximately 21% of the variance in self-generation performance, working memory accounted for an additional 14% of the variation. Further explication of the association between these cognitive factors and semantic knowledge extension requires careful consideration of the specific measures employed. In the knowledge extension task, the behavior assessed was self-derivation of new factual knowledge based on integration of previously learned information (i.e., stem facts). Of note, individuals did not possess knowledge of the stem facts or of the integration facts prior to their participation in the study (Varga & Bauer, under review). It is therefore reasonable to suggest that self-generative learning relied on a culmination of skills, including comprehension of the meaning conveyed by each newly learned stem fact, identification of the relation between to-be-integrated stem facts, flexible manipulation of the paired stem facts, and verbal communication of the self-generated integrative understanding. Thus, because successful knowledge extension necessarily required both comprehension of the semantic information, as well as the flexible use of that information, it is not altogether surprising that verbal abilities and working memory capacity explained a significant portion of the variance in performance.

To clarify how verbal comprehension and working memory might directly contribute to successful knowledge extension, it is necessary to unpack the specific abilities that these cognitive measures purport to measure. In the case of verbal comprehension, participants were asked to provide verbal labels for pictures (e.g., tourniquet), and to demonstrate relational knowledge through completion of synonyms (e.g., untamed: wild), antonyms (e.g., ancient: modern), and analogies (e.g., wrist is to shoulder, as ankle is to hip). Therefore, this task assessed both the extent of one’s crystallized knowledge, as well as the ability to reason based on acquired knowledge. Although participants were required to learn and extend *novel* semantic material in the context of the present research, it is reasonable to suggest that well-developed comprehension skills would facilitate one’s ability to derive meaning from the newly learned stem facts, as well as the ability to integrate the relation between to-be-integrated stem facts in semantic memory.

We can also begin to understand the differential contribution of working memory to successful self-generation through close examination of the digits backward measure employed. In this task, participants were asked to reverse a sequence of auditorily presented numbers, with the length of each sequence increasing from 2-9 digits across administration. Thus, individuals who performed well on this task were those who could hold many items in immediate awareness at once while also mentally transforming the mental representation. Likewise, in the knowledge extension paradigm participants were similarly required to attend to separate yet related stem facts concurrently in order to integrate them at the time of encoding and/or to derive the novel integration fact at the time of test. In light of the proposed role of working memory in the explicit manipulation of the stem facts during self-generation, it might seem puzzling that our measure of concept formation (which was specifically designed to assess fluid reasoning abilities) did not account for unique variance in successful knowledge extension. Yet in the concept formation task participants were presented with novel patterns for an unlimited amount of time, thereby eliminating the need to hold items in memory. Conversely, the digits backward task and the self-generation task currently employed required a combination of memory and reasoning abilities. Indeed, consistent with our interpretation of the role of working memory, Cowan (2014) argues that one’s sophistication in reasoning about any real-world problem depends on working memory capacity, which is defined as the amount of relevant information that an individual can cull from long-term memory concurrently while performing, “any combination of mental strategies and processes,” that may be used to maintain and/or transform the representation (p. 207). Taken together, we argue that greater working memory capacity facilitated the amount of stem fact information that an individual could simultaneously activate and process in the service of transforming the long-term representations into an explicit, integrated understanding.

It is important to emphasize that the knowledge extension task and standardized cognitive measures employed in the present research were administered one week apart. Consequently, the regression analyses provide important insight into the extent to which verbal comprehension and working memory are associated with variability in self-generation, but they cannot be interpreted as establishing cause-and-effect. Indeed, as suggested by the mediation models, a reciprocal pattern was observed between these cognitive factors and successful knowledge extension through integration. Specifically, self-generation through integration also predicted 27% and 19% of the statistical variance in verbal comprehension and working memory capacity, respectively. Based on converging evidence that knowledge newly derived through memory integration is retained in semantic memory over time (e.g., Varga et al., 2016; Varga & Bauer, 2013), we would expect that one’s propensity for self-generative learning would contribute to the breadth and depth of accumulated knowledge (i.e., verbal comprehension). In addition, it is reasonable that knowledge extension, at least as it was assessed here, would also predict working memory. Working memory (e.g., digit span) has been consistently shown to exhibit correlations from .80 to .90 with a numerous measures of flexible knowledge extension (e.g., analogies, verbal reasoning), leading some to argue that working memory is akin to reasoning ability (e.g., Engle, Ruholski, Laughlin, & Conway, 1999). Moreover, recent findings from behavioral (Bauer, Varga, King, Nolen, & White, 2015; Varga & Bauer, 2013) and neuroimaging (Shohamy & Wagner, 2008; Zeithamova, Dominick, & Preston, 2012b) research provide converging evidence that knowledge extension through integration is characterized by integration of related episodes during initial learning followed by flexible manipulation of that information in response to a demand at the time of test. Of relevance to the present research, individual differences in the capacity to extend knowledge are linked to whether participants integrate separate yet related information in memory prior to an explicit test (Zeithamova & Preston, 2010). Thus, because individual differences in self-generation in the present research were likely associated with how readily participants integrated at the time of learning, our behavioral measure necessarily encompasses integrative encoding abilities, which in turn would predict variability in constructs that reflect broad reasoning ability (i.e., working memory).

Our account of the reciprocal relation between knowledge extension through integration and crystallized and fluid abilities, respectively, also sheds light on our understanding of the association between self-generation and academic success. In the present research, we gathered information on participants success on measures of scholastic aptitude prior to college admittance (SAT) and college achievement at the end of the semester of participation (GPA). Importantly, there was reason to believe that our self-generation task would not contribute directly to performance on these academic measures. That is to say, whereas our task assessed whether participants could acquire, integrate, and flexibly extend *new* information, the SAT and GPA assess acquisition of *known* information that participants should have acquired in high school and college classes. Therefore, we hypothesized that the capacity to self-generate through integration would be associated with a more extensive semantic knowledge base, which in turn would relate directly to aptitude and achievement measures. Consistent with this prediction, the present results indicated that together, self-generation and verbal comprehension account for 19% and 15% of the statistical variation in SAT and GPA, respectively. Moreover, although self-generation was significantly associated with these measures of academic success, this relation was statistically mediated by the association between self-generation and verbal comprehension, which in turn contributed to the statistical prediction of academic success. Again, though our measures do not allow us to draw a causal explanation within this sample, the significance of the statistical mediation models conducted contributes to our theoretical understanding of the process by which the ability to self-generate might indirectly contribute to educational success, namely, through supporting the accrual of an extensive semantic knowledge base.

The mediation analyses conducted in the present research also contributed valuable insight into our understanding of the association between self-generation, fluid cognition, and academic success. Together, self-generation and working memory accounted for 18% of the variance in SAT. Further, much like verbal comprehension, working memory statistically mediated this association. Yet in contrast to verbal comprehension, working memory was not associated with participants' GPAs. A clear explanation of this seemingly disparate pattern of results comes from closer examination of the specific skills assessed on tests of aptitude (SAT) versus achievement (GPA). It has long been established that crystallized intelligence is a better predictor of academic achievement than is its fluid counterpart (Kaufman, Kamphaus, & Kaufman, 1985; Kunina, Wilhelm, Formazin, Jonkmann, & Schroeders, 2007). The reason for this is that measures of academic achievement (e.g., GPA) often reflect recent classroom learning, usually in regard to a much narrower scope (e.g., exams taken over the course of a semester) relative to standardized tests that measure one's capacity to master school tasks more broadly (e.g., math, reading, etc.). With respect to the specific aptitude measure reported here, to achieve the maximum score on the SAT, participants must not only demonstrate knowledge acquired in school, but also the ability to engage in inferential reasoning using that knowledge. Because the verbal and quantitative sections of the SAT undoubtedly require a combination of fluid and crystallized abilities (Engle et al., 1999), the association with working memory likely reflects one's fluid ability to use prior knowledge flexibly in the face of new objectives. On the other hand, the GPA measures obtained in the present research were less likely to reflect fluid reasoning skills. Indeed, the individuals tested in the present research were drawn from introductory psychology courses meaning that many participants were in their first or second year of college. As such, it is reasonable to suggest that the GPAs reflected success in building the basic knowledge of a major, rather than more abstract learning skills assessed in upper-level courses. In sum, the capacity to self-generate knowledge through integration was associated with fluid reasoning (i.e., working memory), which in turn was associated with the ability to use prior knowledge flexibly to perform well on the SAT, a skill that is less essential to attaining a high GPA at the outset of college.

It is important to acknowledge that there were several limitations to the present research. First and foremost, because the measures analyzed here were concurrent rather than longitudinal, the present investigation does not provide insight into the cause-and-effect relation between self-generation, fluid and crystallized cognition, and academic success. Indeed, because only fluid reasoning predicts variability in self-generation through integration in the early school years (Varga & Bauer, 2014), it will be important to determine how the relation between this learning ability and specific cognitive factors changes over the course of formal schooling. Second, we recognize that the cognitive factors assessed in the present research accounted for only 19% of the variance in SAT, 15% of the variance in GPA, and 37% of the variance in self-generation through integration. Therefore, other factors such as motivation and personality characteristics that lead some individuals to make more of an effort in school are likely as important in explaining differences in academic performance as the cognitive factors examined here. Thus, additional research aimed at delineating other critical determinants of knowledge extension and academic successful is clearly warranted.

In conclusion, the present research provides new insights into the cognitive abilities that contribute to variability in the productive extension of semantic knowledge in adulthood. It makes clear that both verbal comprehension skills and working memory capacity underlie variability in this fundamental form of learning. Moreover, self-generation through integration also bidirectionally predicted statistical variance in crystallized knowledge and fluid reasoning factors, which in turn were associated with measures of academic success. These findings are important because they suggest that interventions aimed at promoting successful memory integration might have implications for both fluid and crystallized cognition, and might therefore have the potential to facilitate academic success. Along a similar vein, because crystallized abilities are thought to accrue across educational experiences and are less contingent on rate-limiting biological determinants than are fluid abilities (Horn & Noll, 1997), the present research paints an optimistic picture for future research aimed at promoting this crucial learning ability. Therefore, taken together, the current research provides a theoretically plausible and a practically significant framework from which to guide future research aimed at enhancing this educationally relevant phenomenon.

References

Akshoomoff, N., Beaumont, J. L., Bauer, P. J., Dikmen, S. S., Gershon, R. C., Mungas, D., Slotkin, J., Tulsky, D., Weintraub, S., Zelazo, P. d., & Heaton, R. K. (2013). NIH toolbox cognition battery (CB): Composite scores of crystallized, fluid, and overall cognition. In P. D. Zelazo & P. J. Bauer (Eds.), *National institutes of health toolbox cognition battery (NIH toolbox CB): Validation for children between 3 and 15 years* (pp. 119-132). Monographs for the Society for Research in Child Development.

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), 271-282. doi: http://dx.doi.org/10.1037/a0037405

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. doi: 10.1080/15248372.2013.849707

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* New York: Cambridge University Press.

Cattell, R. B. (1941). Some theoretical issues in adult intelligence testing. *Psychological Bulletin, 38,* 592.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action.* Boston: Houghton Mifflin.

Cattell, R. B. (1987). *Intelligence: Its structure, growth and action.* Amsterdam: North-Holland.

Ceci, S. J., & Williams, W. M. (1997). Schooling, intelligence, and income. *American Psychologist, 52*, 1051-1058.

Cowan, N. (2014). Short-term and working memory in childhood. In P. J. Bauer & R. Fivush (Eds.), *Handbook on the development of children's memory* (pp. 202-229). Oxford, UK: Wiley-Blackwell.

Davis, E. A., (2000). Scaffolding students' knowledge integration: Prompts for reflection in KIE. *International Journal of Science Education, 22(8),* 819-837.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309-331.

Goswami, U. (2011). Inductive and deductive reasoning. In U. Goswami (Ed.), *Childhood cognitive development* (pp. 399-419). Wiley-Blackwell.

Goswami, U., & Szűcs, D. (2011). Educational neuroscience: Developmental mechanisms: Towards a conceptual framework. *Neuroimage*, *57*(3), 651-658.

Heaton, R. K., Akshoomoff, N., Tulsky, D., Mungas, D., Weintraub, S., Dikmen, S., … & Gershon, R. (2014). Reliability and validity of composite scores from the NIH toolbox cognition battery in adults. *Journal of the International Neuropsychological Society*, *20*(06), 588-598.

Horn, J. L. (1965). *Fluid and crystallized intelligence.* Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign.

Horn, J. L. (1968). Organization of abilities and the development of intelligence. *Psychological Review, 75,* 242-259.

Horn, J. L., & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In D. P. Flanagan & J. L. Genshaft (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 53-91). New York: Guilford.

Kaufman, A. S., Kamphaus, R. W., & Kaufman, N. L. (1985). New directions in intelligence testing: The Kaufman Assessment Battery for Children (K-ABC). In B. B. Wolman (Ed.), *Handbook of intelligence* (pp. 663-698). New York: Wiley.

Kunina, O., Wilhelm, O., Formazin, M., Jonkmann, K., & Schroeders, U. (2007). Extended criteria and predictors in college admission: Exploring the structure of study success and investigating the validity of domain knowledge. *Psychology Science*, *49*(2), 88-114.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence, 14*, 389-433.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology, 58*, 593-614.

Doi: 10.1146/annurev.psych.58.110405.085542

Meltzoff, A. N., Kuhl, P. K., Movellan, J., & Sejnowski, T. J. (2009). Foundations for a new science of learning. *science*, *325*(5938), 284-288.

Preston, A. R., & Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. *Current Biology*, *23*(17), R764-R773.

Reynolds, C. R., & Voress, J. (2007). Test of Memory and Learning (TOMAL-2). *TX*: PRO-ED.

Schlichting, M. L., Zeithamova, D., & Preston, A. R. (2014). CA1 subfield contributions to memory integration and inference. *Hippocampus, 24*(10), 1248-1260.

Shohamy, D., & Wagner, A. D. (2008). Integrating memories in the human brain: Hippocampal-midbrain encoding of overlapping events. *Neuron, 60*(2), 378-389. doi:10.1016/j.neuron.2008.09.023

Sweegers, C. C. G., Takashima, A., Fernández, G., Talamini, L. (2014). Neural mechanisms supporting the extraction of general knowledge across episodic memory. *NeuroImage, 87*, 138-146.

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. (2014). Conceptual knowledge extension: An examination of its development and the underlying cognitive processes involved. In R. Chen (Ed.), *Cognitive development: Theories, stages & processes and challenges* (pp. 1-16). Huntington, NY: Nova Science Publishers.

Varga, N. L., & Bauer, P. J. (under review). Knowledge extension through memory integration: Factors and conditions that promote the long-term accumulation of knowledge.

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. doi: 10.1016/j.jecp.2015.11.015

Woodcock, R. W. (1991). *Woodcock Language Proficiency Battery-Revised: English and Spanish Forms.* Itasca, IL: Riverside Publishing.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Cognitive Abilities*. Itasca, IL: Riverside Publishing.

Zeithamova, D., Dominick, A. L., & Preston, A. R. (2012b). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron, 75*(1), 168-179.

Zeithamova, D., & Preston, A. R. (2010). Flexible memories: Differential roles for medial temporal lobe and prefrontal cortex in cross-episode binding. *The Journal of Neuroscience, 30*(44), 14676-14684.

Zeithamova, D., Schlichting, M. L., & Preston, A. R. (2012a). The hippocampus and inferential reasoning: building memories to navigate future decisions. *Frontiers in Human Neuroscience*, *6*, 1-14.

Table 1.

Descriptive Statistics for All Measures

Measure	*N*	*M*	*SD*	*Range*	Skew	SE_{Skew}	Z_{Skew}	Kurtosis	SE_{Kurtosis}	Z_{Kurtosis}
SG-Int	117	0.50	0.21	0.03-0.93	-0.10	0.22	-0.45**	-0.97	0.44	-2.20*
SAT	85	1340.82	114.29	1010-1560	-0.37	0.26	-1.43**	-0.29	0.52	-0.56**
GPA	107	3.46	0.45	1.97-4	-1.32	0.23	-5.64	1.72	0.46	3.72
Verbal	114	57.92	4.17	46-67	-0.12	0.23	-0.52**	-0.17	0.45	-0.38**
Reading	116	35.31	2.88	29-42	-0.05	0.23	0.21**	-0.49	0.45	-1.11**
Reasoning	116	35.72	3.62	23-40	-1.18	0.23	-5.26	1.30	0.45	2.92
LT-Retrieval	111	0.08	0.06	0-0.28	1.28	0.23	5.57	2.04	0.46	4.48
STM	113	58.58	12.04	27-81	-0.69	0.23	-3.04	0.46	0.45	1.03**
WM	116	35.28	12.88	13-64	0.13	0.23	0.56**	-1.05	0.45	-2.36*

Note. SG-Int = proportion of successfully self-generated integration facts; SAT = Scholastic Aptitude Test; GPA = grade point average; Verbal = verbal comprehension; Reading = reading comprehension; Reasoning = fluid reasoning; LT-Retrieval = long-term memory retrieval; STM = short-term memory; WM = working memory.

* Non-significantly different from normal at $p < .01$
** Non-significantly different from normal at $p < .05$

Table 2.

Correlation Matrix for all Measures

Measure	SG-Int	SAT	GPA	Verbal	Reading	Reasoning	LT-Retrieval	STM	WM
SG-Int	--								
SAT	.31**	--							
GPA	.25**	.45**	--						
Verbal	.54**	.42**	.37**	--					
Reading	.45**	.26*	.20*	.57**	--				
Reasoning	.33**	.34**	.25**	.34**	.29**	--			
LT-Retrieval	-.32**	-.20	-.20*	-.44**	-.27**	-.27**	--		
STM	.19*	.34**	.20*	.17	.29**	.001	-.15	--	
WM	.35**	.38**	.09	.26**	.37**	.19*	-.37**	.42**	--

Note. SG-Int = proportion of successfully self-generated integration facts; SAT = Scholastic Aptitude Test; GPA = grade point average; Verbal = verbal comprehension; Reading = reading comprehension; Reasoning = fluid reasoning; LT-Retrieval = long-term memory retrieval; STM = short-term memory; WM = working memory.

* $p < .05$
** $p < .01$

Table 3.

Summary of Multiple Regression Analysis Explaining Variability in Self-generation

Predictor	*B*	*SE B*	β	95% CI
Verbal	0.019	0.005	0.368**	0.008, 0.029
Reading	0.009	0.007	0.122	-0.006, 0.023
Reasoning	0.007	0.005	0.122	-0.003, 0.017
LT-Retrieval	-0.020	0.350	-0.005	-0.714, 0.674
STM	<0.001	0.002	-0.003	-0.003, 0.003
WM	0.004	0.002	0.217*	0.000, 0.007

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

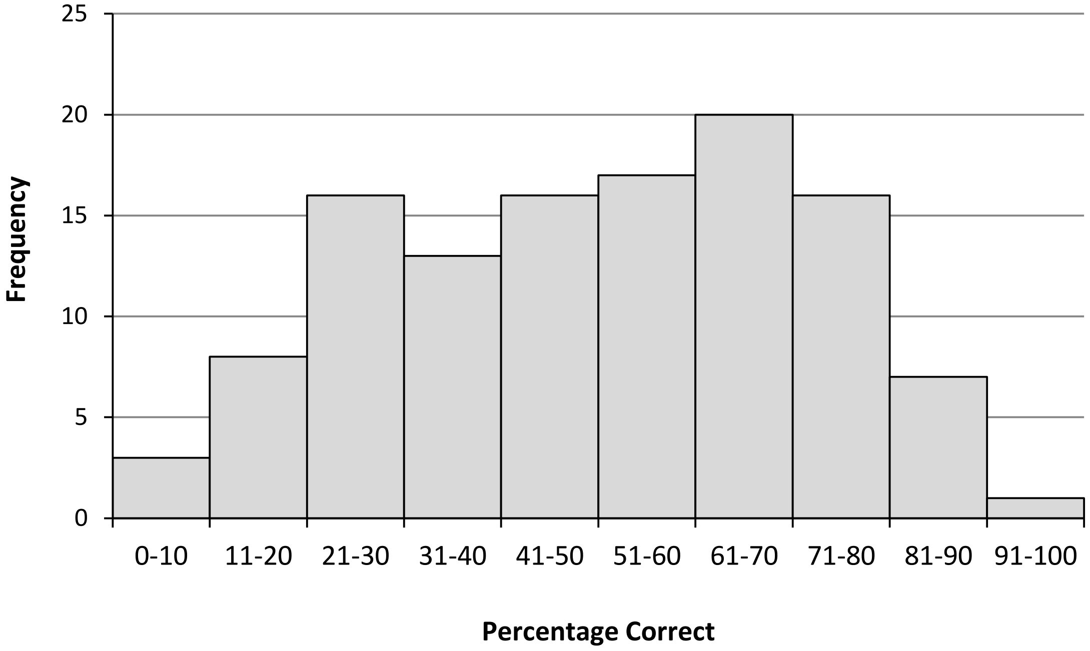
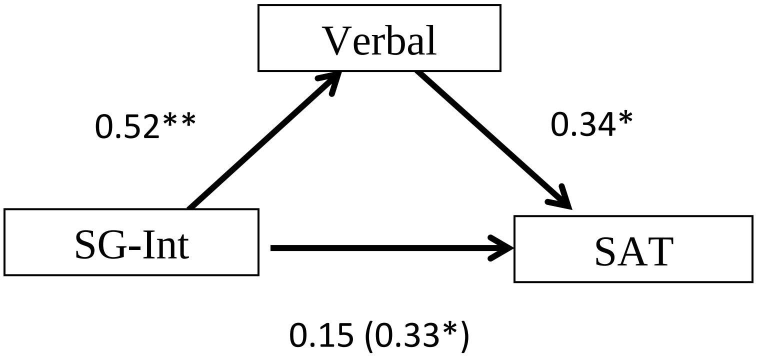


Figure 1. Frequency distribution of self-generation through integration performance among participants. The X-axis shows increments of mean percentage of successfully generated facts.

Panel A: Verbal Comprehension



Panel B: Working Memory

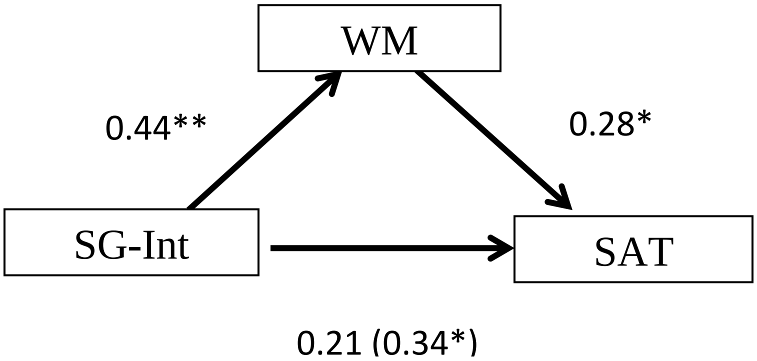
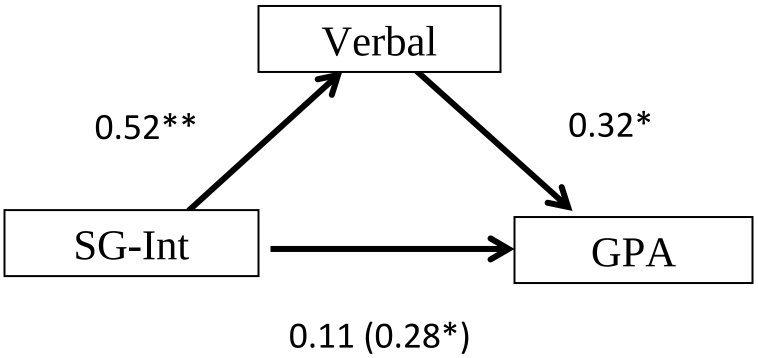


Figure 2. Standardized regression coefficients for the relation between self-generation through integration (SG-Int) and SAT score as mediated by verbal comprehension (Panel A) and working memory (Panel B). The original effect of knowledge extension on scholastic aptitude, without controlling for the mediating variable, is in parentheses. **p* < .05, ***p* < .001.

Panel A: Verbal Comprehension



Panel B: Working Memory

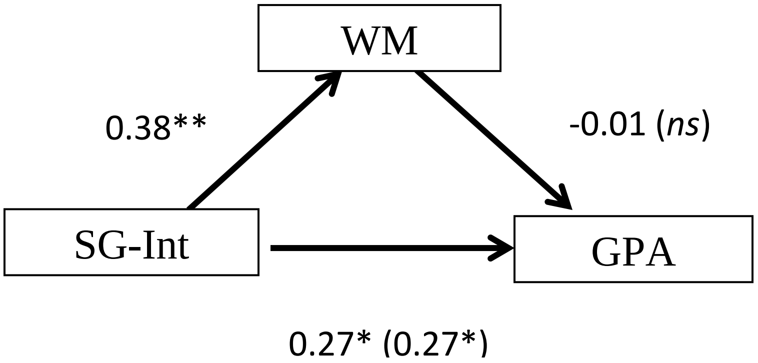


Figure 3. Standardized regression coefficients for the relation between self-generation through integration (SG-Int) and GPA (inverse-transformed scores) as mediated by verbal comprehension (Panel A) and working memory (Panel B). The original effect of knowledge extension on scholastic aptitude, without controlling for the mediating variable, is in parentheses.

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

General Discussion

In this dissertation, I examined knowledge extension through memory integration in young adults as a function of three factors: behavioral manipulations at the time of encoding and test (Study 1), underlying neural correlates (Study 2), and domain-general cognitive abilities and academic outcome (Study 3). The research was designed to address three main questions:

(1) What are the factors and conditions that impact knowledge extension through memory integration?; (2) What are the cognitive and neural processes associated with self-generation of new knowledge through integration?; and (3) How does variability in self-generation through integration contribute to real-world academic outcomes?

Factors and Conditions that Impact Knowledge Extension through Integration

Study 1 provided insight into how knowledge extension through memory integration operates under conditions that mimic those encountered in the world outside the laboratory. The findings indicated that adults self-generate and retain new factual knowledge through integration of separate yet related episodes of new learning, yet striking individual differences are apparent. Although some loss of information was observed over time, integrated knowledge remained highly accessible after a 1-week delay, thus providing empirical support for the theoretically assumed role of memory integration in the long-term accumulation of a semantic knowledge base (Bauer & Varga, 2015; Preston & Eichenbaum, 2013). Findings from Study 1 also indicated that knowledge extension through memory integration was not affected by the degree of lag between temporally distributed stem facts, at least within a single learning session. Given that temporal spacing has been shown to influence the accessibility of separate episodes to one another (e.g., Kahana & Howard, 2005) and to impact other forms of flexible knowledge extension tested within a single study session (Kornell & Bjork, 2008; Kornell, Castel, Eich, & Bjork, 2010; Vlach, Sandhofer, & Kornell, 2008), future research should address whether the capacity for memory integration in adults is sensitive to other episodic manipulations, such as recency judgments of the temporal order of studied items (e.g., Milner, 1971; Petrides, 1991) or the degree of surface similarity between episodes (Bauer et al., 2012).

In Study 1, I also explored the possibility that individual differences in successful self-generation through memory integration might be linked to whether individuals spontaneously detected the relational structure of the incidental learning task. The question was initially motivated by Barsalou and Prinz (1997) who suggested that individuals who exhibit exceptional self-generative learning abilities (i.e., “exceptional creativity”) may perceive subtle structural relations in the world that others do not. Consistent with this hypothesis, 62% of the individuals reported explicit awareness of the opportunity to integrate. Moreover, perception of the structural relations between to-be-integrated facts was strongly associated with successful self-generation. Differential response latencies during the test for knowledge extension further clarified the effect of explicit awareness on successful performance. Specifically, high-performing individuals were faster on successful trials and marginally slower on unsuccessful trials. Importantly, however, explicit awareness of the opportunity to integrate was only correlated with the amount of time spent on unsuccessful trials, not on successful trials. This pattern of results provides some clues about the time-course of this learning process. That is to say, it is possible that high-performing individuals engaged in integrative encoding during initial learning, thereby leading to faster responses and superior performance when presented with a demand to further extend that knowledge at test. Yet the finding that explicit awareness was only associated with unsuccessful trials suggests that high-performers spent significantly longer on unsuccessful trials because they were aware of the possibility to link facts in order to generate new understandings. However, if individuals failed to integrate the relation between stem facts during encoding, self-generation was still likely to be unsuccessful at the time of test. Notwithstanding, equipped with knowledge regarding the strategy to employ, high-performers persevered for longer in comparison to low-performers. In an effort to better explain the neurocognitive processing responsible for the construction of an integrated knowledge base, Study 2 was designed to delineate the point in time in which the brain first identifies the opportunity to integrate new knowledge in memory.

Neurocognitive Processes Associated with Successful Self-generation through Integration

In Study 2, I capitalized on the temporal resolution of ERPs to directly investigate the time-course of neurocognitive processing responsible for successful knowledge extension through memory integration. Specifically, ERPs were measured during encoding of the to-be-integrated stem facts and during the test for knowledge extension, affording observation of the distinct, temporally-staged processes underlying successful behavior. Because the goal was to dissociate the neural mechanisms involved in successful (as opposed to unsuccessful) knowledge extension through memory integration, the question was tested in participants in the middle of the performance distribution, who contributed comparable levels of successful and unsuccessful trials. The results indicated that the opportunity to integrate was recognized by as early as 400-600 msec after the onset of a second, related stem fact. This early effect on successful trials resembled the prototypical N400, a component that indexes whether a semantic deviation between newly and previously learned information is detected, with larger amplitudes indicative of a greater mismatch (Federmeier & Laszlo, 2009; Kutas & Federmeier, 2011; Kutas & Hillyard, 1980). The conclusion that this early effect constituted an N400 was further substantiated by the finding that ERPs to correct (but not incorrect) Stem 2 facts were positively associated with standardized measures of the ability to identify relations between concepts and to predict meaning in context according to cloze probability, the primary factor shown to modulate the N400 amplitude (Kutas & Hillyard, 1984). Thus, when presented with a second, related stem fact (*cyanide is found in pips)*, participants must identify the relation between newly and previously learned information (e.g., they must recognize they have prior knowledge about pips) and also that the newly learned information deviates from prior knowledge (e.g., pips should be associated with apple seeds, not cyanide). It is this semantic deviation between the current stimulus and prior, related knowledge which serves as a precursor to subsequent encoding processes recruited between 1100-1350 msec and 1350-1700 msec.

Whereas a deviation between separate yet related stem facts was detected by as early as 400 msec, processes involved in interpreting and integrating the relation between newly and previously learned information occurred later in the encoding window. Between 1100-1350 msec participants worked to interpret the meaning of the second, related stem fact, and to resolve the deviation previously detected. Specifically, ERP responses to facts that were unsuccessfully integrated (and Stem 1 facts) elicited more negative responses, which indexed the greater effort required to comprehend the meaning conveyed. If participants activated an intact, conflicting representation between 400-600 msec (e.g., *apple seeds are called pips*), explicit resolution of the new meaning (e.g., *cyanide is found in pips*) was facilitated. But if participants failed to activate related knowledge previously, or if the activated representation was insufficient, significantly more effort was required to select relevant content in order to understand the meaning in this time window. Between 1350-1700 msec participants integrated the new semantic understanding in memory. That is, greater posterior positivity which resembled the LPC was observed on trials in which participants failed to integrate, thereby suggesting that more processing resources were needed to represent the relation between overlapping stem facts. Together, these three temporally-staged encoding processing contributed to subsequent performance on the explicit test for self-generation through integration.

Encoding-related processes recruited during the later encoding windows (i.e., semantic interpretation and integration) were reinstated during the test for self-generation. Specifically, greater posterior positivity was observed on unsuccessful trials between 1000-1500 msec, which reflected the greater effort associated with extracting the integrated understanding. Moreover, ERPs on successful trials were associated with standardized measures of the ability to identify relations between concepts and to reason based on previously learned information. It is therefore reasonable to conclude that if participants successfully integrated related stem facts in memory at the time of encoding, self-generative processing was facilitated and subsequently successful at the time of explicit testing.

The findings from Study 2 clarify the findings and conclusions from Study 1, which raised questions about the time-course of successful knowledge extension through memory integration and the role of explicit awareness. Whereas the semantic deviation detected by 400msec might occur with little or no conscious awareness (e.g., Brualla, Romero, Serrano, & Valdizan, 1998; Rolke, Heil, Streb, & Henninghausen, 2001), pronounced differences observed between successful and unsuccessful integration during later encoding (1100-1700 msec) and at the time of test (1000-1500 msec) likely were reflective of more strategic processes associated with explicit comprehension. Moreover, although no significant correlations were observed between behavioral self-generation performance and neural responses, there was a marginally significant association between behavioral task performance and Stem 2 processing from 400-600 msec ($p = 0.09$). This pattern of results further substantiates the finding that explicit awareness was associated with response latencies on unsuccessful trials, but not on successful trials. Put concretely, when participants successfully reactivated the related yet seemingly conflicting Stem 1 fact between 400-600 msec, this initial process facilitated a cascade of subsequent processes required to integrate and further extend that knowledge. Whereas later, strategic processes are argued to occur under explicit awareness, it is the more "automatic" early processing that supported subsequent neural processing and which was marginally related to behavioral task performance. These results help to explain why conscious awareness was not associated with response latencies on successful trials—participants either successfully reactivated the knowledge necessary for integration or they did not, regardless of whether they were high or low performers. Yet if participants failed to reactivate and integrate the relevant prior knowledge during the encoding window, processing during the explicit test was significantly more effortful. Moreover, if participants explicitly recognized the opportunity to integrate, they continued to engage in strategic processing for significantly longer than participants who were unaware of the task structure. Although this might have led to success in some cases, without having integrated prior to the explicit test, performance was still likely to be unsuccessful. Indeed, participants spent an average of 13.46 sec on unsuccessful self-generation trials, as compared to approximately 4.54 sec on successful trials, further suggesting that participants persisted on trials in which strategic processing returned no results.

Taken together, Studies 1 and 2 provide complementary sources of evidence regarding the time-course of the distinct neurocognitive processes supporting self-generation performance, as well as insight into what differential response times are likely to reflect. However, it is important to re-emphasize that the findings from Study 2 were obtained from participants who comprised the middle of the performance distribution. Thus, although this study delineated a process model for successful versus unsuccessful self-generation through integration, it did not directly address the cognitive factors that contribute to person-to-person variability in this fundamental form of learning. We explored this issue in Study 3.

An Explanation of Variability in Self-generative Learning and Academic Outcome

In Study 3, we identified the cognitive factors that contribute to individual differences in the extension of knowledge through memory integration, and whether this form of learning was associated with real-world academic metrics, including SAT scores and college GPA. The findings indicated that variability in behavioral self-generation performance was jointly explained by superior verbal comprehension skills and working memory capacity. Whereas verbal comprehension likely supported the ability to comprehend the relation between newly and previously learned stem facts and to integrate that knowledge in semantic memory, working memory likely facilitated the amount of stem fact information that one could simultaneously activate and process in the service of transforming discretely stored representations in an explicit, integrated representation. A reciprocal relation between self-generation and these crystallized and fluid cognitive constructs was also observed, suggesting that one's propensity for self-generative learning also contributes to the breadth and depth of one's semantic knowledge base (i.e., crystallized cognition) as well as the ability to flexibly apply prior knowledge to meet the demands of novel learning situations (i.e., fluid cognition).

The findings from Study 3 also clarify the consequences that individual differences in self-generative learning have for educational success. Specifically, mediational analyses elucidated the association between self-generation, crystallized cognition, fluid cognition, and academic outcome. Whereas the behavioral self-generation task measured how well participants acquired, integrated, and flexibly extended *new* information, the academic measures assessed how well individuals demonstrated *prior* knowledge directly learned in high school and college (i.e., SAT and GPA) and how successfully they engaged in inferential reasoning based on previously acquired knowledge (i.e., SAT). As such, the findings indicated that self-generation was associated with a more extensive knowledge base (i.e., crystallized knowledge), which together explained 19% and 15% of variability in SAT and GPA, respectively. Additionally, self-generation was also associated with broad reasoning ability (i.e., fluid cognition), which together accounted for an additional 18% of the variance in SAT. Conversely, working memory was not associated with GPA which primarily reflects learning on a much narrower scope (i.e., recent classroom learning). In summary, Study 3 extends beyond our understanding of the basic factors and processes that contribute to knowledge extension through integration and provides a theoretically plausible and a practically significant framework from which to understand how person-to-person variability in this fundamental form of learning indirectly impacts whether individuals do well or poorly on real-world academic measures.

Limitations

The present research was not without limitations. Firstly, we did not have the opportunity to examine initial knowledge extension as a function of whether participants exhibited memory for the individual stem facts (because assessment of stem fact memory following the test for self-generation at Session 1 would have contaminated our test for retention at Session 2). This measure would permit direct examination of the role of memory for Stem 1 on subsequent identification, interpretation, and integration of the semantic deviation between newly and previously learned information. This measure also would have directly informed the association between verbal comprehension and behavioral self-generation. Secondly, it remains unclear how much knowledge participants had of the to-be-integrated stem facts prior to participation in the study. Although the use of real-world stimuli allows for direct assessment of how knowledge extension through integration contributes to semantic memory development, it is impossible to control for what every individual may know prior to participation in the study. Nevertheless, despite the noise that this ecological challenge inevitably introduces, the absence of a relation between self-reported prior knowledge and behavioral performance in most of the experiments suggests that participants typically over-estimated how much they knew about the target facts. Lastly, the finding that the lag manipulation in Study 1 had no impact of knowledge extension through integration was surprising. Thus, future research remains to be done to test if effects are observed under other episodic manipulations.

Conclusions and Future Directions

The current research furthers our understanding of the behavioral, neural, and cognitive factors associated with self-generation and retention of new factual knowledge through integration of separate yet related episodes of new learning in young adults. Most of the extant literature on knowledge extension through integration in adults has employed arbitrary paired associates. The stimuli used in the present research represented real-world factual knowledge, thus the findings provided empirical support for the assumption that memory integration supports the long-term accumulation of a semantic knowledge base. Moreover, the data suggest that individual differences in self-generative learning through memory integration in adulthood are related to both crystallized skills as well as fluid cognitive abilities. This finding is important in light of previous research with 4- and 6-year-old children which suggested that only fluid reasoning abilities are associated with the capacity to extend new knowledge through memory integration prior to formal schooling (Varga & Bauer, 2014). These findings suggest that the relative roles of fluid and crystallized cognition on the capacity to extend knowledge through integration transform substantially over the school-age and college years. That is to say, as fluid and crystallized cognition become increasingly differentiated across the lifespan (e.g., Akshoomoff et al., 2013), domain-specific crystallized skills come to explain more variance in the capacity to acquire, integrate, and further extend new knowledge. Therefore, to chart the developmental trajectory of knowledge extension through integration, further research should continue to examine how the component cognitive processes that contribute to successful behavior and the domain-general cognitive abilities that contribute to individual differences develop across the school-age years and adolescence. Examination of these questions in a broader spectrum of ages will contribute valuable insight into how to promote this process early in life, and thus has the potential to impact long-term educational outcome.

Addendum 1

Self-generation Performance by each Type of Knowledge Extension (Study 1; Exp2)

Type of Relation	Fact	Successful Trials	Total Trials	% Correct
Substitution Based on Equivalence Relations (1-Place Arguments)	Spain	13	16	81%
	Self	8	15	53%
	Stalin	10	16	63%
	Skeleton	8	15	53%
	Truckers	10	16	63%
	Intelligent	8	15	53%
	Age	6	16	38%
	Fencing	5	16	31%
	Savannas	6	15	40%
	Linguist	4	16	25%
	Pyramids	1	13	8%
	Malnutrition	6	16	38%
	Vitamin	1	16	6%
	Total	**86**	**201**	**43%**
Substitution Based on Equivalence Relations (2-Place Arguments)	Smallpox	10	15	67%
	Tibet	9	16	56%
	Cyanide	4	16	25%
	Ethiopia	7	15	47%
	Stress	5	15	33%
	Greeks	5	16	31%
	Beer	2	15	13%
	Total	**42**	**108**	**39%**
Transitive Inference Relations	Spears	10	14	71%
	Belgium	11	15	73%
	Night	8	16	50%
	Fleas	8	16	50%
	Urinal	4	16	25%
	Cigarettes	2	15	13%
	Total	**43**	**92**	**47%**
General to Specific Relations	Cavalry	5	16	31%
	Romanticism	7	15	47%
	Total	**12**	**31**	**39%**
Non-Classified Relations	England	11	15	73%
	Total	**11**	**15**	**73%**

Addendum 2

Self-generation Performance by each Type of Knowledge Extension (Study 1; Exp 3)

Type of Relation	Fact	Successful Trials	Total Trials	% Correct
Substitution Based on Equivalence Relations (1-Place Arguments)	Spain	103	120	86%
	Self	88	120	73%
	Stalin	85	119	71%
	Skeleton	78	120	65%
	Truckers	62	120	52%
	Intelligent	61	120	51%
	Age	59	120	49%
	Fencing	56	118	47%
	Savannas	55	120	46%
	Linguist	45	120	38%
	Pyramids	36	98	37%
	Malnutrition	25	120	21%
	Vitamin	23	120	19%
	Total	**776**	**1535**	**51%**
Substitution Based on Equivalence Relations (2-Place Arguments)	Smallpox	88	120	73%
	Tibet	63	119	53%
	Cyanide	58	120	48%
	Ethiopia	44	120	37%
	Stress	37	120	31%
	Greeks	33	120	28%
	Beer	26	120	22%
	Total	**349**	**839**	**42%**
Transitive Inference Relations	Spears	86	120	72%
	Belgium	80	120	67%
	Night	79	120	66%
	Fleas	64	120	53%
	Urinal	51	120	43%
	Cigarettes	23	120	19%
	Total	**383**	**720**	**53%**
General to Specific Relations	Cavalry	61	118	52%
	Romanticism	60	120	50%
	Total	**121**	**238**	**51%**
Non-Classified Relations	England	90	120	75%
	Total	**90**	**120**	**75%**