Young adults self-derive and retain new factual knowledge through memory integration

Nicole L. Varga[a] and **Patricia J. Bauer**[a]

[a]Department of Psychology, Emory University, 36 Eagle Row, Atlanta, GA, 30322, USA

Abstract

The present research investigated the retention of new factual knowledge derived through integration of information acquired across temporally distributed learning episodes. Young adults were exposed to novel facts as they read long lists of seemingly unrelated information, one sentence at a time. They then were presented open-ended questions, the answers to which could be self-derived through integration of pairs of facts from the list. Experiment 1 was the first test of self-derivation of new factual knowledge through integration in adults using open-ended testing (as opposed to forced-choice testing). Participants successfully self-derived integrated knowledge under these more challenging conditions. Experiment 2 was a test for long-term retention of newly self-derived information. Newly derived knowledge remained accessible after a 1-week delay. Striking individual differences were also observed, which were related to whether individuals spontaneously identified the relational structure of the learning task. Insight into the relation between explicit task knowledge and strategic processing was also revealed through examination of response speed at the time of test. Specifically, knowledge of the task structure was associated with response latencies on unsuccessful (but not successful) trials, such that participants who were aware of the opportunity to integrate spent longer when they were subsequently unsuccessful, presumably reflecting directed search strategies and heightened perseverance when those processes failed. Together, the present findings provide direct evidence for the role of memory integration in the long-term accumulation of a semantic knowledge base and have theoretical implications for our understanding of this fundamental form of learning.

Keywords

Semantic Memory; Episodic Memory; Memory Integration; Knowledge Extension; Associative Learning

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 be impossible. Critically, the formation of an integrated knowledge base also permits flexible extension beyond direct experience, enabling self-derivation of new thoughts, ideas, and understandings. Prior research on

Correspondence may be addressed to Nicole L. Varga, Department of Psychology, 36 Eagle Row, Emory University, Atlanta GA, 30322, USA; nvarga@emory.edu; 404-712-8330.

productive knowledge extension provides important insight into the mechanisms involved in memory integration in nonhuman animals (e.g., Bunsey & Eichenbaum, 1996; Dusek & Eichenbaum, 1997; Tse et al., 2007; 2011), in adults (e.g., Bauer & Jackson, 2015; Kumaran, Summerfield, Hassabis, & Maguire, 2009; Preston, Shrager, Dudukovic, & Gabrieli, 2004; Schlichting, Zeithamova, & Preston, 2014; Shohamy & Wagner, 2008; Sweegers, Takashima, Fernández, & Talamini, 2014; Zeithamova, Dominick, & Preston, 2012a; Zeithamova & Preston, 2010), and in children (e.g., Bauer, King, Larkina, Varga, & White, 2012; Bauer, Varga, King, Nolen, & White, 2015). Yet although productive extension is presumed to serve as a key mechanism through which a knowledge base is formed (Bauer, 2012; Bauer & Varga, 2016; Preston & Eichenbaum, 2013; Siegler, 1989), the self-derivation and later retention of factual knowledge newly derived through integration has not been examined in adults. Furthermore, few studies with adults have investigated how these processes operate under conditions that mimic those encountered in the world outside the laboratory. To address these gaps, in the present research, we examined whether memory integration underlies self-derivation of new factual knowledge (Experiment 1) and whether factual knowledge newly derived through integration becomes incorporated into the semantic knowledge base as evidenced through long-term accessibility (Experiment 2).

The primary methods used to study knowledge extension through memory integration include transitive inference and associative inference, both of which necessitate integration of overlapping yet arbitrary stimulus pairs. 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), 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 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>D). Success on the untrained pairs depends 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. For example, subjects are presented with temporally distributed pictorial 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., AC: *Chair = Blender or Butterfly?*) (though see Schlichting & Preston 2014, for an exception using cued-recall). Converging evidence from these paradigms indicates that the capacity to form 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, little is known about how this mechanism operates under conditions that mimic everyday learning situations in which the target of learning is factual knowledge versus arbitrary associations. To begin to bridge the distance, Bauer and Jackson (2015) designed an ecologically valid paradigm in which adults incidentally learned true but novel stem facts (e.g., *Animals that are nocturnal and diurnal are cathemeral; The red lemur is both nocturnal and diurnal*) and then were tested for self-derivation of new factual knowledge through integration of the target information (i.e., integration facts: *The red lemur is cathemeral*).

Relative to other inference paradigms, this method has the advantage of being about real-world factual knowledge, and thus is directly relevant to how a semantic knowledge base is built over time. Specifically, during the learning phase, individuals read separate yet related sentences that could be combined to form novel integration facts (i.e., 2-stem condition) as well as individual sentences that conveyed one-half of the information necessary to form novel integration facts (i.e., 1-stem control condition). At 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 based on the 1-stem facts, and 10 were novel. Adults selected the novel integration fact on 56% of the trials in which integration was possible (2-stem condition), which is consistent with patterns observed in other forced-choice paradigms (e.g., Schlichting et al., 2014). 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 extended it to investigate knowledge extension through integration under several other ecologically valid learning conditions. In Experiment 1 we moved beyond forced-choice measures to examine whether memory integration supports derivation of new factual knowledge when subjects must respond to open-ended questions. Current theories characterize memory integration as a critical support for a host of other flexible behaviors, including the derivation of new knowledge and creative thinking more broadly (Bauer & Varga, 2016, 2017; Schlichting & Preston, 2015). Yet to date, studies with adults have primarily examined the products of memory integration using forced-choice measures. Thus, although it is widely assumed that memory integration enables *self-derivation* of new understandings, strong tests of the assumption using real-world factual knowledge have yet to be conducted. Accordingly, in the present research we tested the frequency with which adults self-derive new factual knowledge through integration in an open-ended format. This question is of both practical and theoretical significance. In the world outside the laboratory, individuals extend their knowledge without the provision of options from which to choose—the construction of new knowledge depends upon it. From a theoretical perspective, some have proposed that forced-choice permits accurate responding based on a weaker memory trace (see Squire, Wixted, & Clark, 2007 for review). Consistent with this suggestion, superior memory in forced-choice measures (compared to open-ended measures) is well-documented for directly experienced events (e.g., Haist, Shimamura, & Squire, 1992). Thus, primary focus on forced-choice measures might over-estimate the extent to which individuals successfully engage in more demanding forms of knowledge extension. As such, Experiment 1 provides the first empirical test of the extent to which adults successfully derive new factual knowledge under more challenging, open-ended conditions.

To foreshadow the results of Experiment 1, young adults derived new factual knowledge through integration in open-ended testing. Importantly, they did not derive new factual knowledge when integration was not possible (i.e., in a 1-stem control condition), thereby validating the paradigm as a test of integration. In Experiment 2, we tested whether newly

self-derived knowledge is retained. For memory 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 a key mechanism through which a knowledge base is formed (e.g., Bauer, 2012; Bauer & Varga, 2016; Preston & Eichenbaum, 2013). Yet although the extant literature is presumed to capture integrative mechanisms involved in the long-term accumulation of knowledge, this outcome has not been examined directly, owing to the fact that the most common paradigms employed with adults rely on arbitrary stimuli that are unlikely to be incorporated into the knowledge base. Findings from Bauer and Jackson (2015) suggest that new factual knowledge derived through integration is incorporated into semantic memory within a single study session. However, we do not know whether the information is retained in memory over time. We addressed this question in Experiment 2 by testing retention after a 1-week delay. We selected 1 week because this is a period over which children have demonstrated memory for new factual knowledge derived through integration (Varga & Bauer, 2013; Varga, Stewart, & Bauer, 2016) and over which adults have demonstrated memory for directly learned educational material (Roediger & Karpicke, 2006a).

Finally, we also used Experiment 2 to begin to identify the source(s) of individual differences observed in the present research. The capacity for memory integration varies markedly in adults (Shohamy & Wagner, 2008; Schlichting et al., 2014; Zeithamova & Preston, 2010). One potential source of variance is the extent to which individuals spontaneously detect the relational nature of the learning task. Prior research focused on this question has produced mixed findings. For example, one incidental acquired equivalence study suggested that integration of faces and scenes might occur in the absence of explicit knowledge of the relations between items (i.e., only two of the 24 participants reported explicit awareness in a brief post-test questionnaire) (Shohamy & Wagner, 2008). Conversely, in another incidental learning task using arbitrary paired associates, all individuals exhibited explicit knowledge of the relations (Schlichting & Preston, 2014). Furthermore, when individuals are directly instructed to attend to the relational structure in rule-governed tasks, such as predicting the weather based on the spatial structure of fractals (e.g., Kumaran et al., 2009) or determining the patterns governing face-location pairings (e.g., Sweegers et al., 2014), there is variability in the extent to which they express and deploy their explicit knowledge of the task structure, which is positively associated with integration. To extend beyond these studies which have exclusively relied on arbitrary materials, in Experiment 2, we assessed whether explicit awareness of the opportunity to integrate was associated with the capacity to self-derive new factual knowledge through integration. Additionally, we examined response speed at the time of test to determine if explicit knowledge of the task was related to strategic processing, as indexed by differential response latencies on correct and incorrect trials as a function of awareness. Together, the present research contributes valuable insight into the later accessibility of factual knowledge newly self-derived through memory integration, in addition to potential sources of variability in this fundamental learning process.

Experiment 1

Method

Participants—Participants were 31 adults between 18–24 years ($M = 19.63$ years, $SD = 1.26$; 24 females) enrolled in undergraduate psychology courses at a private university. 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. Written informed consent was obtained prior to the start of the study. Individuals received course credit for participation. In this and the subsequent study, the protocol and procedures were approved by the university Institutional Review Board.

Stimuli—Encoding phase stimuli consisted of 75 sentences 5–10 words in length. Sixty sentences featured 30 pairs of related stem facts that could be combined to derive 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 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 presented during encoding. 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-derivation 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.

Unlike the factual stimuli employed in prior research (Bauer & Jackson, 2015), the integration facts in the present study always ended in a sentence-final word that should have been familiar to participants (e.g., *urinal versus cathemeral*). This additional constraint on the stimulus set allowed for assessment of self-derivation of new knowledge through integration in open-ended testing (as opposed to forced-choice selection).

Procedure—The procedure had two phases: encoding and test.

<u>Encoding:</u> At the start of the session, participants were told we were interested in whether memory for newly learned factual information differs as a function of subject domain. 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. To continue with the previous example, in the 1-stem condition, participants saw either "A popular sculpture made from a urinal is called Fountain," or "Duchamp's most well-known work is named Fountain," but not both (see Supplemental Materials for additional examples and information regarding validation of the stimulus set). As depicted in Figures 1A and 1B, 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 knowledge-status 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 2–4 intervening sentences. Lag created temporal distance between to-be-integrated information and prevented participants from anticipating the content of the next fact. Across the sample, each fact was tested in each lag an equal number of times. Additionally, fact assignment was counterbalanced such that pairs of stem facts were tested in the 1-stem and 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: After a break of 5–10 minutes (filled with demographic questionnaires), participants were presented with 30 facts derived through integration of the previously presented stem facts (i.e., 15 2-stem pairs and 15 1-stem facts) using PowerPoint® software. As depicted in Figure 1C, the sentences were presented in the form of questions by omitting the final word of each fact. Participants were asked to provide a one-word answer that could accurately fill in the blank. Participants were given an unlimited amount of time. 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. The experimenter wrote down and scored the answer online (see Supplemental Materials for details regarding scoring criteria), and then presented the next question. Following the open-ended questions, participants received forced-choice questions for any integration facts that were not successfully self-derived. 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., *Duchamp's most popular work consisted of a urinal*) 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, thereby providing a subjective measure of participant's pre-existing knowledge of the integration facts on which they were tested.

Scoring and Analysis—Participants received a score of 1 or 0 for each target integration fact (correctly or incorrectly self-derived in open-ended testing or selected in forced-choice testing), and for each stem/integration prior knowledge judgment (previously known or unknown). Because the number of integration facts included in analyses differed across participants, a proportion score was calculated (see Supplemental Materials for description

of why 1.29% of the total 1-stem and 2-stem trials were omitted and why one stimulus pair was excluded from all analyses). To reduce noise and increase reliability, stem and integration prior knowledge judgments corresponding to omitted integration trials were also excluded, and a prior knowledge proportion score was calculated (the total remaining trials for each prior knowledge measure is reported below).

Results

Self-derivation across Fact Conditions—The main purpose of the present experiment was to test whether participants self-derived the integration facts in open-ended testing when they 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 2 (Panel A), participants self-derived significantly more integration facts in the 2-stem than the 1-stem condition, $t(30) = 7.55$, $p < 0.001$, $d = 1.62$ (see Supplemental Materials for details regarding how 1-stem and 2-stem performance was analyzed within each individual fact pair). A parallel pattern of results was observed in forced-choice testing, $t(28) = 8.15$, $p < 0.001$, $d = 1.87$. Although performance was reliably higher in the 2-stem than 1-stem condition, in both conditions, performance was greater than chance (25%), $t(28) = 14.10$, $p < 0.001$ and $t(30) = 6.45$, $p < 0.001$ in the 2-stem and 1-stem conditions, respectively.

Prior Knowledge of the Stem and Integration Facts—As shown in Figure 2 (Panel B), there was substantial variability in open-ended performance in the 2-stem condition, ranging from 7% to 100% correct. One potential source of variability is prior knowledge of the stimulus facts. To assess prior knowledge, we first examined participants' self-reported prior knowledge of the individual *stem* facts. When ratings were collapsed across facts presented in the 1-stem and 2-stem conditions, 17% (220 out of 1328 trials) of the total facts were reported as previously known ($M = 0.16$; $SD = 0.10$). Ratings of prior knowledge did not differ as a function of whether the fact appeared in the 1-stem ($M = 0.17$; $SD = 0.14$; 76 out of 438 trials) or the 2-stem condition ($M = 0.16$; $SD = 0.11$; 144 out of 890 trials), $t(30) = 0.51$, $p = 0.61$, $d = 0.10$. Importantly, for facts presented in the 1-stem condition, there was not a significant relation between reported prior knowledge and open-ended self-derivation performance, $t(29) = .10$, $p = 0.59$. Thus, participants' reported prior knowledge of single stem facts did not predict production of the novel integration facts. Conversely, the correlation between reported prior knowledge and subsequent open-ended self-derivation when both members of the stem fact pair were presented (i.e., the 2-stem condition) approached significance $t(29) = .35$, $p = 0.053$. This relation implies that prior knowledge of one or the other of the stem facts facilitated self-derivation performance. This same pattern of results was obtained when we examined the within-participant relation between prior knowledge and self-derivation (see Supplemental Materials).

We also examined participants' self-reported prior knowledge of the *integration* facts. When ratings were collapsed across integration facts that corresponded to stem facts presented in the 1-stem and 2-stem conditions, 11% (99 out of 886 trials) of the facts were reported as previously known ($M = 0.11$; $SD = 0.09$). Ratings of prior knowledge did not differ as a function of whether the fact corresponded to stem facts that had appeared in the 1-stem ($M = 0.10$; $SD = 0.09$; 43 out of 439 trials) or the 2-stem condition ($M = 0.13$; $SD = 0.12$; 56 out

of 447 trials), $t(30) = 1.24$, $p = 0.23$, $d = 0.26$. There was not a significant correlation between self-derivation and self-reported knowledge of the integration facts, $r(29) = .24$, $p = 0.20$, indicating that participants' ratings were largely unrelated to their actual production of the integration fact.

Discussion

The present experiment is the first to demonstrate that adults self-derive new factual knowledge through integration of separate yet related information in open-ended testing. 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 small number of trials, integration of the related stem facts was necessary for successful knowledge extension. Nevertheless, unlike the pattern reported in Bauer and Jackson (2015), in forced-choice testing participants selected the correct answer at above-chance levels in both the 1-stem and 2-stem conditions. We attribute above-chance performance in the 1-stem condition to the fact that the sentences ended in a word that was familiar to participants. We designed the stimuli this way to allow for assessment of open-ended performance. The feature also may have increased the likelihood of responding in forced-choice testing based on familiarity of the final word of each fact. Importantly, we presented the forced-choice analyses to provide a fuller picture of participants' performance —they are not the primary foundation on which conclusions are based.

With respect to self-reported prior knowledge, although participants reported knowing 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 open-ended self-derivation performance. There was, however, a relation between prior knowledge of the individual 2-stem facts and self-derivation performance. This is not surprising given that, if participants knew one stem fact, it could reasonably be expected to facilitate processing of the related stem fact and thus self-derivation of integrated knowledge. To corroborate this interpretation, the same pattern of results was observed when we examined the relation between prior knowledge and successful open-ended self-derivation in the 2-stem condition *within* participants (see Supplemental Martials). This raises the interesting possibility and need for future research examining self-derivation 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-derivation. Thus, prior knowledge of only one of the stem facts did not support production of the novel integration facts.

The present experiment also revealed substantial individual differences in self-derivation of new factual knowledge through integration. Given the sensitivity of the paradigm to individual differences, in Experiment 2, we increased the sample size and extended the paradigm to examine measures that might inform our understanding of this variability, including explicit awareness of the task structure and response latencies at test. We also

conducted the first test of whether newly self-derived knowledge persists in semantic memory over time in young adults.

Experiment 2

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 Experiment 1. None of the participants had taken part in Experiment 1. 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 within the specified delay interval (N = 1). Written informed consent was obtained prior to the start of the study. Participants received course credit upon completion of the second visit.

Stimuli—The stimuli were the same as in Experiment 1. Based on significantly poorer performance in the 1-stem control condition reported in the previous experiment, all stem facts were presented in a 2-stem condition here.

Procedure—Participants completed two sessions separated by 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. 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.

<u>Session 1:</u> The procedure was the same as in Experiment 1 with some exceptions. First, in order to collect reaction times, stimuli were presented using E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). As depicted in Figure 1 Panel C, reaction time measures were time-locked to the question mark that first cued participants to derive an answer. Second, although open-ended test phase procedures were identical to those in Experiment 1, 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 session, thereby ensuring an uncontaminated measure of long-term retention. Lastly, following the open-ended questions at Session 1, participants completed a survey inquiring about their explicit perceptions of the task. The measure was designed to assess whether individuals recognized the opportunity to integrate separate yet related facts (see Appendix A). The survey was added to the protocol midway through data collection, after several participants spontaneously commented on the relational structure of the task. Consequently, self-reported awareness was only assessed for the final 78 participants.

To reiterate, though participants were told that their memory would be tested, at no time were they informed that the sentences were related. Moreover, the same counterbalancing

scheme from Experiment 1 was employed. Thus, each fact was presented in a lag of 2, 3, and 4 approximately equally often (see Supplemental Materials for analysis of lag effects) and each fact from a target pair appeared equally often in the first or second serial position.

Session 2: Participants returned to the laboratory approximately one week later. After completion of several standardized tasks, including the Visual-Auditory Learning subtest of the *WJ-III COG* (Woodcock, McGrew, & Mather, 2001) which served as control measure of memory for directly learned paired associates (see Supplemental Materials), memory for the integration facts was assessed. First, participants were tested for recall of the integration facts using the same open-ended questions as Session 1; the order of the questions was different to reduce carry-over effects. The facts were presented on a laptop using PowerPoint® software; reaction time was not recorded. 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 1.

Scoring and Analysis—As in Experiment 1, a proportion score was calculated for each target integration fact and prior knowledge judgment to account for the different number of trials across participants (0.83% of the trials were excluded, see Supplemental Materials). Additionally, participants received a score of 1 or 0 for explicit awareness (did or did not indicate knowledge of the opportunity to integrate on the survey, respectively; Appendix A). In cases in which data were missing for some participants (e.g., awareness ratings), degrees of freedom were adjusted accordingly.

Results

Self-derivation and Retention—At Session 1, participants self-derived the novel integration facts on 50% of the trials ($M = 0.50$; $SD = 0.21$). As depicted in Figure 3 (Panel A), substantial individual differences were observed with performance ranging from 3% to 93% correct across the sample.

The primary question of interest was whether young adults successfully retained the newly self-derived knowledge over the 1-week delay. At Session 2, when tested in an open-ended format, participants recalled 42% of the total integration facts ($M = 0.42$; $SD = 0.21$). Despite high levels of recall and similar patterns of variability to that observed at Session 1 (Figure 3, Panel B), a significant loss of information was observed between sessions, $t(114) = 10.04$, $p < .001$, $d = 0.53$. Yet as depicted in Figure 4, when initial and delayed performance was compared at the individual level, the modal number of integration facts lost between sessions was only 1 or 2, with many individuals exhibiting no loss at all. Consistent with this observation, performance at Session 2 was significantly correlated with self-derivation performance at Session 1, $r(113) = .92$, $p < .001$. Finally, for the facts that were not successfully recalled after the delay, participants selected the correct answer on 51% of the forced-choice trials ($M = 0.51$; $SD = 0.15$), which significantly differed from chance (25%), $t(114) = 18.98$, $p < .001$. In total, participants either recalled the integration facts or selected them in forced choice on 71% of the trials ($M = 0.71$, $SD = 0.02$).

Individual Differences in Self-derivation—To investigate variability in self-derivation performance, we examined the latency to derivation of the novel integration facts. As depicted in Table 1, a dependent samples *t*-test indicated that participants were significantly faster to respond on successful versus unsuccessful trials, $t(106) = 13.06$, $p < .001$, $d = 1.50$. Moreover, self-derivation performance and reaction time were significantly negatively correlated, $t(105) = -.24$, $p = .01$. That is, high performers derived 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-derivation performance was found, $t(105) = -.23$, $p = .02$. However, examination of unsuccessful trials revealed a marginally significant positive correlation between reaction time and self-derivation performance, $t(105) = .19$, $p = .052$. Thus, high-performing participants not only responded faster on successful trials, but also were marginally slower on unsuccessful trials. Importantly, this pattern of results was also obtained when the effect of memory for directly learned items was controlled, indicating that individual differences in self-derivation and relations to response time are not accounted for by memory alone (see Supplemental Materials).

We next addressed whether awareness of the opportunity to integrate was related to variability in self-derivation performance. Of the 78 participants who completed the survey, 62% ($N = 48$) reported explicit awareness that some of the facts were related. A Spearman's rho indicated that explicit awareness was significantly correlated with the proportion of integration facts self-derived 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 derive 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, $r_s(71) = .02$, $p = .87$. Interestingly, however, a significant positive correlation was found between explicit awareness and mean reaction time on unsuccessful trials, $r_s(71) = .34$, $p = .004$, such that participants who were aware of the opportunity to integrate spent longer on trials in which they were unsuccessful. These relations remained when the effect of memory for directly learned paired associates was controlled, as well as when these measures were treated as dichotomous variables (see Supplemental Materials for partial correlations and chi square analyses).

Prior Knowledge of the Stem Facts—As in Experiment 1, we also examined participants' self-reported prior knowledge of the individual stem facts. On average, 21% of the facts were identified as known ($M = 0.21$; $SD = 0.13$). The proportion of facts previously known was significantly correlated with self-derivation of the novel integration facts, $t(114) = .18$, $p = .049$. The same pattern of results was observed when we examined the relation between prior knowledge of the stem facts and subsequent self-derivation within participants (see Supplemental Materials).

Discussion

The present experiment replicated and extended Experiment 1. That is, adults extended new knowledge under open-ended testing conditions and they exhibited striking variability in self-derivation of new knowledge through memory integration. Although newly self-derived

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 (for a total retention score of 71%). Therefore, information self-derived through integration persisted in memory over time.

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 failed, namely, on unsuccessful trials. Support for this conclusion comes from the finding that explicit awareness of the opportunity to integrate was 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. In contrast, on successful trials, stronger encoding during the learning phase might have led to quicker access of knowledge needed for self-derivation.

General Discussion

The present research was an investigation of self-derivation and retention of new factual knowledge through integration of separate yet related episodes of new learning. We extended prior research by examining this learning process under conditions that mimic those encountered in everyday learning situations. The primary question concerned the extent to which adults were successful at extending knowledge through more challenging means than have typically been examined, namely, in open-ended as opposed to forced-choice testing. Adults successfully self-derived integrated knowledge under these conditions. We next assessed whether newly self-derived knowledge was retained over time. Although some loss was observed, 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 factual knowledge. Finally, striking individual differences were evident in the extent to which adults successfully self-derived 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 understanding of the nature of memory integration, particularly with respect to its implications for self-derivative behavior. Memory integration enables individuals to establish links between separate yet related traces of information, forming the building blocks of a semantic knowledge base. Importantly, formation of an integrated semantic knowledge base then enables the striking capacity to derive new knowledge; acts ranging from basic creativity to the derivation of scientific theories depend upon this generative capacity. The present research constitutes the first test of the extent to which adults self-derive understandings typically demanded in everyday learning situations, namely, generation of *factual knowledge* as opposed to arbitrary associations. In both

experiments, participants self-derived the novel integration facts on approximately 50% of the trials. Moreover, substantial individual differences were observed, with successful self-derivation ranging from 3% to 100% across experiments. This range of performance is similar to that observed for integration of arbitrary associations using cued-recall (Schlichting & Preston, 2014; *Range* = 6.7 to 83.3%). Studies employing open-ended, cued-recall procedures are important because, as discussed previously, they require stronger memory traces than those that may support forced-choice responding (Squire et al., 2007). The present research makes clear that individuals extend new factual knowledge through integration under more challenging testing conditions than are typically employed. What is more, the extent of variability observed is comparable across studies that employ arbitrary or naturalistic materials.

The present experiments also take an important step toward furthering our understanding of the long-term retention of self-derived knowledge. In everyday learning contexts, delays between initial learning and later use are commonly encountered. Although many researchers have acknowledged the 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 (though see Roediger & Karpicke, 2006a for extensive review of evidence for long-term retention of directly learned materials). Yet 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 & Bauer, 2013; Varga et al., 2016), Experiment 2 of the present research indicated that young adults retain knowledge newly derived through memory integration. That is, at Session 1 individuals self-derived 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 (indicating total retention of 71% of the integration facts). Relatively high accessibility following a delay suggests that knowledge newly self-derived through integration had been incorporated into the knowledge base. This conclusion is consistent with results from Bauer and Jackson (2015). 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, there was significant loss of information between sessions in Experiment 2. This finding is consistent with Sweegers and colleagues (2014), the only analogous investigation of long-term retention of integrated knowledge conducted with adults to date. They found that directly learned, cross-episode relations were preferentially consolidated in memory as compared to isolated memories. Yet even these integrated

representations exhibited significant degradation over 48 hours. Indeed, diminished retention in the face of a 1-week delay in the present research is not altogether surprising. For instance, in the testing-effect literature, retrieving previously learned material confers benefits for the longevity of memory across free- and cued-recall paradigms (e.g., Tulving, 1967; Lachman & Laughery, 1968; Allen, Mahler, & Estes, 1969; Jacoby, 1978) and for educationally-relevant materials (e.g., Karpicke & Roediger, 2006b; McDaniel & Fisher, 1991). Nevertheless, forgetting constitutes the rule rather than the exception. For instance, even when participants' memory of educationally-relevant prose passages was tested three times after initial learning, individuals still forgot 14% of the material after a 1-week delay (Karpicke & Roediger, 2006b). Thus, we might view it as remarkable that undergraduate students retained as much knowledge as they did, especially given that they acquired the novel integration facts through single-trial procedures. Notwithstanding, because significant loss was still apparent, additional research aimed at promoting the long-term accessibility of new factual knowledge derived through memory integration is warranted.

The current research also advances our understanding of potential sources of variability in extension of new knowledge through integration. Specifically, 62% of the individuals in Experiment 2 reported explicit awareness of the opportunity to integrate; whereas, 38% made no mention of the possibility to do so. What is more, perception of the structural relations between to-be-integrated facts was associated with self-derivation. This finding is particularly interesting in light of conflicting findings in the literature. As discussed above, when cross-episode integration was elicited through trial-and-error, reinforcement training of face-scene equivalencies, Shohamy and Wagner (2008) found that explicit awareness of the task structure was exceedingly rare (also see Daw & Shohamy, 2008; Greene, Gross, Elsinger, & Rao, 2006 for similar results in acquired equivalence paradigms). Moreover, in an examination of explicit awareness using the paired associate learning paradigm, Schlichting and Preston (2014) found that all participants were aware of the relational structure of the learning task, thereby precluding examination of whether it was associated with inferential performance (though, importantly, participants were trained to criterion on the individual pairs which likely heightened awareness).

Based on the present research alone, we cannot determine the source (or sources) of differential relations between explicit awareness and memory integration in the present research versus previous research. We speculate that the nature of the stimuli contributed to the differential findings. 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 consider. Implicit in this proposal is the assumption that flexible behavior depends on the capacity to detect subtle relations across various forms of knowledge, much like in the present research. That is to say, to recognize that they could link separate yet related facts, individuals were required to detect the relation between various elements of knowledge spanning history, art, biology, and more. Moreover, explicit recognition of the opportunity to integrate could also facilitate performance by allowing those same individuals to process related stem facts at a deeper level, thereby enabling extraction of the relation and facilitating performance and response speed on the subsequent test. In contrast, even when individuals become aware of the

prototypical AB:BC paired associate structure, the arbitrary nature of the relations likely precludes deeper processing that could further enhance integration performance. Consistent with this interpretation, explicit awareness is also associated with integration in pattern-learning tasks in which individuals must extract subtle rules, such as how the spatial and non-spatial orientation of various fractals determine weather outcomes (Kumaran et al., 2009) or rules regarding face attributes repeatedly encountered at specific locations (Sweegers et al., 2014). Thus, it is possible that the nature of the stimuli might account for greater variability in explicit awareness of the relational structure relative to research utilizing more contrived, arbitrary items. 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. 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 new knowledge.

The current research also provides some clues about 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 relations are only established when discrete memories are retrieved and recombined in response to a demand 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 later, even though participants made inferential judgments (e.g., Shohamy & Wagner, 2008). However, retrieval-based signatures have 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, 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, such as in the present research, integrative encoding is necessary but not sufficient for explicit knowledge extension. Consistent with this notion, in a child analogue of the paradigm employed here, Varga and Bauer (2013) inserted delays throughout the learning process and found that self-derivation through integration consisted of a two-step process, one of integrative encoding followed by flexible manipulation of that information at test (see Bauer, Blue, Xu, & Esposito, 2016 for similar encoding results using eye-tracking). It is therefore reasonable to suggest that related stem facts may be integrated in advance of a test probe which then further prompts explicit self-derivation.

The present research cannot directly address the mechanisms involved during encoding and test. Yet, 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-derived the novel integration fact. Response latency advantages for high performers (relative to low performers) have similarly been observed in other memory integration studies. For instance, Shohamy and Wagner (2008) found that the lower half of performers showed significantly more slowing on trials assessing integration versus trials testing memory for trained items, relative to participants in the upper half of the distribution. Consistent with the interpretation proffered in the previous study, one possible explanation for this pattern is that high-performing individuals in the current study 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 test. This could explain both slower response times and less success overall. Second, high performers were nominally slower on trials in which they failed to self-derive, 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 derive new understandings. However, if they failed to integrate during initial encoding, self-derivation 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. It is important to emphasize that although several investigations have converged on the finding that integration at the time of learning confers benefits for response time during a subsequent test for knowledge extension (e.g., Kumaran et al., 2009; Sweegers et al., 2014), analyses are typically limited to correct trials. Thus, the present research contributes novel empirical support for the idea that strategic processing might play an important role in the capacity for integration.

In conclusion, the present experiments took important steps toward furthering our understanding of the self-derivation and long-term retention of factual knowledge newly derived through integration, as well as some of the sources of individual variability in this fundamental learning process. In addition to contributing to our understanding of self-derivative learning 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, 2017 for discussion). The results reported here demonstrate that even among young adults, there is substantial variability in performance that depends on integrating and further generating new knowledge, consistent with what has previously been reported in forced-choice paradigms. 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 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 newly self-derived knowledge was less accessible after a 1-week delay, future research must determine means of further promoting the long-term accessibility of self-derived knowledge.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

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?

References

Allen GA, Mahler WA, Estes WK. Effects of recall tests on long-term retention of paired associates. Journal of Verbal Learning and Verbal Behavior. 1969; 8:463–470.

Barsalou, LW., Prinz, JJ. Mundane creativity in perceptual symbol systems. In: Ward, TB.Smith, SM., Vaid, J., editors. Creative thought: An investigation of conceptual structures and processes. Washington, DC: American Psychological Association; 1997. p. 267-307.

Bauer, PJ. The life I once remembered: The waxing and waning of early memories. In: Bensten, D., Rubin, DC., editors. Understanding autobiographical memory: theories and approaches. Cambridge, UK: Cambridge University Press; 2012. p. 205-225.

Bauer PJ, Blue SN, Xu A, Esposito AG. Productive extension of semantic memory in school-aged children: Relations with reading comprehension and deployment of cognitive resources. Developmental Psychology. 2016; 52(7):1024–1037. DOI: 10.1037/dev0000130 [PubMed: 27253263]

Bauer PJ, Jackson FL. Semantic elaboration: ERPs reveal rapid transition from novel to known. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2015; 41(1):271–282. DOI: 10.1037/a0037405

Bauer PJ, King JE, Larkina M, Varga NL, White EA. Characters and clues: Factors affecting children's extension of knowledge through integration of separate episodes. Journal of Experimental Child Psychology. 2012; 111(4):681–694. DOI: 10.1016/j.jecp.2011.10.005 [PubMed: 22153911]

Bauer, PJ., Varga, NL. The developmental cognitive neuroscience of memory: Implications for education. In: Tardif, É., Doudin, P., editors. Collective work on the topics of neuroscience, cognition and education. Oxford, UK: De Boeck; 2016.

Bauer PJ, Varga NL. Similarity and deviation in event segmentation and memory integration. [Peer commentary on the paper Event perception: Translations and applications by L.L. Richmond, D. A. Gold, & J. M. Zacks]. Journal of Applied Research in Memory and Cognition. 2017

Bauer PJ, Varga NL, King JE, Nolen AM, White EA. Semantic elaboration through integration: Hints both facilitate and inform the process. Journal of Cognition and Development. 2015; 16(2):351–369. DOI: 10.1080/15248372.2013.849707

Bunsey M, Eichenbaum H. Conservation of hippocampal memory function in rats and humans. Nature. 1996; 379(6562):255–257. [PubMed: 8538790]

Daw ND, Shohamy D. The cognitive neuroscience of motivation and learning. Social Cognition. 2008; 26(5):593–620. DOI: 10.1521/soco.2008.26.5.593

Dusek JA, Eichenbaum H. The hippocampus and memory for orderly stimulus relations. Proceedings of the National Academy of Sciences. 1997; 94(13):7109–7114.

Gentner, D., Smith, L. Analogical reasoning. In: Ramachandran, VS., editor. Encyclopedia of Human Behavior. 2nd. Oxford, UK: Elsevier; 2012. p. 130-136.

Greene AJ, Gross WL, Elsinger CL, Rao SM. An fMRI analysis of the human hippocampus: Inference, context, and task awareness. Journal of Cognitive Neuroscience. 2006; 18:1156–1173. [PubMed: 16839289]

Haist F, Shimamura AP, Squire LR. On the relationship between recall and recognition memory. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1992; 18(4):691–702. DOI: 10.1037/0278-7393.18.4.691

Heckers S, Zalesak M, Weiss AP, Ditman T, Titone D. Hippocampal activation during transitive inference in humans. Hippocampus. 2004; 14(2):153–162. [PubMed: 15098721]

Jacoby LL. On interpreting the effects of repetition: Solving a problem versus remembering a solution. Journal of Verbal Learning and Verbal Behavior. 1978; 17(6):649–667. DOI: 10.1016/S0022-5371(78)90393-6

Jee BD, Uttal DH, Gentner D, Manduca C, Shipley TF, Tikoff B, Sageman B. Commentary: Analogical thinking in geoscience education. Journal of Geoscience Education. 2010; 58(1):2–13. DOI: 10.5408/1.3544291

Kumaran D. What representations and computations underpin the contribution of the hippocampus to generalization and inference. Frontiers in Human Neuroscience. 2012; 6:1–11. [PubMed: 22279433]

Kumaran D, McClelland JL. Generalization through the recurrent interaction of episodic memories: A model of the hippocampal system. Psychological review. 2012; 119(3):573–616. [PubMed: 22775499]

Kumaran D, Summerfield JJ, Hassabis D, Maguire EA. Tracking the emergence of conceptual knowledge during human decision making. Neuron. 2009; 63:889–901. DOI: 10.1016/j.neuron.2009.07.030 [PubMed: 19778516]

Lachman R, Laughery KR. Is a test trial a training trial in free recall learning? Journal of Experimental Psychology. 1968; 76:40–50. DOI: 10.1037/h0025278

McDaniel MA, Fisher RP. Tests and test feedback as learning sources. Contemporary Educational Psychology. 1991; 16:192–201.

Preston AR, Eichenbaum H. Interplay of hippocampus and prefrontal cortex in memory. Current Biology. 2013; 23(17):R764–R773. [PubMed: 24028960]

Preston AR, Shrager Y, Dudukovic NM, Gabrieli JD. Hippocampal contribution to the novel use of relational information in declarative memory. Hippocampus. 2004; 14(2):148–152. DOI: 10.1002/hipo.20009 [PubMed: 15098720]

Roediger HL, Karpicke JD. The power of testing memory: Basic research and implications for educational practice. Perspectives on Psychological Science. 2006a; 1(3):181–210. [PubMed: 26151629]

Roediger HL, Karpicke JD. Test-enhanced learning: Taking memory tests improves long-term retention. Psychological Science. 2006b; 17(3):249–255. DOI: 10.1111/j.1467-9280.2006.01693.x [PubMed: 16507066]

Schlichting ML, Preston AR. Memory reactivation during rest supports upcoming learning of related content. Proceedings of the National Academy of Sciences of the United States of America. 2014; 111(44):15845–15850. DOI: 10.1073/pnas.1404396111 [PubMed: 25331890]

Schlichting ML, Preston AR. Memory integration: Neural mechanisms and implications for behavior. Current opinion in behavioral sciences. 2015; 1:1–8. [PubMed: 25750931]

Schlichting ML, Zeithamova D, Preston AR. CA1 subfield contributions to memory integration and inference. Hippocampus. 2014; 24(10):1248–1260. [PubMed: 24888442]

Shohamy D, Wagner AD. Integrating memories in the human brain: Hippocampal-midbrain encoding of overlapping events. Neuron. 2008; 60(2):378–389. DOI: 10.1016/j.neuron.2008.09.023 [PubMed: 18957228]

Siegler RS. Mechanisms of cognitive development. Annual Review of Psychology. 1989; 40(1):353–379.

Squire LR, Wixted JT, Clark RE. Recognition memory and the medial temporal lobe: A new perspective. Nature Reviews Neuroscience. 2007; 8(11):872–883. DOI: 10.1038/nrn2154 [PubMed: 17948032]

Sweegers CCG, Takashima A, Fernández G, Talamini L. Neural mechanisms supporting the extraction of general knowledge across episodic memory. NeuroImage. 2014; 87:138–146. [PubMed: 24215973]

Tse D, Langston RF, Kakeyama M, Bethus I, Spooner PA, Wood ER, Morris RG. Schemas and memory consolidation. Science. 2007; 316(5821):76–82. [PubMed: 17412951]

Tse D, Takeuchi T, Kakeyama M, Kajii Y, Okuno H, Tohyama C, Morris RG. Schema-dependent gene activation and memory encoding in neocortex. Science. 2011; 333(6044):891–895. [PubMed: 21737703]

Tulving E. The effects of presentation and recall of material in free recall learning. Journal of Verbal Learning and Verbal Behavior. 1967; 6:175–184.

Varga NL, Bauer PJ. Effects of delays on 6-year-old children's self-generation and retention of knowledge through integration. Journal of Experimental Child Psychology. 2013; 115(2):326–341. [PubMed: 23563162]

Varga NL, Stewart RA, Bauer PJ. Integrating across episodes: Investigating the long-term accessibility of self-derived knowledge in 4-year-old children. Journal of Experimental Child Psychology. 2016; 145:48–63. DOI: 10.1016/j.jecp.2015.11.015 [PubMed: 26774259]

Woodcock, RW., McGrew, KS., Mather, N. Woodcock-Johnson III Tests of Cognitive Abilities. Itasca, IL: Riverside Publishing; 2001.

Zeithamova D, Dominick AL, Preston AR. Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. Neuron. 2012a; 75(1):168–179. [PubMed: 22794270]

Zeithamova D, Preston AR. Flexible memories: Differential roles for medial temporal lobe and prefrontal cortex in cross-episode binding. The Journal of Neuroscience. 2010; 30(44):14676–14684. [PubMed: 21048124]

Zeithamova D, Schlichting ML, Preston AR. The hippocampus and inferential reasoning: Building memories to navigate future decisions. Frontiers in Human Neuroscience. 2012b; 6:1–14. [PubMed: 22279433]

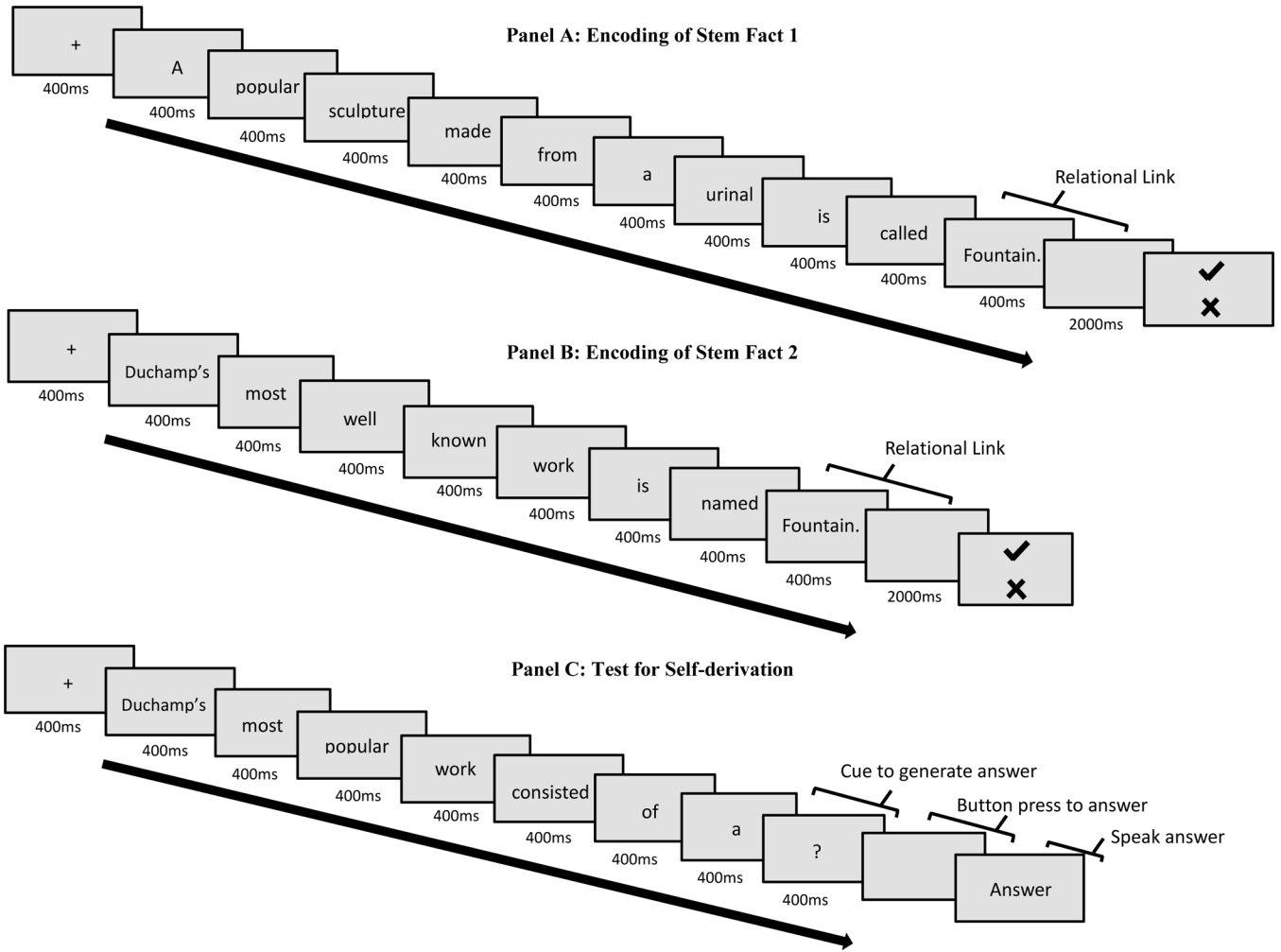**Panel A: Encoding of Stem Fact 1**

+ 400ms
A 400ms
popular 400ms
sculpture 400ms
made 400ms
from 400ms
a 400ms
urinal 400ms
is 400ms
called 400ms
Fountain. 400ms

Relational Link

2000ms

Panel B: Encoding of Stem Fact 2

+ 400ms
Duchamp's 400ms
most 400ms
well 400ms
known 400ms
work 400ms
is 400ms
named 400ms
Fountain. 400ms

Relational Link

2000ms

Panel C: Test for Self-derivation

+ 400ms
Duchamp's 400ms
most 400ms
popular 400ms
work 400ms
consisted 400ms
of 400ms
a 400ms
? 400ms

Cue to generate answer

Button press to answer

Speak answer

Answer

Figure 1.

Schematic of encoding (Panels A and B) and test phase (Panel C) procedures in Experiments 1 and 2. Reaction time was recorded from the onset of the "?" and terminated when a button press response was made during the test phase in Experiment 2.

Panel A: Performance across Conditions
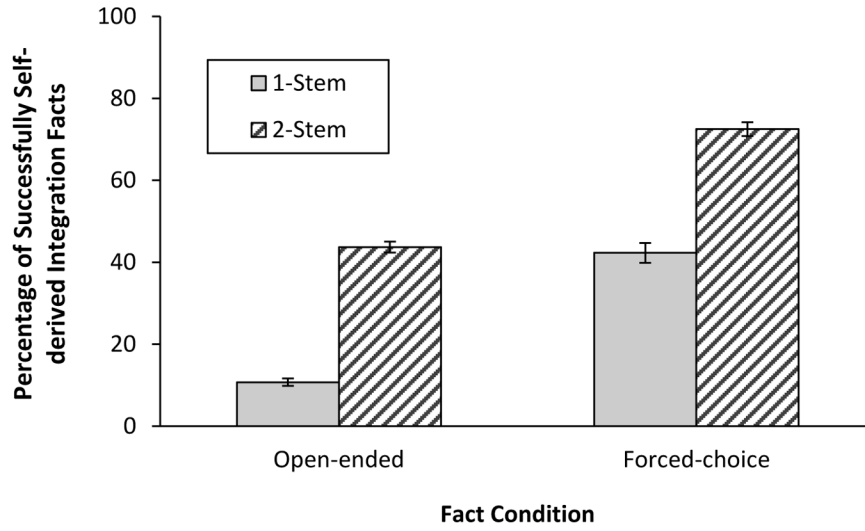
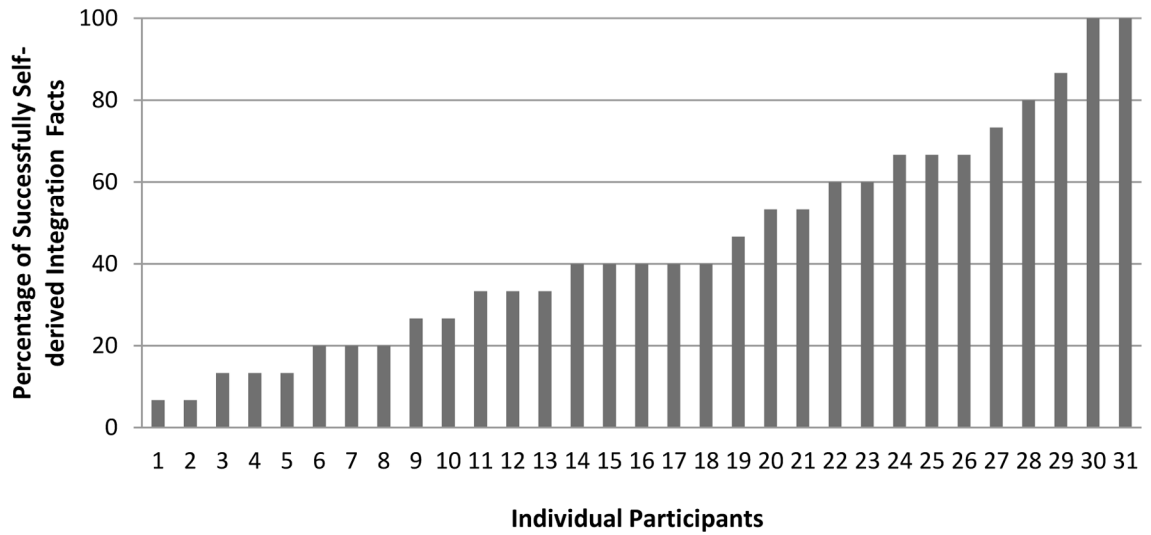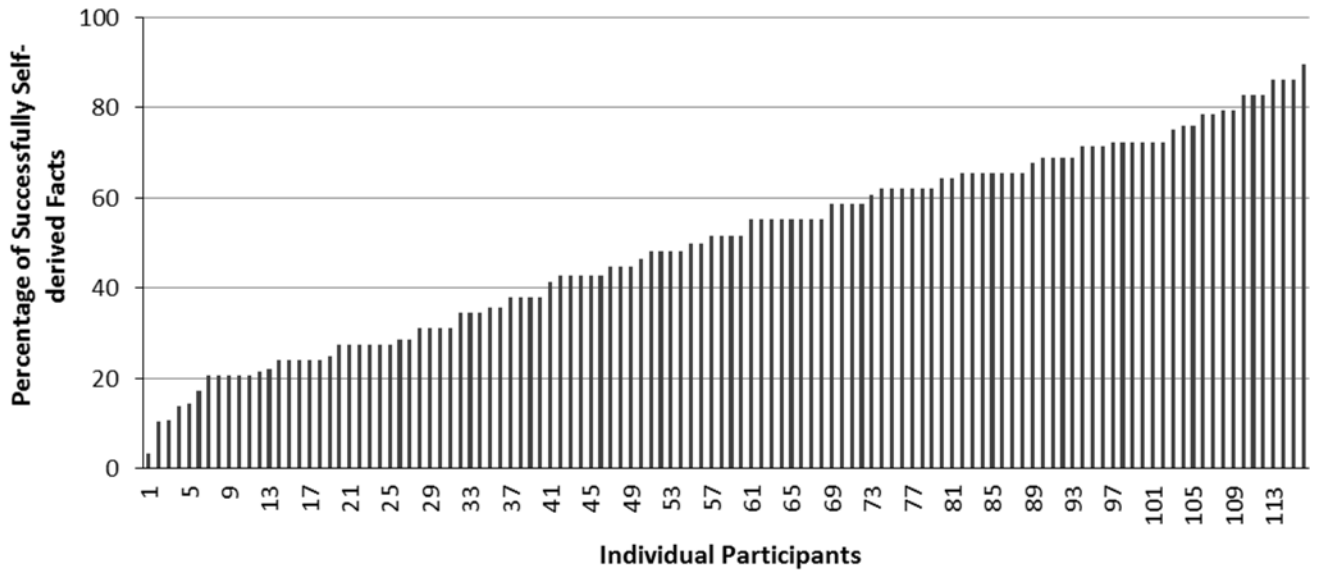**Panel B: Variability across Participants**

Figure 2.
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-derived by individual participants in the 2-stem condition (Panel B) in Experiment 1. Error bars represent standard error of the mean.

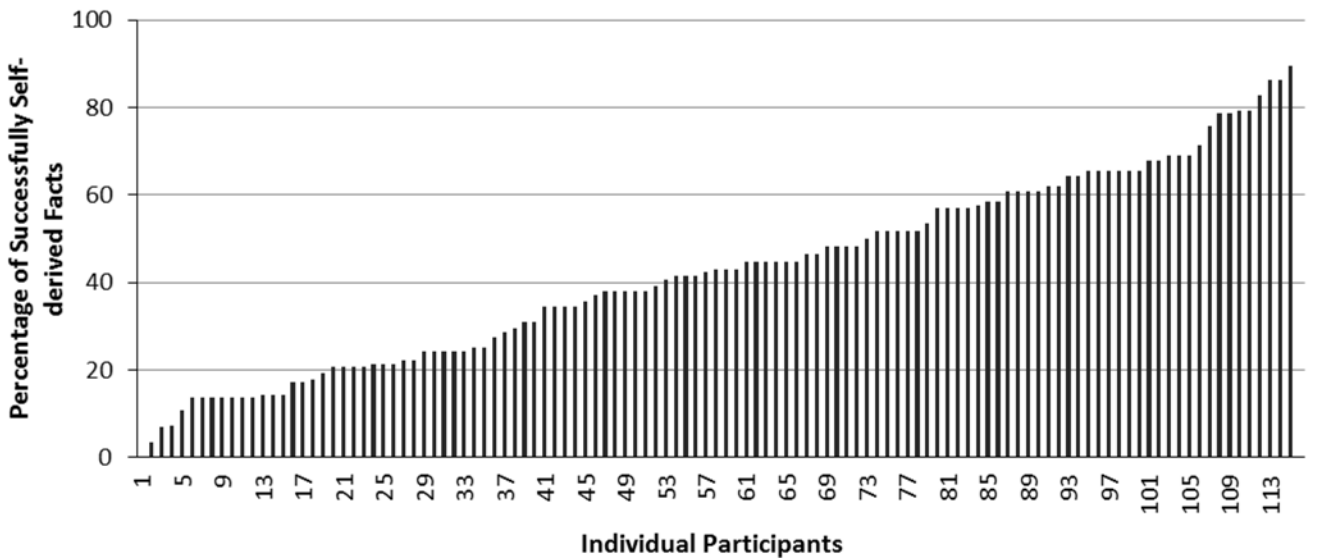## Panel A: Session 1

Panel B: Session 2

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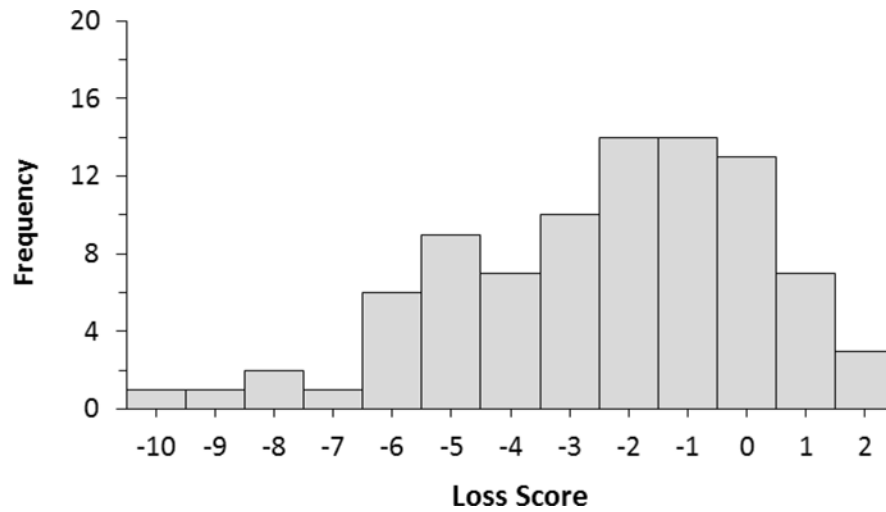participant performance arranged from lowest to highest.

Figure 4.

Frequency distribution of the number of integration facts participants failed to recall after a 1-week delay. The X-axis shows the loss score, which was calculated by subtracting the number of integration facts successfully self-derived at Session 1 from the number of integration facts successfully recalled at Session 2 (negative scores indicate a decrease in performance, whereas, positive scores indicate an increase in performance). As is suggested by the negative skew, the majority of participants exhibited minimal loss over the 1-week delay.

Table 1

Mean Response Time (ms) on Successful and Unsuccessful Trials in Experiment 2

	M	*SD*
Successful Trials	4528	2193
Unsuccessful Trials	13498	8172
All Trials	8609	4738