Effects of delays on 6-year-old children’s self-generation and retention of knowledge through integration

Nicole L. Varga and Patricia J. Bauer
Department of Psychology, Emory University

Abstract

The present research was an investigation of the effect of delay on self-generation and retention of knowledge derived through integration by 6-year-old children. Children were presented with novel facts from passages read aloud to them (stem facts) and tested for self-generation of new knowledge through integration of the facts. In Experiment 1, children integrated the stem facts at Session 1 and retained the self-generated memory traces over 1 week. In Experiment 2, 1-week delays were imposed either between the to-be-integrated facts (between-stem delay) or after the stem facts but before the test (before-test delay). Integration performance was diminished in both conditions. Moreover, memory for individual stem facts was lower in Experiment 2 than in Experiment 1, suggesting that self-generation through integration promoted memory for explicitly taught information. The results indicate the importance of tests for promoting self-generation through integration as well as for retaining newly self-generated and explicitly taught information.

Keywords

cognitive accessibility; learning; retention; self-generation; semantic integration; semantic memory

The question of how knowledge is acquired, organized, and retained is fundamental to an understanding of cognitive development. These processes also are critical for informing educational practice and have been identified as the primary concern of educational psychologists (Beliner, 2006). In everyday contexts and educational settings alike, individuals are bombarded with a vast amount of information and faced with the burden of sorting the material, connecting it to previously learned content, and building upon what is already known. Semantic memory serves a primary role in the storage and retrieval of this newly and previously acquired information, respectively. Semantic memory also exhibits a capacity for extending itself such that new information can be self-generated through processes such as semantic integration—the combination of two or more separate, but related traces of information to yield new knowledge (Bauer, 2012). Preschool-age children have been shown to generate new knowledge through integration of information explicitly taught in separate passages of text (e.g., Bauer & San Souci, 2010). However, whether the newly self-generated knowledge persists over time has not been examined. Additionally, how the temporal spacing of to-be-integrated information and the timing of the subsequent
test for integration affect self-generation are largely unknown. Thus, in the present research we tested whether newly self-generated knowledge remains cognitively available over a 1-week delay (Experiment 1), and how delays imposed either between presentation of to-be-integrated information or after presentation of the information but before the test for integration impact self-generation of new knowledge (Experiment 2).

The capacities to store and organize information about the world and to generate meaningful connections and understandings are present from infancy. Within the first year of life, infants demonstrate retention of information about the properties of objects, as evidenced by long-term recall of unique actions and temporal sequences of actions such as in elicited and deferred imitation paradigms (e.g., Carver & Bauer, 1999, 2001; Meltzoff, 1988). At least by the second year, infants have created organized semantic knowledge structures based on taxonomic categories of animals and vehicles, for example (e.g., Mandler, Bauer, & McDonough, 1991), as well as categories based on spatial-temporal contiguities (so-called contextual categories), such as kitchen things and bathroom things (e.g., Mandler & Bauer, 1988; Mandler, Fivush, & Reznick, 1987). Infants also generate meaningful connections and understandings by generalizing what they know to novel instances of categories. For example, they induce that a novel vehicle such as a forklift starts with a key, based on knowledge that familiar vehicles, such as cars, exhibit this property (Mandler & McDonough, 1996). The capacity for flexible and productive use of existing knowledge through logical processes such as induction (e.g., Gelman & Markman, 1987), deduction (Dias & Harris, 1988), and analogy (Goswami, 2011) increases over the preschool years (see Goswami, 2011 for a review). These capacities are foundational to establishment of a knowledge base.

The preschool years also are marked by changes in another important means of knowledge expansion, namely, self-generation of new knowledge through integration of separate episodes of experience. This capacity is called upon in a variety of learning contexts, and is especially salient in curricula that require children to learn subject matter in units spaced over days and weeks. Children are expected not only to retain the information from separate episodes but also to integrate the material across lessons. The productive aspects of semantic memory are apparent when the process of integration results in self-generation of new knowledge that was not present in the individual episodes. To test this ability, Bauer and San Souci (2010) employed a paradigm in which 4- and 6-year-old children were read passages of text, each of which conveyed a novel fact (i.e., “stem” facts). The novel stem facts could be combined or integrated with one another to generate new information that was not presented in the texts (i.e., “integration” facts). For example, children were read stories in which they were explicitly taught the stem facts that “dolphins talk by clicking and squeaking” and that “dolphins live in groups called pods.” They then were asked the question “How does a pod talk?” Four-year-old children generated the novel integration fact that “pods talk by clicking and squeaking” on only 13% of the trials, yet they recognized the correct answer from among distracters on 62% of the trials. Six-year-olds generated the novel integration facts on 67% of the trials. Memory for the individual stem facts facilitated productive self-generation of the integration facts. That is, children who remembered more stem facts also generated more integration facts. Memory for the stem facts was not sufficient, however, as evidenced by the fact that 4-year-olds who were brought to a criterion level of learning of the stem facts nevertheless failed to achieve the high level of self-generation demonstrated by the 6-year-olds (4-year-olds generated the integration facts on 33% of the trials whereas 6-year-olds generated the integration facts on 67% of the trials).

For semantic integration to be psychologically, cognitively, and educationally meaningful, its products must persist over time. That is, the continual development of a knowledge base
requires that individuals not only generate links between related concepts, but also that they retain this newly self-generated knowledge. A large body of literature demonstrates that preschool-age children remember information over delays (e.g., see Bauer, in press; and Bauer, Larkina, & Deocampo, 2011, for reviews). Yet to date, studies examining memory over a delay have focused on material that is explicitly taught or directly experienced. For example, children recall semantically related items that are explicitly presented via list-learning paradigms (see for example Haden, in press; Roebers, in press, for reviews) as well as the features of directly experienced events such as medical exams (e.g., Ornstein, Gordon, & Larus, 1992). Tests of memory for knowledge self-generated through integration have not been conducted. To address this void in the literature, in Experiment 1 of the present research, we tested whether information that is self-generated through semantic integration remains cognitively available over time. Specifically, we taught children information in separate passages of text (stem facts) and then tested them for self-generation of novel integration facts. We then imposed a 1-week delay after which we tested children’s memory for the self-generated knowledge.

We hypothesized that children would retain integration facts that they had successfully generated. We based this hypothesis on findings that the act of self-generating known lexical or semantic content promotes memory for the self-generated items (Gardiner, Gregg, & Hampton, 1988; e.g., Gardiner & Hampton, 1985). This phenomenon, known as the generation effect, is commonly examined using verbal material in which participants are provided with a rule (e.g., the word pairs are associates) and asked to either read a stimulus-target word pair (e.g., rapid—fast) or to generate the target word when given the stimulus but only a portion of the target word (e.g., rapid—f_ _ _) (Slamecka & Graf, 1978; see Mulligan & Lozito, 2004 for review). Recall and recognition of self-generated material is consistently higher than the control condition in which participants read the full stimulus-target word pairs. This effect has been observed as early as 7 years of age and continues to develop throughout the school years (McFarland, Duncan, & Bruno, 1983). Whether the facilitative effect of self-generation extends beyond known semantic content to promote retention of newly derived knowledge is not known. We tested the question in Experiment 1.

In Experiment 2 of the present research, we tested the impact of another kind of delay, namely, between to-be-integrated information or after presentation of the information but before the test for integration of it. Often, as in educational contexts, related information is distributed over time, rather than presented in a single session. For the knowledge base to grow, experiences that are temporally separated from one another must be integrated. In the between-stem delay condition in Experiment 2, we tested the impact on self-generation of new knowledge of a 1-week delay between presentation of to-be-integrated information. That is, we presented one member of a stem-fact pair in one session and then imposed a 1-week delay before presentation of the related stem fact and test for integration of the members of the pair. Similarly, in the world outside the laboratory, individuals are presented with a wealth of related information in the absence of explicit cues or demands to integrate. The cue or demand to integrate may come only later in time. In the before-test delay condition in Experiment 2, we tested the impact on subsequent self-generation of a 1-week delay imposed after presentation of to-be-integrated information, but before the test for integration of it. Specifically, we presented both members of a pair of stem facts in a single session but then did not test for integration of the stem facts until 1 week later.

We hypothesized that the between-stem delay (i.e., Stem Fact 1 in Session 1 and Stem Fact 2 in Session 2) would have a detrimental effect on children’s self-generation of new knowledge. We based this hypothesis on the expectation that productive self-generation through integration depends on the cognitive accessibility of the to-be-integrated information (Bauer, 2012). Operation of this principle in the context of self-generation
through integration is apparent in the finding that manipulations that increase cognitive accessibility of relevant information, such as learning the individual stem facts to criterion, increases integration performance (Bauer & San Souci, 2010). It also is apparent in the finding that decreasing cognitive accessibility of relevant information, by reducing the surface similarity between to-be-integrated material, decreases integration performance (Bauer, King, Larkina, Varga, & White, 2012). When children hear two passages of text in a single session and then experience the demand to integrate (in the form of a question about the passages), the relevant information is readily accessible (in principle, if not in practice).

In the case of a between-stem delay, however, we may expect that although the information from one of the passages will be highly accessible at the time of the demand for integration, the information from the other passage—the one encountered one week earlier—will be less accessible. The expected result is a lower level of self-generation through integration.

Hypotheses regarding the effect of the before-test delay are less straightforward because of lack of information about the time-course of integration processes. That is, because this means of knowledge extension is only beginning to be investigated, it is unknown at this time whether integration processes occur upon experience of related information or only with a demand for integration. If integration processes are set into motion upon experience of related information (likely based on simultaneous high cognitive accessibility), we would expect the temporal proximity of the individual stem facts to promote self-generation of integration facts, even in the absence of an explicit demand to integrate. As above, we would expect the new information to be retained over the delay, resulting in a high level of performance at the time of explicit test 1 week later. However, there is the distinct possibility that integration processes occur only on demand. This suggestion follows from the principle of cognitive economy—that resources such as would be required to generate new information are not expended without motive or demand (Wyer & Srull, 1986).

Together, the principle of cognitive economy and the expectation that productive self-generation through integration depends on the cognitive accessibility of to-be-integrated information, lead to the hypothesis that performance in the before-test delay condition will suffer, owing to the fact that at the time of the demand for integration, the relevant individual stem facts are relatively cognitively inaccessible (having been presented 1 week previously).

In both experiments, we selected the delay of 1 week. One week is a period of time over which we may expect children to remember (e.g., Baker-Ward, Gordon, Ornstein, Larus, & Clubb, 1993), yet to exhibit some degradation of the memory trace. We focused the inquiry on 6-year-olds because the research was intended to build on the fledging literature on self-generation of new information through integration that has focused on 4- and 6-year-old children (Bauer et al., 2012; Bauer & San Souci, 2010). The age group is especially appropriate for the inquiry because in the later preschool and early school years, children increasingly find themselves in settings and situations that demand that they integrate information that has been acquired at different times and in different contexts. Moreover, we tested only 6-year-olds because the 4-year-olds in Bauer and San Souci (2010) had low levels of integration in the self-generation testing format (they showed evidence of integration on only 13% of trials). We expected imposition of delays to depress self-generation performance, which likely would have resulted in floor effects for 4-year-olds.

**Experiment 1**

**Method**

**Participants**—Participants were 16 6-year-old children (8 girls and 8 boys; mean age = 6 years 6 months, range = 6 years 2 months to 6 years 8 months). Children were recruited from a volunteer pool consisting of families who had expressed interest in participating in
research. The sample was 6% African American, 6% Asian, and 88% Caucasian. None of
the participants was of Hispanic descent. Children had typical levels of verbal
comprehension, as measured by Test 1 of the Woodcock-Johnson III Tests of Cognitive
Ability (M standard score = 119.81, SD = 14.84; see below; Woodcock, McGrew, &
Mather, 2001). Children participated in two sessions separated by 1 week (M = 6.81 days,
SD = 0.40, range = 6 days to 7 days). Three additional participants were tested but excluded
from analysis due to an insufficient delay period (n = 1), failure to complete the task (n = 1),
and experimental error (n = 1). At the end of each hour-long session, children received a
small toy to acknowledge their participation and children’s parents were given a $10 gift
card after the second session. Prior to the start of each session, the experimenter thoroughly
explained the method and obtained verbal assent from the child as well as written informed
consent from the child’s parent or guardian. These procedures were approved by the
university Institutional Review Board.

Stimuli—The stimuli were the same as in Bauer and San Souci (2010) and Bauer et al.
(2012). They were six “stem” facts that could be combined to produce three “integration”
facts determined to be novel to children 6 years of age. There were two stem facts pertaining
to dolphins (dolphins talk by clicking and squeaking; dolphins live in groups called pods),
two pertaining to kangaroos (all baby kangaroos are called joeys; some kangaroos are called
blue flyers), and two pertaining to a volcano (the world’s largest volcano is in Hawaii;
Mauna Loa is the world’s largest volcano). Each pair of related stem facts could be
combined to generate a novel integration fact: Pods communicate by clicking and squeaking,
baby blue flyers are called joeys, and Mauna Loa is in Hawaii. The results of Bauer and San
Souci (2010) demonstrated that both stem facts of a pair were required to support self-
generation of integration facts. That is, when 6-year-olds were exposed to just one of the two
paired stem facts, they generated only 17% of the novel integration facts (compared with
67% when both paired stem facts were presented). Thus exposure to the information in both
stem facts is necessary to generate the corresponding novel integration fact.

Each stem fact was conveyed via a short passage read aloud by the experimenter. Passages
ranged between 82 and 89 words in length. Each passage consisted of four pages with a
hand-drawn picture on each individual page depicting the main actions of the spoken text.
Each of the six passages followed a similar framework in which a character (e.g., a ladybug)
learned something new. Within a pair of passages, the stem facts were presented by the same
character (e.g., both dolphin passages featured a ladybug) which created high surface
similarity between the to-be-integrated passages (see Bauer et al., 2012, for effects of
surface similarity manipulations). The novel stem fact was first conveyed on the second or
third page of each passage and then repeated on the final page. A sample pair of passages is
provided in Appendix A. Importantly, the novel integration facts were not presented in the
passages.

Procedure—Participants were tested individually in a room equipped with a table, two
chairs, and a small couch for parents. Two female experimenters (including N.L.V.)
conducted the sessions with each experimenter testing an approximately equal number of
female and male participants. Children were tested by the same experimenter at each
session. The experimenters followed a detailed written protocol and regularly reviewed
video-recorded sessions to ensure protocol fidelity.

Session 1: In the learning phase of Session 1, children were presented with three pairs of
related passages, for a total of six passages. At the beginning of the session, children were
read three passages, each of which conveyed one of the two stem facts from each of the
paired passages about dolphins, kangaroos, and volcanoes. Children were instructed to listen
to the stories and look at the pictures. The passages were read continuously without
interruption and each passage was read twice before moving on to the next passage. After presentation of the three passages, participants engaged in approximately 15 minutes of filler activity comprised of Test 1 of the WJ III COG which served as a measure of verbal comprehension (Woodcock et al., 2001).

Following the intervening task, children were presented with the second set of three passages, each of which conveyed the second of the two stem facts from each of the paired passages about dolphins, kangaroos, and volcanoes. The passages were presented in the same order in the first and second segments of the session (i.e., if order A, B, C in segment 1 then order A, B, C in segment 2). Across subjects, the order of presentation of the passages was counterbalanced such that each was presented approximately equally often in each serial position. Children then completed an unrelated filler task for approximately 15 minutes to attenuate primacy and recency effects during the test phase.

The test phase consisted of open-ended and recognition questions pertaining to the integration facts. First, children were tested for generation of the novel integration facts in an open-ended format. The experimenter asked each of the three integration questions: “How does a pod talk?”, “What is a baby blue flyer called?”, and “Where is Mauna Loa located?” The integration fact questions were interspersed among five facts commonly known to 6-year-olds, including “What is a baby cow called?” and “Where does Mickey Mouse live?” (for a total of 8 questions). The known questions were provided so that children had a successful experience, even if they failed to generate the integration facts. The experimenter recorded the child’s responses as they made them.

Children then were tested for recognition of the correct answers to any integration questions they failed to generate during open-ended testing. That is, if the child generated the correct response to the integration fact question when it was posed in the open-ended format, the recognition question for that integration fact was not presented. All recognition questions had three choices, one of which was correct. The integration recognition questions were interspersed among questions to which children in the target age range were likely to know the answers. Because children were asked the recognition questions only for the integration facts that they failed to generate in the open-ended phase of testing, children received a different number of recognition questions. No corrective feedback was provided at any point in the testing phase. Following completion of the testing phase, the children completed a brief picture matching task in order to end the session with an activity unrelated to the purpose of the second session. Parents were explicitly asked not to discuss the passages or the questions with their children between the first and second study sessions. Parents reported that they complied with this request.

Session 2: One week later, children returned to the laboratory and experienced a four-step testing procedure (see Appendix B for sample integration and stem fact questions). First, children were tested for recall of the integration facts. As in Session 1, the experimenter asked each of the three integration questions interspersed among the five questions the answers to which children were likely to know. The questions were asked in a novel order, relative to Session 1. Second, children were asked six stem fact recall questions, assessing their recall of the individual stem facts to which they had been exposed via the passages at Session 1 (e.g., “What is a joey?”; “What is a baby blue flyer?”). Third, children were tested for recognition of the correct answers to the integration fact questions. Fourth, children were tested for recognition of the correct answers to the stem fact questions. For both the integration and stem facts, recognition questions were asked only for the facts that children failed to produce during the open-ended phase of testing. Thus children received a different number of recognition questions, depending on the number of facts produced in the open-ended phase of testing. All recognition questions had three alternatives, one of which was
correct. The questions were interspersed among questions to which children in the target age range were likely to know the answers.

Testing for the integration and stem facts was administered in the fixed order just described, although the order of questions presented in each section was counterbalanced among the participants. The open-ended format provided the strongest test of children’s ability to remember the integration and stem facts. To be valid, the open-ended format test had to be administered before the test of recognition. Within each test format (open-ended, recognition), the integration fact questions were asked before the stem fact questions, so that children were not reminded of the stem facts immediately prior to testing of the integration facts.

Scoring—For the integration facts, children received self-generation (Session 1) and recall (Session 2) scores based on their performance on the three open-ended integration questions (max = 3 per session). Children also received an integration total score at each session which was the sum of correct responses in the open-ended and recognition testing formats. Because children were only presented with the follow-up recognition questions for items that were answered incorrectly during the open-ended phase, the maximum total score was 3. For the stem facts, children received 1 point for each pair of stem facts recalled at Session 2 (e.g., both stem facts about dolphins recalled; max = 3). They also received a stem fact pair total score, which was the sum of correct responses in the open-ended and recognition testing formats (max = 3).

Results

Generation and Retention of Integration Facts—The 6-year-olds both generated and recognized the novel integration facts during the first session. Specifically, participants generated 63% (M self-generation score = 1.88, SD = 1.09) of the novel integration facts in the open-ended format and they either spontaneously generated or correctly recognized 85% (M total score = 2.56, SD = 0.63) of the integration facts. Self-generation and total scores were comparable to those reported in Bauer and San Souci (2010; 67% and 93%, respectively). Further, as reflected in Figure 1, Panel a, 81% of the children (n = 13) successfully generated 1 or more of the novel integration facts.

The major question posed in this experiment was whether the children retained the self-generated memory traces over the 1-week delay. At the second session, children recalled 60% (M recall score = 1.81, SD = 0.98) of the novel integration facts in the open-ended format and they either recalled or correctly recognized 83% (M total score = 2.50, SD = 0.52) of the integration facts. A one-way repeated-measures analysis of variance (ANOVA) indicated that self-generation and recall of the integration facts at Sessions 1 and 2, respectively, did not differ significantly [F(1, 15) = 0.10, p = .751, η² = 0.007, α = 0.05]. Thus there was no evidence of forgetting of the self-generated integration facts over the 1-week delay. Statistical comparison of total performance at Session 1 and Session 2 also failed to reveal evidence of forgetting of the self-generated integration facts over the delay [F(1, 15) = 0.19, p = .669, η² = 0.013, α = 0.05]. On 83% of the trials on which children self-generated the integration facts at Session 1, they recalled the integration fact 1 week later at Session 2. Over all trials, children exhibited the same performance at Session 1 and Session 2 on 36 trials (75%). On 7 trials (15%) they showed decreased performance across

1The results from 1 trial (of 3) from each participant’s Session 1 open-ended integration fact testing was published in Bauer et al. (2012) to lend support to the consistency of self-generation frequency across samples. This is the first report of all trials from Session 1 and none of the data from Session 2 have been published.
sessions (e.g., self-generating at Session 1 but only recognizing at Session 2), and on 5 trials (10%) they showed increased performance across sessions.

**Memory for Stem Facts**—Though the children were not tested for recall or recognition of the individual stem facts during Session 1, they had high levels of memory of them after the 1-week delay. Specifically, at Session 2, the children recalled 63% of the total pairs of stem facts ($M = 1.88$, $SD = 1.02$). Additionally, stem pair total scores (i.e., the sum of correct responses in the recall and recognition testing formats) approached ceiling levels, with children either recalling or recognizing 90% of the total pairs ($M = 2.69$, $SD = 0.60$). Further, as depicted in Table 1, of the 30 total trials in which both of the stem facts from a pair of passages were recalled (i.e., Stem Facts Recalled = 2), children recalled the corresponding novel integration facts at Session 2 on 25 trials (i.e., 83%) and failed to recall the novel integration facts on only 5 trials (i.e., 17%).

**Discussion**

The primary purpose of Experiment 1 was to test whether information children self-generated through integration of separate passages of text remained available to them 1 week later. The results were clear—children self-generated new knowledge at Session 1 and had high levels of recall of the self-generated knowledge 1 week later at Session 2. In fact, performance at the two sessions did not differ statistically, indicating no significant forgetting of the newly self-generated information over the delay. Moreover, although the children were not tested for recall of the individual stem facts at Session 1, they recalled approximately two-thirds of the complete pairs 1 week later at Session 2. Thus over a 1-week delay, both newly self-generated information and explicitly taught facts remained highly accessible.

The situation posed in Experiment 1, in which children experienced a delay before being tested for retention of knowledge they self-generated through integration, is but one of the types of delay that occurs in the course of construction and use of a semantic knowledge base. Other common types of delays are between to-be-integrated information and after presentation of the information but before the test for integration of it. The purpose of Experiment 2 was to test integration processes in these circumstances. In a before-test delay condition, we tested the impact on subsequent self-generation of new knowledge of a 1-week delay imposed after presentation of to-be-integrated information, but before the test for self-generation. In a between-stem delay condition, we tested the impact on self-generation of a 1-week delay between presentation of to-be-integrated information.

**Experiment 2**

**Method**

**Participants**—Participants were 32 6-year-old children (16 girls and 16 boys; mean age = 6 years, 6 months, range = 6 years 2 months to 6 years 10 months). Equal numbers of children were pseudo-randomly assigned to either the before-test or between-stem delay conditions (see below). Assignment was random with the constraint of gender balance. The children were drawn from the same research pool as in Experiment 1. The sample was 9% African American, 6% Asian, 81% Caucasian, and 3% Pacific Islander. Three-percent of the participants were of Hispanic descent. Children were typical in language, as measured by WJIII ($M$ standard score = 116.31, $SD = 14.64$). Children participated in two sessions separated by 1 week ($M = 7$ days, $SD = 0.36$, range = 6 days to 8 days). Five additional children participated but were excluded from the final analysis due to an insufficient delay period ($n = 2$), failure to return for the second session ($n = 1$), failure to complete the task ($n = 1$), and experimental error ($n = 1$). None of the children had participated in Experiment 1.
Children’s participation was acknowledged as in Session 1. Informed consent and children’s assent were obtained as in Experiment 1.

**Stimuli and Procedure**—The stimuli were the same as in Experiment 1. Children were tested individually in the same testing room and by the same two female experimenters as Experiment 1. Children were tested by the same experimenter at each session. Further, each experimenter tested an approximately equal number of female and male participants across conditions.

As reflected in Table 2, the before-test and between-stem delay conditions were designated by the placement of a 1-week delay period between experimental sessions. In the before-test delay condition, following the procedures described in Experiment 1, children experienced three pairs of related passages of text in Session 1. The members of the pairs of passages were separated by approximately 15 minutes during which Test 1 from the WJ III COG (Woodcock et al., 2001) was administered. In addition, after the second set of three passages, children completed an unrelated filler task for approximately 15 minutes. However, unlike in Experiment 1, the children were not tested for integration of the passages. Instead, children returned to the laboratory 1 week later ($M = 7.06$ days). Following a 15 minute filler activity to reacclimate them to the laboratory, they were tested for generation and recognition of the integration facts and for recall and recognition of the stem facts following the procedure in Experiment 1.

In the between-stem delay condition, at Session 1, children experienced only one member of each of three pairs of related passages (the Stem 1 passages). That is, they experienced one story about dolphins, one story about kangaroos, and one story about volcanoes. They did not experience the second member of the pairs of related passages (the Stem 2 passages) until Session 2 (the specific passage that served as Stem 1 and Stem 2 was counterbalanced across sessions, across participants). The test for integration also was administered at Session 2. Specifically, at Session 1, children heard three Stem 1 passages and then completed Test 1 from the WJ III COG (Woodcock et al., 2001). One week later ($M = 6.94$ days) they returned to the laboratory and first engaged in 15 minutes of filler activity to reacclimate them to the laboratory. They then were read the three Stem 2 passages. Children then completed another 15-minute filler activity, after which they were tested for generation and recognition of the integration facts and for recall and recognition of the stem facts following the procedure in Experiment 1.

It is important to note that maintaining the same task order between groups and manipulating only the placement of the 1-week delay resulted in sessions of different length for the two groups. The before-test delay group experienced a first session of approximately 45 min and a second session of approximately 30 min. Conversely, the between-stem delay group experienced a first session of approximately 30 min followed by a second session of approximately 45 min. However, all children experienced one shorter and one longer session. Although it may be argued that children in the between-stem delay condition may have been fatigued prior to the test, note that the length of the session was shorter than the first session of Experiment 1. Also, as is apparent from Table 2, both groups experienced (a) 15 minutes of filler activity after each set of three passages, (b) 15 minutes of filler activity to reacclimate them to the laboratory at Session 2, and (c) 15 minutes of activity prior to the test.

**Scoring**—Scoring was conducted as in Experiment 1 with the exception that stem fact recall was additionally categorized according to the order in which the individual stems were presented via the passages. That is, the facts conveyed through the first passage presented about dolphins, kangaroos, and volcanoes were coded as Stem 1 facts ($max = 3$) and the
facts from the corresponding three passages were coined Stem 2 facts (max = 3). This designation was made because in the between-stem delay condition, half of the stem facts were presented in Session 1 and half in Session 2, making it likely that recall might differ.

**Results**

**Generation and Retention of Integration Facts**—Children in the before-test delay condition self-generated the novel integration facts on 21% (M self-generation score = 0.63, SD = 0.89) of the trials and children in the between-stem delay condition self-generated 23% (M self-generation score = 0.69, SD = 0.95) of the novel integration facts. An independent means t-test indicated no significant difference in performance between the groups. As shown in Figure 1, Panel b, only 14 of the 32 children (7 in each condition) produced at least one of the three integration facts in the open-ended format. This strikingly different pattern from that observed in Experiment 1 (Figure 1, Panel a) suggests that imposition of a delay—whether between the presentation of to-be-integrated material or before test for integration of previously learned material—severely interfered with the process of self-generation through integration.

Consideration of children’s total scores (self-generation plus recognition) revealed higher levels of performance relative to the open-ended, self-generation format. Children in the before-test and between-stem delay conditions self-generated or correctly recognized 67% (M total score = 2.0, SD = 0.73) and 75% (M total score = 2.25, SD = 0.86) of the novel integration facts, respectively. An independent means t-test revealed that total performance did not differ significantly between groups. Although children’s total performance was higher than their self-generation performance, it was nominally lower than total performance in Experiment 1 (85% and 83% at Sessions 1 and 2, respectively).

**Memory for Stem Facts**—As reflected in Table 1, across conditions, on 20 of the 21 trials on which the children generated the integration facts, they also recalled both of the stem facts from the corresponding pair of passages. Recall of the stem facts was not sufficient to support self-generation of new knowledge through integration, however. Across conditions, on 17 trials, children recalled both of the stem facts from the corresponding passages yet did not spontaneously generate the novel integration fact. This pattern is consistent with the findings of Bauer and San Souci (2010, Experiment 2) that recall of the stem facts is necessary, but not sufficient, to ensure self-generation of new knowledge through integration.

Similar to the self-generation findings, memory for the stem fact pairs also did not differ between conditions. Specifically, children in the before-test delay condition recalled 31% of the total pairs of stem facts (M = 0.94, SD = 1.06), whereas, children in the between-stem delay condition recalled 46% of the total pairs (M = 1.38, SD = 1.09). Furthermore, stem pair total score performance did not significantly differ. Children in the before-test and between-stem delay conditions recalled or recognized 65% (M = 1.94, SD = 0.68) and 79% (M = 2.38, SD = 0.62) of the total stem pairs, respectively.

Finally, a 2 x 2 mixed-model ANOVA (Stem Fact x Condition) revealed a significant interaction [F(1, 30) = 5.05, p = .032, r^2 = 0.14, α = 0.05]. To examine the interaction, we conducted follow-up one-way repeated-measures ANOVAs for each condition separately. The analyses revealed that recall for the individual stem facts was significantly different in the between-stem condition [F(1, 15) = 8.45, p = .011, r^2 = 0.36, α = 0.05], such that children recalled more Stem 2 facts (M = 2.44, SD = 0.73) than Stem 1 facts (M = 1.56, SD = 1.15). Conversely, in the before-test delay condition, recall did not differ for the Stem 1 and Stem 2 facts (M = 1.75, SD = 1.06; M = 1.63, SD = 0.96, respectively) [F(1, 15) = 0.146, p = .71, α = 0.05]. This finding was to be expected given that children in the
between-stem condition were presented with the Stem 2 facts shortly before the test phase. However, heightened memory for the second set of facts did not boost either self-generation performance or memory for complete stem fact pairs.

**Comparison of Samples in Experiments 1 and 2**—The lower levels of performance of the children in the present experiment relative to Experiment 1 cannot be attributed to differences in the samples. The samples were drawn from the same source, were approximately equally matched on demographic composition, and were the same mean age across conditions. Further, there were no differences in performance across the three delay groups on standard scores of verbal comprehension ($F_{(2, 45)} = 0.48, p = 0.624$) indicating equivalence between conditions on basic cognitive functioning more generally.

**Discussion**

The primary purpose of Experiment 2 was to test the impact of imposition of 1-week delays before test for integration of pairs of related stem facts and between presentation of members of the pairs of related stem facts. Both delays proved a challenge to self-generation of new knowledge through integration. In contrast to Experiment 1, in which children self-generated the integration facts on 63% of trials, in the present experiment, performance was 21% and 23% in the before-test and between-stem delay conditions, respectively. Moreover, children’s levels of recall of the stem facts also seemingly suffered as a result of the delays. Whereas in Experiment 1, on 63% of the trials, children recalled both members of the pair of related stem facts, in the present experiment, the corresponding values were 31% and 46% in the before-test and between-stem delay conditions, respectively. Importantly, the lower level of integration performance observed in the present experiment relative to Experiment 1 cannot be attributed solely to lower levels of recall of the stem facts. When children recalled both members of the pair of stem facts, they generated the novel integration fact on 54% of the trials and failed to generate it on 46% of the trials. Thus even when children recalled both members of the pair of stem facts self-generation through integration was not ensured.

**General Discussion**

The purpose of the present research was to investigate the effects of delay on both the self-generation and retention of new knowledge derived through integration. In the first experiment, we extended prior research by examining whether knowledge self-generated through integration remained accessible after a 1-week delay. The 6-year-old children demonstrated high retention of the self-generated memory traces. In fact, there was no evidence of forgetting of the material over the delay. In Experiment 2, we examined the impact on self-generation of a temporal delay between presentation of to-be-integrated information (between-stem delay condition) or after presentation of the information but before the test for integration of it (before-test delay condition). The frequency with which children integrated was one-third of that observed in Experiment 1 and prior research (Bauer & San Souci, 2010). Thus although performance was unaffected by the delay in the first experiment, it was severely negatively impacted across both conditions in the second experiment.

It is noteworthy that the present research is the third study (encompassing 7 experiments) in which we have observed self-generation of new knowledge through integration of separate episodes of experience. Moreover, the phenomenon has been observed under different task conditions, including different levels of support for recall of stem facts, differences in surface similarity, and the presence or absence of hints. The finding that children reliably integrate separate episodes under varying conditions places self-generation of new knowledge through integration in the same category as other means of knowledge extension such as induction, deduction, and analogy (e.g., Goswami, 2011). Yet to date, it has received
far less research attention. Although we have begun to identify some of the factors that promote this cognitive process within a single learning session (Bauer et al., 2012; Bauer & San Souci, 2010), the present research is the first to examine the phenomenon under the more challenging—yet ecological valid—conditions of distributed learning and demand for integration.

The present research also advances our understanding of the processes involved in self-generation of new knowledge through integration of separate episodes. At the most general level, like other forms of knowledge extension (i.e., analogy, deduction, induction), integration relies on flexible manipulation of information in order to self-generate novel understandings. At a more fine-grained level of analysis, integration is similar to other types of reasoning in that it also depends on knowledge about the to-be-manipulated information as well as an understanding of the higher-order relational commonalities amongst this material (e.g., Gentner, 1983; Goswami, 2011). For instance, successful self-generation via integration relies on knowledge about the pairs of to-be-integrated stem facts in addition to understanding of their relatedness. That is, children must both retrieve a pair of stem facts and also recognize the opportunity to extend beyond them in order to integrate. It is thus reasonable to conclude that memory for the stem facts must be quite robust to support self-generative learning via integration.

The present research also furthers our understanding of the time-course of integration processes. When to-be-integrated material and the demand to integrate (in the form of questions about the material) were imposed within the same session (i.e., Experiment 1, Session 1), children successfully self-generated new knowledge through integration. Conversely, administering the test 1 week after the necessary information was acquired (as in the before-test delay condition of Experiment 2) resulted in self-generation performance that was indistinguishable from that of children who had no opportunity to integrate until after the delay (as in the between-stem delay condition in Experiment 2). This pattern suggests that integration does not occur until children are presented with a demand to do so. The alternative explanation, namely, that in the before-test delay condition integration occurred upon experience of the related information (at Session 1), but that the newly derived knowledge was forgotten over the 1-week delay, also is possible. However, we argue that this conclusion is not likely.

First, a large body of research demonstrates that the act of self-generating already known semantic information confers benefits to later memory retrieval (see Mulligan & Lozito, 2004, for review). Consistent with this generation effect, in Experiment 1 of the present research, children who successfully self-generated new knowledge through integration at Session 1 retained the information on more than 80% of the trials at Session 2. It is therefore reasonable to assume that if children in the before-test delay condition of Experiment 2 spontaneously engaged in self-generation processes upon exposure to related information at Session 1, integration performance at Session 2 would have been more comparable to that observed in Experiment 1. Instead, performance was substantially lower. Moreover, performance in the before-test and between-stem delay conditions of Experiment 2 was almost indistinguishable. Second, even when children in the before-test delay condition successfully retrieved a complete pair of stem facts at Session 2, they failed to integrate them on 40% of the trials. This pattern is quite different from that observed in Experiment 1, in which retrieval of both members of a pair of stem facts was associated with integration of them on 83% of the trials. Based on the strikingly different patterns of performance between experiments, in the face of similar behavior in the two conditions of Experiment 2, it is reasonable to conclude that—at least under the conditions of the present research—6-year-old children do not spontaneously integrate upon experience of related information. It will
be left to future research to determine whether different conditions might promote spontaneous integration, whether among 6-year-olds or other age groups.

The findings of the present research have implications for the promotion of knowledge development. Specifically, the present research highlights the importance of encouraging cognitive processing of information by children at a time when it is highly cognitively accessible (Bauer, 2012). In Experiment 1, children were called upon to use newly-acquired information shortly after exposure to it. They not only rose to the challenge of self-generating new knowledge through integration, but also retained the information over time thus demonstrating that the generation effect extends beyond known information to newly acquired knowledge. In contrast, in the before-test delay condition of Experiment 2, when there was no demand for integration at the time of experience of related information, the opportunity for integration was largely lost—children had low levels of performance at Session 2. Similarly, in the between-stem delay condition of Experiment 2, whereas the stem facts introduced at Session 2 were highly accessible at the time of test, those introduced at Session 1 were less available, resulting in low levels of integration. These findings provide insight into how daunting the task of integration of information may be for young learners, both inside and outside the classroom. That is, delays throughout the learning process are the rule, not the exception. Children acquire related information across separate learning sessions and experiences and across different contexts more generally. Furthermore, the task demands experienced by children in the present paradigm closely approximate typical learning circumstances in which children encounter new information embedded within a rich context. That is, they were read passages consisting of several characters, in stories with unfolding plots, accompanied by illustrations. They were faced with the challenge of identifying, sorting, and connecting the relevant information. These conditions clearly were challenging for the 6-year-olds in the present research. It will be left to future research to determine whether supports such as provided in prior research, including “hints” to think about related information (Bauer et al., 2012) and learning stem facts to a criterion (Bauer & San Souci, 2010), would increase the cognitive accessibility of information over a delay, thus facilitating integration.

The results of the present research also have implications for educational practice. Specifically, the findings highlight the importance of tests for promoting retention of explicitly taught information. That is, children’s memory of the stem facts in Experiment 1 was superior to that of the children in Experiment 2. We maintain that the differences in accessibility of the explicitly encoded facts can be attributed to administration of a test for integration at the time of initial learning. That is, it seems that in Experiment 1, the prompt to access the stem facts in order to respond to the demand to integrate facilitated retention of the facts over the 1-week delay. The finding that tests improve retention is not novel in and of itself (see Roediger & Karpicke, 2006 for an extensive review). The act of retrieving encoded information has been shown to increase the longevity of memory across varying paradigms (e.g., Lachman & Laughery, 1968; Thompson, Wenger, & Bartling, 1978), throughout childhood (e.g., Gates, 1917; Spitzer, 1939), and over many different retention intervals (e.g., Landauer & Bjork, 1978). However, the current study contributes the novel finding that a test for self-generation of new knowledge results in enhanced retrieval of explicitly encoded material, even in the absence of a direct initial test of that material.

In conclusion, across development, the task of learning specific facts while also recognizing their relatedness poses significant challenges (Davis, 2000). Yet Experiment 1 demonstrated that children as young as 6 years of age successfully meet this challenge when cued to the relevant content on which to manipulate via a test. Further, the demand to integrate not only increased self-generation and retention of new knowledge, but also produced high levels of recall of information that was explicitly taught. Hence, these studies suggest that tests may
promote the extension of knowledge and should be given when information is still cognitively accessible, rather than after delays, in order to facilitate retention of both self-generated and explicitly taught information. Moreover, given the prevalence of delays in the learning process, both inside and outside the classroom, prompts to generate connections between novel information as well as between old and new content may be crucial to construction of an integrated knowledge base.

Acknowledgments

This research was funded by NIH HD57291 and HD67359, to Patricia J. Bauer, and by Emory College of Arts and Sciences. The authors thank Gabriella Cooper and Ayzit Doydum for their illustrations, and Elizabeth A. White for her help with various aspects of the research. The authors also extend their gratitude to the children and families who took part in the research, without whom this work would not have been possible.

References


Carver LJ, Bauer PJ. When the event is more than the sum of its parts: Nine-month-olds’ long-term ordered recall. Memory. 1999; 7:147–174. [PubMed: 10645377]


Gates AI. Recitation as a factor in memorizing. Archives of Psychology. 1917; 6(40)


Appendix A

The Traveling Lady Bug

Page 1: As a ladybug slept one night a strong wind came and blew her out of bed.

Page 2: She woke up and found she was at sea. A friendly dolphin came up and said “hello” to her by clicking and squeaking.

Page 3: Before the ladybug could say much more than “hello,” the very strong wind blew again and she was swept back home.

Page 4: The ladybug was sad she didn’t get to play with the friendly dolphin. But now the ladybug knew how all dolphins talk—by clicking and squeaking.
The Lonely Lady Bug

Page 1: One day, a ladybug went to the zoo so that she could make some new friends.
Page 2: At the zoo, she met some friendly dolphins playing in the water. “Friendly dolphins,” she asked, “may I be part of your group?”
Page 3: The dolphins said, “We’d love to have you join our pod. But you’ll have to live in the water with us.”
Page 4: The ladybug shook her head sadly and then she left to go home. But now she knew that a group of dolphins was called a pod.

Novel Integration Fact

Pods talk by clicking and squeaking

Appendix B

Open-Ended Integration Question and Correct Answer

How does a pod talk?

By clicking and squeaking

Stem Fact Recall Questions and Correct Answers

1. How does a dolphin talk?
   By clicking and squeaking

2. What is a group of dolphins called?
   A pod
Highlights

- Tests promote integration of information from separate sources by school-aged children
- Explicit demands to integrate must be given when information is cognitively accessible.
- Self-generation of knowledge promotes memory for integrated and explicitly encoded information.
Figure 1.
Frequency of children who generated 0, 1, 2, or 3 integration facts correctly a) before and after the delay in Experiment 1 and b) across the two delay conditions in Experiment 2.
### Table 1
Session 2 Integration Performance and Corresponding Stem Fact Recall across all Passages and each Delay Condition$^{a,b}$

<table>
<thead>
<tr>
<th>Delay Condition</th>
<th>Stem Facts Recalled</th>
<th>Integration Fact Generated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>(%)</td>
</tr>
<tr>
<td>After Test (Experiment 1)</td>
<td>0</td>
<td>(0)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>(31)</td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>(83)</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>(60)</td>
</tr>
<tr>
<td>Between Stem (Experiment 2)</td>
<td>0</td>
<td>(0)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>(0)</td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>(50)</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>(23)</td>
</tr>
<tr>
<td>Before Test (Experiment 2)</td>
<td>0</td>
<td>(0)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>(4)</td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>(60)</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>(21)</td>
</tr>
</tbody>
</table>

Note. n = number of trials; % = percentage of trials successfully or unsuccessfully integrated.

$^a$Because no differences in self-generation performance were observed by passage, the values listed in Integration Fact Generated reflect all possible integration trial counts (i.e., 48 per each delay condition).

$^b$Integration trials were categorized based on whether neither (i.e., 0), 1, or both (i.e., 2) of the stem facts from a corresponding pair of passages were recalled.
Table 2

Schematic of the Before-Test and Between-Stem Delay Manipulations in Experiment 2a

<table>
<thead>
<tr>
<th>Session 1</th>
<th>Session 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passages</td>
<td>Filler</td>
</tr>
<tr>
<td>Before-Test</td>
<td>Stem 1</td>
</tr>
<tr>
<td>Between-Stem</td>
<td>Stem 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Filler</th>
<th>Passages</th>
<th>Filler</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stem 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Test</td>
</tr>
</tbody>
</table>

aAll filler tasks were unrelated to the integration task and took approximately 15 minutes.