Productive extension of semantic memory in school-aged children: Relations with reading comprehension and deployment of cognitive resources

Patricia Bauer, Emory University
Shala N. Blue, Emory University
Aoxiang Xu, Emory University
Alena Esposito, Emory University

Journal Title: Developmental Psychology
Volume: Volume 52, Number 7
Publisher: American Psychological Association | 2016-07-01, Pages 1024-1037
Type of Work: Article | Post-print: After Peer Review
Publisher DOI: 10.1037/dev0000130
Permanent URL: https://pid.emory.edu/ark:/25593/ts3c7

Final published version: http://dx.doi.org/10.1037/dev0000130

Copyright information:
© 2016 American Psychological Association.

Accessed December 5, 2019 12:55 PM EST
Productive Extension of Semantic Memory in School-aged Children: Relations with Reading Comprehension and Deployment of Cognitive Resources

Patricia J. Bauer,
Department of Psychology, Emory University
Shala N. Blue, Aoxiang Xu, and
Department of Psychology, Emory University
Alena G. Esposito
Department of Psychology, Emory University

Abstract
We investigated 7- to 10-year-old children’s productive extension of semantic memory through self-generation of new factual knowledge derived through integration of separate yet related facts learned through instruction or through reading. In Experiment 1, an experimenter read the to-be-integrated facts. Children successfully learned and integrated the information and used it to further extend their semantic knowledge, as evidenced by high levels of correct responses in open-ended and forced-choice testing. In Experiment 2, on half of the trials, the to-be-integrated facts were read by an experimenter (as in Experiment 1) and on half of the trials, children read the facts themselves. Self-generation performance was high in both conditions (experimenter- and self-read); in both conditions, self-generation of new semantic knowledge was related to an independent measure of children’s reading comprehension. In Experiment 3, the way children deployed cognitive resources during reading was predictive of their subsequent recall of newly learned information derived through integration. These findings indicate self-generation of new semantic knowledge through integration in school-age children as well as relations between this productive means of extension of semantic memory and cognitive processes engaged during reading.

Keywords
knowledge extension; learning; reading comprehension; self-generation; semantic integration; semantic memory

Learning and memory are dynamic cognitive processes that support development of an understanding of the world. Much of our storehouse of facts, meanings, and concepts—so-called semantic memory—is acquired through direct experience (e.g., firsthand observation,
instruction, reading). Semantic memory also is productively extended through a number of processes, including self-generation of new factual knowledge through integration of information acquired in separate yet related episodes of new learning (Bauer & Varga, 2015; Varga & Bauer, 2014). This process has been assessed in 4- and 6-year-old children (e.g., Bauer, King, Larkina, Varga, & White, 2012; Bauer & San Souci, 2010; Varga & Bauer, 2013; Varga, Stewart, & Bauer, in press), and in college students (Bauer & Jackson, 2015). In the present research, we extended the focus to school-age children, an age range in which there are substantial demands to acquire knowledge in both formal and informal educational settings. We tested self-generation of new semantic content through integration of separate learning episodes by 7- to 10-year-old children both when information was read to them by an experimenter (Experiments 1 and 2) and when they read material for themselves (Experiments 2 and 3). As a window on the processes involved in productive extension of knowledge through integration, in Experiment 3, we used measures of cognitive effort during reading to predict recall of information formed through integration.

Children and adults alike are regularly called upon to learn new information from a variety of sources, such as stories, textbooks, and lectures, to name a few. Yet not all that is known is directly learned. Some new information is “discovered” through action on objects (e.g., Sobel & Sommerville, 2010), and through problem solving (e.g., Fireman, Kose, & Solomon, 2003), for example. Other sources of knowledge are the productive logical processes of analogy, deduction, and induction, which permit going beyond what was explicitly learned or discovered to generate new understandings (see Goswami, 2011, for a review).

The productive process that is the subject of the present research is self-generation of new semantic knowledge through integration of information acquired in two (or more) related learning episodes that are separated by time, context, and so forth. Without integration, information acquired in discrete episodes (e.g., a textbook chapter and a lecture) would remain separate and would not readily accumulate into a knowledge base (Bauer & Varga, 2015). Moreover, integration permits the productive process of self-generation of new semantic knowledge derived from the combination of the episodes but not conveyed in either of the episodes separately. For example, in one episode, an individual may acquire the new information that dolphins communicate by clicking and squeaking. In a separate episode, the individual may acquire the new information that dolphins live in groups called pods. Integration of the episodes supports further extension of semantic memory through self-generation of the new knowledge that pods communicate by clicking and squeaking. The information was not conveyed in either learning episode, yet can be self-derived once the episodes are integrated with one another.¹

¹Self-generation of new knowledge as discussed in this work differs from the so-called “generation effect.” The generation effect is the advantage in memory derived from generating a known category associate or rhyming word, for example, relative to hearing the word produced by another (e.g., Jacoby, 1978; McFarland, Duncan, & Bruno, 1983; Slamecka & Graf, 1978). The generation effect is obtained on content already committed to semantic memory (e.g., TOOL-N __ __: “Name a tool that begins with the letter N”). In contrast, in the present research, participants generate new semantic content—facts that they previously did not know—based on integration of true but previously unknown facts explicitly taught in two different learning episodes.
The process of integration within an episode has been extensively studied in the domain of text comprehension, in particular. There, integration is in the service of making inferences regarding information only implied in the text (e.g., Cain & Oakhill, 1999; Graesser, Singer, & Trabasso, 1994; Paris & Upton, 1976). Some types of inferences maintain textual integrity through integration of information within the text, to understand anaphoric reference, for example. Other types of inferences permit filling of gaps in the text through recruitment of information from beyond it (i.e., background knowledge), to understand information within the text. Importantly, most of these inferences survive in working memory only long enough to aid comprehension (McKoon & Ratcliff, 1992). As such, though they could in principle lead to new knowledge, they are not investigated as a source of extension of semantic memory.

To date, the processes of integration of separate yet related episodes of new learning and self-generation of new semantic knowledge derived through the integration have been tested in two developmentally disparate populations—children 4 and 6 years of age and college-age adults. Children have been tested with a story-passage paradigm (e.g., Bauer & San Souci, 2010; Bauer, Varga, King, Nolen, & White, 2015). Using the example above, children listen to two stories about dolphins. In each, they learn a true but previously unknown fact (a “stem” fact). They then are asked a question that can only be answered by integrating the newly learned facts and using them to derive yet another new fact (an “integration” fact). The questions first are presented in an open-ended format, requiring that children self-generate the novel fact. In the case of the dolphin example, the question is “How does a pod talk?” On trials on which children fail to self-generate the novel fact (pods talk by clicking and squeaking), the question then is presented in a forced-choice format, requiring that children select the integration fact from among distracters.

In the first test of self-generation of new semantic knowledge through integration of separate episodes of new learning (Bauer & San Souci, 2010), in open-ended testing, 6-year-olds generated the novel integration facts on 67% of the trials. In a control condition in which only one of the passages was presented, children produced the integration facts on only 17% of the trials. In open-ended testing, 4-year-olds generated the novel integration facts on only 13% of the trials. Yet in forced-choice testing, they selected the integration facts from among distracters on 62% of trials. In contrast, in the 1-stem condition, 4-year-olds selected the novel integration facts 33% of the time, the value expected by chance (3-item forced-choice). Thus both 4- and 6-year-olds productively extended semantic knowledge. Six-year-olds generated the new information in an open-ended format; 4-year-olds showed the ability in a forced-choice format.

Subsequent research has revealed that young children not only extend semantic memory as a result of self-generation of new factual knowledge through integration of separate learning episodes, they also retain the new information over time. In Varga and Bauer (2013), in response to open-ended questions, 6-year-olds self-generated new semantic knowledge on 63% of trials (on par with 67% in Bauer & San Souci, 2010). The children were tested for recall of the self-generated information one week later. They recalled the information on 60% of total trials. This high level of retention of information newly self-generated through integration demonstrates the psychological significance of the process and its potential as a
mechanism for the rapid accumulation of semantic knowledge (Bauer & Varga, 2015; see also Varga et al., in press).

The second population that has been tested for self-generation of new semantic knowledge through integration is college-age adults (Bauer & Jackson, 2015). Adults were tested in a paradigm in which they read a large number of true but previously unknown facts (N = 120). They were not made aware that some of the facts in the long list were related to one another and when integrated, supported self-generation of new knowledge. Later in the session, adults were presented with declarative sentences the last word of which was omitted. They were required to choose the most appropriate ending to the sentence, from among 4 choices. For trials on which the correct responses could be derived through integration of the facts in the pairs of related sentences read earlier (i.e., the equivalent of the 2-stem condition used with 4- and 6-year-olds), participants selected the correct response 56% of the time. In contrast, in the equivalent of the 1-stem condition used in developmental research, they selected the correct response only 27% of the time, which did not differ from the rate expected by chance (25%). A separate group of adults participated in an event-related potential (ERP) study in which, after exposure to stem facts, they read a series of declarative sentences, some of which conveyed information formed through stem-fact integration. Based on the first 400ms presentation of the facts, ERP responses to the novel integration facts were intermediate between well-known facts (e.g., “Washington DC is the capital of the USA”) and entirely novel facts. Based on a second 400ms presentation, the response became indistinguishable from that to facts that were well-known. The rapid transition of newly self-generated facts to the status of “well-known” suggests that the new information had become incorporated into the semantic knowledge base.

The research summarized thus far indicates that the productive process of self-generation of new semantic knowledge through integration of separate episodes of new learning is operational in children 4 and 6 years of age and in college-age adults. To date, the work has not been extended to older school-age children, a population for which acquisition of world knowledge is a primary task. Accordingly, in Experiment 1 of the present research, we developed a new paradigm to test self-generation of new semantic knowledge through integration in 7- to 10-year-olds. The motivation for filling the void in the literature was not that we thought the processes would operate fundamentally differently in school-age children, relative to younger children and adults. Indeed, in all likelihood, such a fundamental means of productive extension of knowledge is conserved over ontogeny (as well as phylogeny: Preston & Eichenbaum, 2013). Rather, the purpose was to enable investigation of the processes throughout development, permitting examination of the conditions under which knowledge is productively extended, the variables that facilitate or interfere with the processes, and the efficiency with which the processes are carried out. It is these variables that likely change with development, even as the fundamental means of productive extension of knowledge is conserved.

The paradigm developed for the present research was an adaptation of that used in Bauer and Jackson (2015), in which college-age adults read long lists of true but previously unknown facts some of which could be integrated to generate new knowledge. For school-age children, we embedded the facts in a task in which “players” used cards to navigate around a
game board. On each space of the board, children learned a true but previously unknown fact. After the game, they answered questions; some of the answers could be self-derived through integration of information acquired in the course of the game. A desirable feature of the task, relative to the story-passage paradigm used with 4- and 6-year-olds, is that it did not entail time-consuming reading of full passages of text, permitting administration of a larger number of test trials. Whereas in the story passage paradigm, the maximum number of trials tested has been four, in the new paradigm developed for the present research, we tested 14 to 20 trials (see below).

In Experiment 2, we used the newly-developed paradigm to test self-generation of new knowledge under conditions children routinely experience in educational settings—children were challenged to integrate information they read on their own. Specifically, an experimenter read the facts for half of the trials (experimenter-read condition, as in Experiment 1), and the children read the other half themselves (self-read condition). Within this age range, children take on increasing responsibility for their own learning, as they learn to read and acquire information from written material. Over the first years of formal schooling, they progress from identifying words to comprehending texts (e.g., Perfetti, 1999; Perfetti, Landi, & Oakhill, 2005). Yet in this period, reading can be very demanding of cognitive resources. As such, we hypothesized that performance in the self-read condition would be lower than in the experimenter-read condition, especially for children with less well developed reading comprehension skills. To evaluate this potential relation, we included a standardized measure of reading comprehension. We expected an association between reading comprehension and self-generation in the self-read condition. In light of links between reading and listening comprehension in the target age range (e.g., Nation & Snowling, 1997; Spooner, Baddeley, & Gathercole, 2004), we also predicted a correlation between reading comprehension and self-generation in the experimenter-read condition.

To anticipate the results of Experiment 2, we found that children successfully self-generated new semantic knowledge both when the true but previously unknown stem facts were read to them and when they read the facts themselves. Successful performance in the self-read condition created the opportunity to use reading for insight into the cognitive processes involved in productive extension of knowledge through integration. In research with both children (e.g., Nation, Marshall, & Altmann, 2003; Van Der Schoot, 2008) and adults (e.g., Just & Carpenter, 1980; Rayner, Chace, Slattery, & Ashby, 2006), the ebb and flow of reading has been used to gain insight into the cognitive processes involved as readers comprehend texts. Briefly, reading times tend to slow under conditions of high comprehension demand. Readers who are more successful making the inferences required to comprehend texts show greater reductions in reading time, relative to readers who are less successful (see Oakhill & Yuill, 1996, for a review). To capitalize on this potential, in Experiment 3, we adapted the task for use on an eye tracker, thus permitting examination of the number and duration of fixations that children made as they read stem facts and facts derived through integration. We used the variation in fixations to inform the cognitive processes involved as children read and are called upon to remember new knowledge derived through integration. We predicted that fixations during reading of the integration facts would be related to subsequent recall of them. We also made the more tentative prediction that fixations during reading of the second member of a pair of related stem facts also would be
related to subsequent recall of integration facts, based on the expectation that inferential processes would be triggered by the second member of a stem-fact pair.

In all three experiments, children were 7 to 10 years of age. We selected this range for three reasons. First, the range was sufficiently narrow to permit derivation of facts that were both accurate and likely unknown to the children prior to the study. In the metropolitan area in which the research was conducted, children transition to middle school in Grade 6. Given substantial changes in curriculum between the elementary- and middle-school years, we elected Grade 5 (age 10 years), as the upper bound for the age range. Second, for Experiments 2 and 3, we wanted to be reasonably assured that the children would have sufficient reading proficiency to take part in the reading-based paradigm. This placed the lower limit at age 7 years. Third, the age range was sufficiently wide to capture variability in reading comprehension and thus adequately address the research questions. In this regard it is important to note that the primary variable of interest in Experiments 2 and 3 was reading comprehension, rather than age, per se (which is only a proxy variable in any case). Thus we were not primarily concerned with age-related differences within the age range but rather, with potential relations with reading comprehension.

**Experiment 1**

**Method**

**Participants**—Participants were 22 children aged 7 and 10 years (10 females, 12 males; \( M = 8.6 \) years; range: 7 years, 2 months to 10 years, 4 months). The participants were recruited from a pool of parents who expressed interest in having their children take part in research studies. Nine percent of the children were African American, 86% were Caucasian, and 5% were of mixed race, based on parental report. Nine percent of the children were identified as Hispanic. No specific data on socioeconomic status were collected though the pool from which the participants were recruited is primarily middle- to upper-middle socioeconomic range. All of the participants were native English speakers, based on parental report. In this and the subsequent experiments, parents provided written informed consent and children provided verbal assent prior to initiation of the study session. At the end of the session, children were given an age-appropriate toy and parents were given a gift card to a local merchant. The institutional review board reviewed and approved all study protocols and procedures in this and the subsequent studies in this report.

**Materials**—Materials were a board game in the context of which children learned true but previously unknown facts, and open-ended and forced-choice questions to test for self-generation of new semantic knowledge through integration of the facts. Children also participated in a brief non-game-related activity (“find the difference” or word-search worksheets). In the game, “Move Ahead” cards indicated the number of spaces a child was to move on the board, and “Fact” cards featured a fact 4 to 10 words in length. There were 20 pairs of Fact cards, for a total of 40 cards. A pair consisted of two facts (stem facts) that when integrated with one another, supported self-generation of new knowledge not given in either of the facts (integration facts). An example stem-fact pair is “A wombat is a marsupial” and “Marsupials keep their babies in a pouch.” When integrated with one
another, the pair of stem facts supported self-generation of the novel integration fact that “Wombats keep their babies in a pouch.” All of the stem and integration facts were accurate and determined novel for 7- to 10-year-old children based on pilot testing.

**Procedure and Scoring**—Children were tested individually in a single session in a laboratory setting by one of two experimenters. Each experimenter tested half of the children. To ensure protocol fidelity, the experimenters followed an explicit written study protocol. They regularly reviewed one another’s sessions and discussed any departures from protocol. The session consisted of two phases.

**Exposure to stem facts:** Children were introduced to and received instructions on how to play the board game. To play, children selected a Move Ahead card indicating how far to advance on the board. Upon arrival at the designated location, the researcher selected a Fact card and read it aloud to the child. In actuality, both the number of spaces advanced and the specific Fact cards read on each trial were predetermined to ensure the appropriate number and spacing of stem-fact presentations. To ensure the child’s attention to the facts as they were read, after each fact, children were asked a question requiring a judgment about the fact. For instance, for the fact “A wombat is a marsupial,” the question was “Is this fact about animals or computers?”

In the course of the game, children were exposed to 30 of the possible 40 stem facts across two conditions: 2-stem and 1-stem. As depicted schematically in Table 1, Panel a, in the 2-stem condition, children were exposed to both members of a pair of related stem facts which, when integrated with one another, supported self-generation of a novel integration fact (20 stem facts total; 10 pairs). In the 1-stem condition, children were exposed to only one of the members of a given stem-fact pair (10 stem facts; 0 pairs); they were not exposed to the other 10 stem facts. Because in the 1-stem condition, children were exposed to only half of the information necessary for self-generation of new knowledge, they were not expected to produce the integration facts for these 10 stimuli. Across participants, the stem fact pairs were counterbalanced such that each was presented equally often as a full pair (2-stem condition) and as only half a pair (1-stem condition). The individual stem facts are presented pseudorandomly such that no two members of a pair of facts were presented in succession. The lag (number of Fact cards presented between related stem-fact pairs) ranged from 2 to 21, with an average spacing of 9 cards between members of a related pair. Although the lag between members of a stem-fact pair could influence performance, lag was not a variable of interest in the present research and thus was not analyzed. Importantly, none of the integration facts was presented and at no time were children informed that any of the facts was related to any others. Indeed, the trials were separated into individual episodes by virtue of the facts that (a) each of the stem facts was experienced on a separate trial of the game, (b) children were asked and answered questions about each fact before selecting a Move Ahead card to begin the next trial, and (c) members of the pairs of stem facts were separated by other stem facts (and related questions) from different conceptual domains.

Following the game, children completed an unrelated activity for roughly 5 minutes (“find the difference” or word-search worksheets). The activity served as a buffer between
presentation of the stem facts and test for self-generation of new knowledge through integration.

**Test for self-generation:** After the 5 minute buffer activity, children were tested for self-generation of integration facts. For each pair of stem facts in both the 2- and 1-stem conditions, they were asked an open-ended question that could be answered by self-generating new knowledge through integration of the stem facts. For instance, the stem facts “A wombat is a marsupial” and “Marsupials keep their babies in a pouch,” could be integrated to answer the experimenter’s question “Where do wombats keep their babies?” The experimenter recorded children’s verbal responses to the questions. To ensure some success during open-ended testing, we interspersed 5 questions to which children were likely to know the answers, such as “What is the hottest season of the year?” Answers to these questions were not scored. The questions were presented in one of 4 pre-determined random orders; each order was used approximately equally often across participants. Across the two conditions, 20 questions were posed (10 2-stem and 10 1-stem); the maximum possible number of correct responses per condition was 10.

Open-ended questions provide the strongest test of self-generation of new knowledge through integration. As an additional measure, for each question not correctly answered in open-ended format, we administered a forced-choice question; all open-ended questions were asked before any forced-choice questions were asked. Each forced-choice question had 3 substantive options, 1 of which was correct; children were not given a “don’t know” response option. For example, for the question, “Where do wombats keep their babies?” the answer choices were: (a) on their backs, (b) in their pouches, (c) in their arms. To ensure that children had some success during forced-choice questioning, we interspersed 5 questions to which children were likely to know the answers. Answers to these questions were not scored. The forced-choice questions were presented in one of 4 pre-determined random orders; each order was used approximately equally often across participants. The number of forced-choice questions answered correctly was added to the number of open-ended questions answered correctly, for a total score, the maximum of which was 10 in each condition (2-stem and 1-stem).

**Results**

Descriptive statistics on children’s performance are provided in Table 2, Panel a. For both open-ended and total performance (number correct in open-ended plus forced-choice question formats), children’s performance was higher in the 2-stem than in the 1-stem condition: ts(21) = 4.91 and 4.40, ps < .0001, respectively. Thus children provided more integration facts in the 2-stem condition than the 1-stem condition, presumably because they self-generated them in the 2-stem condition, an approach not available in the 1-stem condition. Considering forced-choice performance alone, in the 2-stem condition, children’s performance differed from that which would be expected by chance: t(21) = 5.67, p < .0001. In contrast, in the 1-stem condition, children’s forced-choice performance did not differ from chance: t(21) = 1.45, p > .10.
In light of the wide range in age of the children in the study (7;2–10;4), we calculated Pearson Product-Moment correlations to test relations between age and performance. As noted above, the primary reason for including the wide age range was not to evaluate age-related differences, per se, but to provide variability in the primary individual difference variable in Experiments 2 and 3, namely reading comprehension. Yet it was important to determine whether success in open-ended testing was confined to the older children in the sample. In open-ended testing, in the 2-stem condition, age accounted for a non-significant 11% of the variance in performance, \( r(20) = .33, \text{ ns.} \) Similarly, in the 1-stem condition, age was not correlated with production of the integration facts, \( r(20) = .01. \) In the 2-stem condition, age accounted for 31% of the variance in total performance (number correct in open-ended plus forced-choice format), \( r(20) = .56, p < .01, \) possibly reflecting the greater prowess of older children in detecting the relation between the forced-choice alternatives and the stem facts experienced earlier in the session. Age was not correlated with total performance in the 1-stem condition, \( r(20) = -.02. \)

**Discussion**

The paradigm developed for the present experiment provided a valid test of self-generation of new knowledge through integration by 7- to 10-year-old children. In open-ended testing, when both members of a pair of facts had been presented (2-stem condition), children self-generated 48 (22%) of the 220 possible integration facts (22 children \( \times \) 10 2-stem trials per child). In contrast, when only one member of the pair of facts was presented (1-stem condition), children produced only 6 (3%) of the 220 possible integration facts. Success in open-ended testing was not confined to the oldest children and conversely, even the oldest children did not enjoy success in the 1-stem control condition. Additional evidence that integration of stem facts was necessary to support production of the integration facts came from performance in forced-choice testing. When presented with 3 alternatives, in the 2-stem condition, children selected the correct response more frequently than expected by chance. In the 1-stem condition, performance did not differ from chance. Together, performance in open-ended and forced-choice testing provided strong evidence that exposure to both members of a pair of related facts was necessary for production of the integration facts, presumably through self-generation of new knowledge.

In addition to permitting a valid test for self-generation of new knowledge through integration, the new paradigm developed in this research also has the potential benefit of permitting evaluation of the cognitive process when children learn stem facts as they read. In the present experiment, all of the stem facts were read by an experimenter. The paradigm is readily adaptable to permit children to read the stem facts for themselves. Though there are marked individual differences in reading ability among school-age children (e.g., Bast & Reitsma, 1998; Juel, 1988; Whitehurst & Lonigan, 2003), by the second to third grade, children can be expected to read for meaning in a text (e.g., Clay, 1991). As they progress through the school years, expectations for independent reading and extraction of information from texts increase (e.g., Collins & Smith, 1980; Nagy, Anderson, & Herman, 1987; Whitehurst & Lonigan, 2003). Given the importance of reading as a source of information, it is desirable to determine how self-generation of new knowledge through integration is affected when children are required to read stem-fact material themselves. We addressed this
question in Experiment 2. We also assessed children’s levels of reading comprehension, thus permitting test of how reading proficiency relates to self-generation of new knowledge through integration. We hypothesized that performance in the self-read condition would be lower than in the experimenter-read condition, especially for children with less well developed reading comprehension skills.

Experiment 2

Method

Participants—Participants were 20 children aged 7 and 10 years (12 females, 8 males; M = 8.8 years; range: 7 years, 2 months to 10 years, 8 months). Children were recruited from two sources: the participant pool from which the children in Experiment 1 were drawn (N = 8) and an arts afterschool program (N = 12). The sources were comparable in racial and ethnic composition and in family socio-economic status. The children did not differ in their reading comprehension scores (see below). Three additional children were tested but excluded from analyses due to procedural error (N = 1) and inability to participate in the reading manipulation, due to failure on the basal level of the reading assessment (N = 2). Based on parental report, 10% of the children were African American, 85% were Caucasian, and 5% were of mixed race. Ten percent of the children were identified as Hispanic. All participants were native English speakers, based on parental report. None of the children had participated in Experiment 1. At the end of the session, children were given a toy or stickers. The parents of children from the participant pool were given a gift card to a local merchant; the after-school program was given a gift card donation.

Materials—The experimental stimuli were the same as used in Experiment 1. We also used the Woodcock Johnson III Test of Cognitive Abilities (WJ-III) Passage Comprehension test (Woodcock, McGrew, & Mather, 2001) to assess children’s reading abilities.

Procedure—Children were tested individually in a single session in a laboratory setting or a quiet room of their afterschool program. Testing was conducted by one of four experimenters (including the second author), each of whom tested 3–7 children. To ensure protocol fidelity, the experimenters followed a written protocol. They regularly reviewed and observed one another’s sessions and discussed any departures from protocol. The session consisted of three phases.

Exposure to stem facts: Children were exposed to the stem facts in the same paradigm used in Experiment 1, adapted in two ways. First, all stem facts were presented in the 2-stem condition (i.e., children were exposed to both members of all stem-fact pairs). The 1-stem condition was omitted because Experiment 1 made clear that exposure to both members of the stem-fact pair was necessary to support production of integration facts, presumably through self-generation of them. Second, half of the stem fact pairs were read by the experimenter (experimenter-read condition) and half were read out loud by the children themselves (self-read condition). Although a total of 20 stem-fact pairs was available, to maintain consistency in the number of facts read across Experiments 1 and 2, for a given child, 15 of the 20 possible pairs were used, for a total of 30 facts; each pair was used
equally often across participants. Seven pairs (14 stem facts) were used in the experimenter-read condition (both members of the pair were read by the experimenter) and seven pairs (14 stem facts) were used in the self-read condition (both members of the pair were read by the child her/himself). The remaining two related stem facts were used in a split condition in which the experimenter read one stem fact and the child read the other. This split pair was not used in analyses. Across participants, the stem fact pairs were counterbalanced such that each was presented equally often in the experimenter- and self-read conditions. The individual stem facts were presented pseudorandomly such that no two members of a pair of facts were presented in succession. The number of stem facts presented between members of stem-fact pairs (lag) ranged from 2 to 23, with an average spacing of 11. As in Experiment 1, although lag could influence performance, lag was not a variable of interest in the present research and thus was not analyzed. Also as in Experiment 1, to ensure the child’s attention to the facts, after each fact, they were asked a question requiring a judgment about the fact (e.g., “Is this fact about animals or computers?”). At no time were children informed that any of the facts was related to any others; none of the integration facts was presented.

As in Experiment 1, following the game, children completed an unrelated buffer activity for approximately 5 minutes.

**Test for self-generation:** After the buffer activity, children were tested for self-generation following the procedures in Experiment 1. The experimenter read all questions to the child. All open-ended questions were posed prior to any forced-choice questions. Forced-choice questions were posed only when children responded incorrectly to the open-ended questions. The maximum possible number of correct responses in each condition (experimenter- and self-read) was 7. The number of correct forced-choice responses was added to the number of correct open-ended responses, for a total score, the maximum of which was 7 in each condition.

**Test for reading comprehension:** Following open-ended and forced-choice testing, children’s reading comprehension was tested using the WJ-III Passage Comprehension test. In the initial portion of the assessment, children read portions of sentences and complete them with an appropriate word; pictures accompany the sentences. As the assessment progresses, the items increase in difficulty and are not accompanied by pictures. Testing is terminated after 6 consecutive errors. Children received a raw score between 0–42.

**Results**

**Preliminary Analyses**—Preliminary analyses revealed no significant differences in performance on either the experimental tasks or the reading comprehension test for the children drawn from the participant pool and the children drawn from the afterschool program. For all subsequent analyses, the two groups of participants were treated as one sample.

**Self-generation Performance**—Descriptive statistics on children’s performance are provided in Table 2, Panel b. Across conditions, the children self-generated the novel integration facts in open-ended testing on 39% of the trials; they either self-generated or
responded correctly to a forced-choice question regarding the novel integration facts on 82% of the trials. Contrary to the hypothesis that performance in the self-read condition would be lower than in the experimenter-read condition, paired t-tests revealed no statistically significant differences between children’s performance in the two conditions, in either open-ended testing or total performance, ts(19) = −1.82 and 0.46, ps > .05, respectively. Performance in the two conditions was significantly correlated in open-ended testing, r(18) = .49, p = .03. When total performance was considered, the correlation between scores in the two conditions fell below the level of statistical significance, r(18) = .43, p = .062.

**Reading Comprehension and Relations with Self-generation Performance—**

Reading comprehension was variable among the children, with scores ranging from 11–35 reflecting equivalency grade levels 1.7–15.6 (college-level reading; M = 23.25, Estimated Grade Level Equivalency = 5). Because reading comprehension scores were correlated with age, r(18) = .58, p = .007, analyses of relations with self-generation performance were conducted with age controlled through partial correlation. The partial correlation coefficients between reading comprehension scores and self-generation performance are provided in Table 3.

As reflected in Table 3, consistent with our hypotheses, reading comprehension was correlated with self-generation through integration as measured in the open-ended format. Across conditions, it accounted for 58% of the variance in open-ended performance. Reading comprehension scores also were correlated with performance in both the experimenter-read and self-read conditions separately, accounting for 52% and 36% of variance, respectively. Although the association between self-generation performance and reading comprehension scores was nominally higher in the experimenter-read than the self-read condition, the correlations did not differ statistically, Williams T(17) = .76, p = .46. Reading comprehension scores were not significantly related to children’s total performance (open-ended plus forced-choice), either across conditions or for either condition separately. The lack of association was not due to the control for age: age was not significantly correlated with total performance either overall or for either the experimenter-read or child-read conditions separately (see Table 3).

**Discussion**

The major purpose of Experiment 2 was to determine how self-generation of new knowledge through integration was affected when children were required to read stem-fact material themselves. Performance did not differ between the experimenter-read and self-read conditions, thus demonstrating the robustness of the process of self-generation of new semantic knowledge through integration of true but previously unknown facts. As expected, reading comprehension scores were both variable and related to children’s age (within the 7–10 year range). With age controlled, reading comprehension accounted for just over half of the variance in open-ended self-generation performance, across conditions. It explained significant variance in both the experimenter- and self-read conditions. These findings indicate an interplay of reading comprehension and self-generation of new knowledge through integration in school-age children, even when children are not responsible for reading the material themselves.
The present experiment allowed us to test how levels of self-generation of new knowledge through integration were impacted by the requirement that children read the stimulus materials for themselves. Because changes in the rate of reading are indicative of fluctuations in cognitive effort (e.g., Just & Carpenter, 1980; Rayner, 1998), reading also holds the potential for examining the process of integration. To capitalize on this opportunity, in Experiment 3, we transitioned the paradigm from a game format in which children read the stem facts from “Fact” cards, to a monitor that permitted tracking of the number and duration of fixations that the children made as they read. Because longer fixations are indicative of devotion of more cognitive effort (e.g., Just & Carpenter, 1980; Rayner, 1998), we expected that longer fixations would be predictive of successful performance. We also changed the outcome measure. Instead of the number of trials on which children self-generated a novel integration fact, we measured their recall of integration facts they read for themselves. This change permitted examination of relations between processing of the stem facts and processing of facts formed through integration of the stem facts. As in Experiment 2, we also assessed children’s reading comprehension, thus permitting test of how reading proficiency relates to the processing of true but previously unknown stem-facts and integration facts, and subsequent recall of integration facts.

**Experiment 3**

**Method**

**Participants**—Participants were 19 children aged 7 to 10 years (7 females, 12 males; $M = 8.9$ years; range: 7 years, 11 months to 10 years, 1 months). The children were recruited from the same source as in Experiment 1. Four additional participants were tested but excluded from analyses due to procedural error ($N = 2$) and low quality eye movement data ($N = 2$; see below). Based on parental report, 58% of the children were African American, 37% were Caucasian, and 5% reported as other race. No children were identified as Hispanic. All of the participants were native English speakers, based on parental report. None of the children had participated in either Experiment 1 or Experiment 2. At the end of the session, children were given an age-appropriate toy and the parents were given a gift card to a local merchant.

**Materials**—As in Experiment 2, we used the WJ-III Passage Comprehension test to assess children’s reading abilities. The stimuli were the same as in Experiments 1 and 2. For purposes of the present research, we developed “irrelevant” stem facts that were used in a control condition. As reflected in Table 1, Panel b, in the irrelevant control condition, for a given pair of stem facts, the second member of the pair was replaced by a fact that featured the same target concept, but which could not be integrated with the first stem fact to generate the integration fact. For example, for the stem-fact pair “A wombat is a marsupial” and “Marsupials keep their babies in a pouch,” the irrelevant second stem fact was “A joey sleeps in a cozy pouch.” Whereas the first two stem facts could be integrated to generate the integration fact that “Wombats keep their babies in a pouch,” the pair “A wombat is a marsupial” and “A joey sleeps in a cozy pouch,” could not be integrated to generate the integration fact. Because, across participants, each stem fact was used as the first and the second stem fact (see counterbalancing information, below), irrelevant facts were generated...
for both stem facts in a stem-fact pair (e.g., the irrelevant fact for “A wombat is a marsupial” was “A wombat eats at night”). Also, whereas in prior research, probes for the integration facts were presented in the form of questions (e.g., “Where do wombats keep their babies?”), for purposes of the present experiment, the integration-fact questions were transformed to declarative sentences, such as “Wombats keep their babies in a pouch.”

**Procedure**—Children were tested individually in a laboratory setting. Testing was conducted by one female experimenter (the second author). The session consisted of three phases.

---

**Exposure to stem and integration facts:** Children were exposed to the stem and integration facts on the monitor of a Tobii T120 eye tracker. This allowed for recording of both the number and the duration of visual fixations children made as they read the stimulus materials. As in Experiment 2, all of the stem facts were presented in pairs (i.e., on all trials, children were exposed to two stem facts). On half of the trials (N = 10), the two stem facts could be combined to form an integration fact (relevant condition), and on the other half of the trials, the two stem facts could not be combined to form the integration fact (irrelevant condition). Unlike prior experiments, the children also were presented with the integration facts for each stem-fact pair. The integration facts were presented as declarative sentences, and thus were indistinguishable from the stem facts (see above). The three facts for a trial (Stem 1, Stem 2, Integration fact) were presented as a triad, such that for each triad, the first and second stem facts were presented one after another, followed by the integration fact. The children read all of the facts themselves, out loud. Once the child had finished reading each fact, the experimenter advanced the presentation to a question about the fact (e.g., “Is this fact about animals or computers?”). Across participants, the relevant and irrelevant stem-fact pairs were counterbalanced such that each pair was presented equally often in each condition. The relevant and irrelevant fact triads were interspersed in a predetermined random order. Also, across participants, within a fact pair, the facts presented first and second were counterbalanced. At no time were children informed that some of the facts were related to one another or that others were not. As in Experiments 1 and 2, following exposure to the facts, children completed a buffer activity for roughly 5 minutes.

---

**Test for recall of the integration facts:** After the buffer activity, children were tested for recall of the integration facts. For each fact, the experimenter posed an open-ended question that children could answer by recalling the integration facts they had read. For example, the fact “Wombats keep their babies in a pouch,” was rephrased into the question “Where do wombats keep their babies?” Note that the question was the form of testing in Experiments 1 and 2. The maximum possible number of correct responses in each condition (relevant and irrelevant) was 10. Unlike in prior experiments, we did not administer forced-choice questions. The decision to omit forced-choice testing was made to reduce participant burden, which was already high, in that children were required to read 60 facts (10 pairs of related facts, 10 pairs of unrelated facts, 20 integration facts) and answer 20 fact recall questions. As well, because the children already had read all of the integration facts, we expected ceiling performance in forced-choice testing.
Test for reading comprehension: Children’s reading comprehension was tested using the WJ-III Passage Comprehension test, administered as in Experiment 2.

Eye Tracking Data Reduction—Data with a sampling rate lower than 40% were considered low quality and were not analyzed. As noted above, this criterion resulted in exclusion of data from two participants. For participants with data of sufficient quality, we calculated the mean number of fixations that children made to each of the stem facts and the integration facts, as well as the mean duration of fixations. We created separate mean scores for the number of fixations and duration of fixations for the relevant and irrelevant conditions. Within a condition, we created mean scores for the number and duration of fixations for trials on which the children subsequently recalled the integration facts and the trials on which they failed to recall the integration facts.

Results

Recall of Integration Facts—The percentage of integration facts recalled overall, and for the relevant and irrelevant conditions separately, is provided in Table 2, Panel c. Children recalled an average of 6.37 (SD = 1.98) and 6.47 (SD = 2.17) of the 10 integration facts in the relevant and irrelevant conditions, respectively. Performance in the two conditions did not differ statistically, *t*(36) = −0.26, *p* = .79.

Reading Comprehension—Reading comprehension scores ranged from 18–33, reflecting equivalency grade levels 3.3–13.0 (college-level reading; *M* = 24.95, Estimated Grade Level Equivalency = 6.6). Unlike in Experiment 2, the correlation between age and reading comprehension scores was not significant, *r*(17) = .18, *p* = .46, likely due to the more restricted range relative to Experiment 2. Because the variability was not significant, age was not controlled in subsequent analyses.

Number and Duration of Visual Fixations while Reading—Descriptive statistics for the mean number and duration of fixations are provided in Table 4. To test for differences in the mean number and duration of fixations in each condition, for each fact type, as a function of subsequent recall, we conducted separate 2 (condition: relevant, irrelevant) × 3 (fact type: stem 1, stem 2, integration) × 2 (subsequent recall: recalled, not recalled) multivariate analyses of variance (MANOVAs) for each dependent variable (mean number of fixations and mean duration of fixations). For the mean number of fixations, there was a significant main effect for fact type, *F*(2, 207) = 3.72, *p* < 0.01, η² = 0.05. Post-hoc pairwise comparisons with Tukey HSD test revealed that overall, children made more fixations while reading the first stem facts (*M* = 12.13, *SD* = 4.22) than while reading the integration facts (*M* = 9.93, *SD* = 3.07). There were no other significant differences in the mean number of fixations. For the mean number of fixations, there were no other main effects and no interactions. The analysis for mean duration of fixations yielded no significant main effects and no interactions.

Relations among the Variables

Number of integration facts recalled and reading comprehension: The number of integration facts that children recalled was correlated with their scores on the WJ-III Passage
Comprehension test. Across conditions (relevant, irrelevant) the correlation was, \( r(17) = .50, p < .05 \). When the conditions were examined separately, scores on the reading comprehension test were found to correlate with recall of integration facts in the irrelevant condition, \( r(17) = .62, p < .05 \), but not in the relevant condition, \( r(17) = .30, \text{ns} \).

**Number of integration facts recalled and mean number and mean duration of fixations:** The correlations between the number of integration facts recalled and the mean number of fixations and mean duration of fixations are provided in Table 5, Panel a. The number of integration facts recalled was not correlated with the mean number of fixations children made while they read the facts regardless of condition and fact type. As well, the mean duration of fixations on the first stem facts was not correlated with the number of integration facts recalled in either condition. In contrast, in the relevant condition, the number of integration facts recalled was positively correlated with the mean duration of fixations on the second stem facts and on the integration facts. Thus in the condition in which the second stem fact could be integrated with the first stem fact to form the integration fact, the longer children’s average duration of fixation while reading the second stem facts and the integration facts, the higher their levels of subsequent recall of the integration facts. As reflected in Table 5, Panel a, the same relations did not hold in the irrelevant condition. Thus in the condition in which the two stem facts did not form the integration fact, mean duration of fixation was not correlated with subsequent recall.

The importance to subsequent recall of the average duration of fixation on the second stem facts in particular comes from separate analyses of the trials on which the children subsequently recalled the integration facts and trials on which they did not. In the relevant condition only, average fixation durations on the second stem facts were differentially related to subsequent recall. The correlation was statistically significant for trials on which the children subsequently recalled the integration facts, \( r(17) = .51, p < .05 \), but not for trials on which they did not subsequently recall the integration facts, \( r(17) = .27, \text{ns} \). Thus on trials on which children had longer mean fixation durations as they read the second stem facts, they tended to recall the integration facts. On trials on which children did not have longer mean fixation durations as they read the second stem facts, they tended to fail to recall the integration facts. In the relevant condition only, average fixation durations on the integration facts also were positively related to subsequent recall. However, the relation held both for the subset of facts the children subsequently recalled and for the subset of facts they did not recall, \( rs(17) = .50, ps < .05 \). Thus longer average fixation durations on the integration facts were not differentially predictive of whether the facts would be recalled. In effect, cognitive effort while reading the second stem facts was rewarded with higher levels of subsequent recall of the integration facts, whereas effort while reading the integration facts themselves did not as reliably lead to later successful recall. Separate analyses by subsequent recall revealed no relations in the irrelevant condition.

**Reading comprehension and mean number and mean duration of fixations:** The correlations between reading comprehension scores and the number and duration of fixations are provided in Table 5, Panel b. Neither measure correlated with children’s reading comprehension.
Discussion

In the present experiment, levels of recall of novel integration facts did not differ by condition. Nor did either the mean number or the mean duration of fixations differ by condition. Thus differences in relations between processing of the stimuli and subsequent recall cannot be accounted for by differential attention to the stimuli or differential success as a function of condition. Yet against this backdrop of similarities, we observed differences in cognitive processing, as reflected in relations between the mean duration of fixations and recall. Specifically, in the relevant condition only, children who had longer average fixation durations as they read the second stem facts in a triad had higher levels of subsequent recall of the integration facts. The relation obtained only for trials on which the children subsequently recalled the integration facts. On trials on which recall of the integration facts failed, there was not a relation between processing of the second stem fact in a triad and recall. We interpret this pattern as suggestive evidence that in the relevant condition, higher average fixation durations while reading the second stem fact in a triad was indicative of active processing of the relation between the stem fact the children were currently reading and the one they had just read. Conjoint processing of the pairs of related stem facts in turn facilitated processing—and subsequent recall—of the integration facts.

Children who had longer average fixations of the integration facts also had higher levels of recall of them. Yet the relation was observed both on trials on which they recalled the integration facts and on trials on which they did not recall the facts. Thus greater cognitive effort while reading the integration facts did not necessarily result in successful performance (see Kafkas & Montaldi, 2011, for similar results in a memory task). This pattern was not apparent in the irrelevant condition, in which none of the relations between mean fixation duration and recall obtained. Overall, the results suggest that when productive extension of knowledge through integration is possible (as in the relevant condition), greater cognitive effort while reading the second stem of a pair of related facts in particular, is rewarded with higher levels of recall.

General Discussion

Semantic memory functions as a storehouse for knowledge about the world. Many entries in the semantic memory store are made through direct tuition or experience. Others are the result of productive processes such as self-generation of new semantic knowledge through integration of separate yet related episodes of new learning (Bauer & San Souci, 2010). Prior research has demonstrated this productive process in children 4 and 6 years of age who, based on stories read to them by an experimenter, productively extend semantic knowledge through self-generation. As well, college-age adults have been found to productively extend semantic knowledge through self-generation based on facts they read themselves during a knowledge-difficulty judgment task (Bauer & Jackson, 2015). One major purpose of the present research was to extend this line of research to the population in between preschool and the college years—to children in elementary school. School-age children are prime targets for investigation of this productive capacity, given the importance of acquisition of semantic knowledge during this age period. A second purpose of the work was to investigate self-generation of new knowledge through integration of information children read themselves (Experiments 2 and 3). The demand to acquire information through reading
increases throughout the school years, making the consequences of reading for self-generation of new knowledge an important topic for examination. The third purpose was to use variation in children’s reading of stem and integration facts as a “window” on the cognitive processes involved in productive extension of semantic knowledge (Experiment 3).

There are several major findings from the present research. The first major finding was that the new board-game paradigm that we developed for school-age children proved a valid test of self-generation of new semantic knowledge through integration. In Experiment 1, 7- to 10-year-old children successfully self-generated new semantic knowledge through integration of separate, yet related pairs of novel facts read by an experimenter. In contrast, when exposed to only one member of a pair of related facts, they neither produced the integration fact in open-ended testing, nor evidenced above-chance performance during forced-choice testing. Thus integration of separate yet related facts was necessary for production of the novel facts.

In a pattern that violates the expectation of improved performance with age, the level of self-generation performance of the 7- to 10-year-olds in Experiment 1 of the present research was nominally lower than that observed among 6-year-old children in prior research; the level of total performance (including integration facts correctly identified in a forced-choice format) also was lower (Bauer & San Souci, 2010; Varga & Bauer, 2013). The level of performance across open-ended and forced-choice testing was roughly on par with that observed in forced-choice testing of adults in Bauer and Jackson (2015). One potential source of difference in levels of performance may be the format of presentation of the stem facts. In prior research with 4- and 6-year-olds, stem facts were presented in the context of connected passages of text. In contrast, in the present research, as well as in Bauer and Jackson, participants were exposed to long lists of seemingly unrelated facts. Differences in coherence of the information to which participants were exposed are a possible source of variance in performance. Another possible source of variance in performance is the number of stem facts to which participants were exposed. In research with younger children, the maximum number of story passages read is four (typically with two stories featuring information that could be integrated). In the research with school-age children and adults, the numbers of stem facts to which participants were exposed was 30 and 120, respectively. A larger number of stimuli at the time of encoding may have contributed to interference, to cognitive overload, or both. The impacts of different formats of presentation of stem facts and the number of facts to which participants are exposed should be examined systematically for their individual and combined impact on self-generation performance.

The second major finding from the present research was that the 7- to 10-year-old children tested in Experiment 2 were equally facile at productive self-generation of new semantic knowledge through integration of pairs of related facts read to them by an experimenter and pairs of related facts they read themselves. Indeed, overall performance in Experiment 2 (39% and 82% for open-ended and total performance, respectively) was nominally better than performance in Experiment 1 (22% and 65% for open-ended and total performance, respectively). The pattern is unlikely to be attributable to differences in the samples between experiments. The children in the two experiments had the same mean age as well as the same range of ages. The samples also were drawn from comparable populations. One likely
possible explanation for the difference in performance is that the “density” of to-be-integrated information differed between experiments. In both experiments, children were exposed to 30 stem facts. In Experiment 1, 20 of the stem facts formed 10 stem-fact pairs; 10 stem facts were unrelated. In Experiment 2, all of the stem facts formed pairs; none of the stem facts was unrelated. Though the precise means by which a greater density of related information might facilitate performance is unknown, it is possible that the high density of relations among the facts established a cognitive set that primed children for relational processing. This is an intriguing possibility that could be tested in future research.

The third major finding from the present research was of positive relations between self-generation of new knowledge through integration (Experiment 2) and recall of facts derived through integration (Experiment 3) and independent measures of reading comprehension, namely, the Woodcock Johnson III Test of Cognitive Abilities Passage Comprehension test. In Experiment 2, the correlation between reading comprehension and self-generation through integration was observed in the open-ended testing format, for performance across conditions as well as for performance in each of the experimenter-read and self-read conditions separately. With variance in age controlled, reading comprehension accounted for 58% of the variance in open-ended format testing performance across conditions, and 52% and 36% of the variance in performance in the experimenter- and self-read conditions, respectively. Reading comprehension scores were not correlated with total performance (including forced-choice testing), likely due to restricted range of the variable (children’s total performance averaged 82% correct).

An interesting finding in Experiment 2 was that reading comprehension scores were correlated not only with performance in the self-read condition, but also with performance in the experimenter-read condition. Given that within the age range tested in the present research, reading and listening comprehension are positively related with one another (Nation & Snowling, 1997; Spooner et al., 2004), the finding was not unexpected. On the basis of the present research, it is not possible to interpret the findings unequivocally. One possible interpretation is that the ability to self-generate new semantic knowledge through integration of separate learning episodes facilitates both listening and reading comprehension. An equally logical alternative is that both listening and reading comprehension facilitate self-generation of new semantic knowledge through integration. Future research—especially of a longitudinal nature—will help to address this important question.

Reading comprehension scores also were positively related to recall of integration facts in Experiment 3. The correlation between reading comprehension scores and recall of integration facts was observed across the relevant and irrelevant conditions, and for the irrelevant condition alone; the proportions of variance explained were 25% and 38%, respectively. The correlations are understandable in light of the fact that children were required to read the stimulus materials on their own. More proficient reading presumably allowed them to devote more attention to the meaning of the material they were reading, leading to more robust recall. In contrast to overall recall and recall in the irrelevant condition, for the relevant condition alone, the relation between recall of integration facts and reading comprehension scores was not significant. In the relevant condition, rather than
with overall reading comprehension, performance was associated with deployment of cognitive resources while reading, as described next.

The fourth major contribution of the present research was the light it shed on the cognitive processes involved in productive extension of knowledge through integration. In Experiment 3, children’s levels of recall of facts that were formed through integration did not differ from their recall of facts that were not formed in that manner. Nor did the patterns of fixation as the children read the facts differ in the two conditions, either in terms of the number of fixations or their average duration. When integration of the stem facts was not a means to derivation of the integration facts (i.e., in the irrelevant condition), the only variable that predicted children’s subsequent recall of the integration facts was their level of reading comprehension (discussed above). In contrast, when integration of the stem facts was a means to derivation of the integration facts (i.e., in the relevant condition), the mean duration of fixations to both the second stem facts and the integration facts themselves was related to subsequent recall. The total number of integration facts recalled was correlated with the mean duration of fixations as children read the integration facts that they subsequently recalled as well as those that they subsequently did not recall. This indicates that higher effort was not necessarily associated with higher success.

In contrast to the pattern observed for the integration facts, the relation with mean duration of fixations of the second stem fact was only present on trials for which the children subsequently recalled the integration facts. In this case, higher cognitive effort while reading the second stem fact was associated with successful recall. Together, the pattern of results suggests that for children 7 to 10 years of age, the work of production extension of new semantic knowledge begins as soon as the opportunity for integration presents itself (i.e., as children are reading the second, related stem fact). When children capitalize on the opportunity, they tend to have high levels of recall of the new knowledge. This speaks to the psychological significance of the process and its potential as a mechanism for extension of the semantic knowledge base (see Bauer & Varga, 2015, for discussion). Further research will be required to determine the precise nature of the processing that facilitates subsequent recall. Based on the results of the present research, a reasonable hypothesis is that active processing of the relation between the two stem facts—reflected in higher average fixation durations while reading the second stem fact—facilitates processing and subsequent recall of integration facts.

The present research took important steps in informing self-generation of new semantic knowledge through integration of separate learning episodes in school-age children. It also suggests potentially productive lines of future research. In addition to those already mentioned, two especially salient directions for future research are into the causes and consequences of self-generation of new semantic knowledge through integration. In the present research, we examined two potential “causes,” namely, reading comprehension and deployment of cognitive effort during reading. In future research it will be important to examine other potential determinants of self-generation, including such factors as executive function, which may be especially important as children select from among possible competing sources of potentially relevant information. In Bauer and Larkina (in press), this requirement impaired performance among 4- and 6-year-old children, though not among 8-
year-olds. Whether developmental differences in executive function contributed to the pattern is unknown, because no measures of executive function were obtained. We predict that one especially important “consequence” of self-generation of new knowledge through integration is academic achievement. Children who readily extend semantic memory through integration of separate learning episodes should be at an advantage in academic settings. In future research it will be desirable to examine relations between individual differences in self-generation of new semantic knowledge in our experimental task and children’s achievement in a variety of academic domains.

In conclusion, self-generation through integration is increasingly established as one among many productive processes exploited by active learners, including for example, discovery (e.g., Sobel & Sommerville, 2010), problem solving (e.g., Fireman et al., 2003), textual inference (e.g., Paris & Upton, 1976), analogy (e.g., Goswami, 1992), and induction (e.g., Gelman & Markman, 1986). Yet relative to these other productive processes, it has the unique characteristic of involving integration of material acquired in separate episodes during which children learn true but previously unknown facts, thus illustrating a mechanism for accumulating knowledge acquired over time, different contexts, and so forth. It also has the feature of going beyond the information directly acquired, thus illustrating a mechanism for generating new semantic content, thereby facilitating rapid accumulation of a knowledge base. As well, from prior research, we know that the newly-derived content becomes incorporated into semantic stores (Bauer & Jackson, 2015), where it endures over time (Varga & Bauer, 2013; Varga et al., in press). Together, these features nominate self-generation of new semantic knowledge through integration of separate episodes of new learning as a major mechanism of cognitive development (Bauer, 2012; Bauer & Varga, 2015). The process is apparent in the preschool years (Bauer et al., 2012, 2015; Bauer & San Souci, 2010; Varga & Bauer, 2013; Varga et al., in press) and in the college years (Bauer & Jackson, 2015). The present research began to fill the “gap” in between, by focusing on school-age children.

The 7- to 10-year-old children in the present research not only extended their knowledge through self-generation based on integration, they did so while simultaneously challenged to read information for themselves. Indeed, reading comprehension scores were related to productive self-generation of semantic memory both when children read for themselves and when an experimenter read to them. Reading comprehension scores also were related to subsequent recall of integration facts. Subsequent recall of newly learned information derived through integration also was predicted by deployment of cognitive resources during reading. These findings motivate future research to identify other factors that may contribute to productive extension of semantic memory through self-generation, as well as the consequences of this productive capacity for academic and other achievements. As discussed in Bauer and Jackson (2015), this productive process may contribute to more than the extension of semantic knowledge. It also may contribute to the decontextualized nature of the contents of semantic memory: in some cases, we may not remember where or when we acquired knowledge because it was never directly experienced—we generated it ourselves.
Acknowledgments

The work was supported by NICHD HD067359, to Patricia J. Bauer, and by Emory College of Arts and Sciences. The authors also thank Elizabeth White, Ayla Nolen, Nicole Hättenschwiler, Kristen Tuggle, and Catherine Womick, for assistance with participant testing, and other members of the Memory at Emory laboratory group for help at various stages of the research. They also thank the Color Wheel after-school program, Decatur, Georgia, and the children and parents who took part in the research.

References


Table 1
Schematic Representation of Conditions in Experiment 1 (Panel a) and Experiment 3 (Panel b)

<table>
<thead>
<tr>
<th>Experimental Conditions</th>
<th>Stem 1</th>
<th>Stem 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel a: Experiment 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-stem condition</td>
<td><em>A wombat is a marsupial</em></td>
<td><em>Marsupials keep their babies in a pouch</em></td>
</tr>
<tr>
<td>1-stem condition</td>
<td><em>A wombat is a marsupial</em></td>
<td>[not presented]</td>
</tr>
<tr>
<td>Panel b: Experiment 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant condition</td>
<td><em>A wombat is a marsupial</em></td>
<td><em>Marsupials keep their babies in a pouch</em></td>
</tr>
<tr>
<td>Irrelevant condition</td>
<td><em>A wombat is a marsupial</em></td>
<td><em>A joey sleeps in a cozy pouch</em></td>
</tr>
</tbody>
</table>

*Dev Psychol.* Author manuscript; available in PMC 2017 July 01.
Table 2
Means (and Standard Deviations) of Percentage Correct for Open-ended and Total (Open-ended plus Forced-Choice Self-generation) Performance for Experiment 1 (Panel a) and Experiment 2 (Panel b), and Percentage Correctly Recalled in Open-ended Testing for Experiment 3 (Panel c)

<table>
<thead>
<tr>
<th>Condition (max possible score)</th>
<th>Open-ended M percentage (SD)</th>
<th>Total (open-ended plus forced-choice self-generation) M percentage (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel a: Experiment 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-stem (10)</td>
<td>22 (15.6)</td>
<td>65 (19.4)</td>
</tr>
<tr>
<td>1-stem (10)</td>
<td>3 (6.5)</td>
<td>40 (16.0)</td>
</tr>
<tr>
<td>Panel b: Experiment 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Across conditions (14)</td>
<td>39 (21.4)</td>
<td>82 (13.0)</td>
</tr>
<tr>
<td>Experimenter-read (7)</td>
<td>33 (20.9)</td>
<td>79 (16.9)</td>
</tr>
<tr>
<td>Self-read (7)</td>
<td>41 (25.8)</td>
<td>83 (12.7)</td>
</tr>
<tr>
<td>Panel c: Experiment 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Across conditions (20)</td>
<td>64 (18.8)</td>
<td>NA</td>
</tr>
<tr>
<td>Relevant (10)</td>
<td>65 (22.0)</td>
<td>NA</td>
</tr>
<tr>
<td>Irrelevant (10)</td>
<td>63 (19.5)</td>
<td>NA</td>
</tr>
</tbody>
</table>
Table 3
Pearson’s Product Moment Correlations (df = 18) for Age and Reading Comprehension with Self-generation Performance in Experiment 2

<table>
<thead>
<tr>
<th>Condition/Scoring Phase</th>
<th>Measure</th>
<th>Age</th>
<th>Reading comprehension with age controlled (partial correlation)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age</td>
<td>.18</td>
<td>.24</td>
</tr>
<tr>
<td>Across conditions</td>
<td>Open-ended</td>
<td>.23</td>
<td>.76***</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Experimenter-read</td>
<td>.16</td>
<td>.72**</td>
</tr>
<tr>
<td></td>
<td>Open-ended</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>.27</td>
<td>.20</td>
</tr>
<tr>
<td></td>
<td>Self-read</td>
<td>.23</td>
<td>.60**</td>
</tr>
<tr>
<td></td>
<td>Open-ended</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>-.01</td>
<td>.42</td>
</tr>
</tbody>
</table>

Note:

* = p ≤ .05;
** = p ≤ .01;
*** = p ≤ .001.
### Table 4

Means (and Standard Deviations) for Number of Visual Fixations and Duration of Visual Fixations by Condition and Subsequent Recall in Experiment 3

<table>
<thead>
<tr>
<th>Dependent Measure/Condition/ Subsequent Recall Status</th>
<th>Stimulus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stem fact 1</td>
</tr>
<tr>
<td>Panel a: Number of fixations</td>
<td></td>
</tr>
<tr>
<td>Relevant     Recalled</td>
<td>12.30 (5.99)</td>
</tr>
<tr>
<td>Relevant     Not-recalled</td>
<td>13.83 (4.60)</td>
</tr>
<tr>
<td>Irrelevant   Recalled</td>
<td>12.93 (5.24)</td>
</tr>
<tr>
<td>Irrelevant   Not-recalled</td>
<td>12.73 (5.03)</td>
</tr>
<tr>
<td>Panel b: Duration of fixations (ms.)</td>
<td></td>
</tr>
<tr>
<td>Relevant     Recalled</td>
<td>269.07 (52.71)</td>
</tr>
<tr>
<td>Relevant     Not-recalled</td>
<td>253.60 (38.96)</td>
</tr>
<tr>
<td>Irrelevant   Recalled</td>
<td>274.09 (48.60)</td>
</tr>
<tr>
<td>Irrelevant   Not-recalled</td>
<td>257.31 (53.85)</td>
</tr>
</tbody>
</table>
Table 5

Correlations between the Number and Duration of Fixations in the Relevant and Irrelevant Conditions and the Number of Integration Facts Recalled (Panel a) and Reading Comprehension Scores (Panel b) in Experiment 3

<table>
<thead>
<tr>
<th>Dependent Measure/Condition</th>
<th>Stimulus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel a: Integration facts recalled</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stim fact 1</td>
</tr>
<tr>
<td>Number of fixations</td>
<td>Relevant</td>
</tr>
<tr>
<td></td>
<td>Irrelevant</td>
</tr>
<tr>
<td>Duration of fixations</td>
<td>Relevant</td>
</tr>
<tr>
<td></td>
<td>Irrelevant</td>
</tr>
<tr>
<td><strong>Panel b: Reading comprehension</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of fixations</td>
<td>Relevant</td>
</tr>
<tr>
<td></td>
<td>Irrelevant</td>
</tr>
<tr>
<td>Duration of fixations</td>
<td>Relevant</td>
</tr>
<tr>
<td></td>
<td>Irrelevant</td>
</tr>
</tbody>
</table>

* = p ≤ .05; ** = p ≤ .01.

Note: