CHILD DEVELOPMENT

Child Development, January/February 2017, Volume 88, Number 1, Pages 247–262

Realizing Relevance: The Influence of Domain-Specific Information on Generation of New Knowledge Through Integration in 4- to 8-Year-Old Children

Patricia J. Bauer and Marina Larkina
Emory University

In accumulating knowledge, direct modes of learning are complemented by productive processes, including self-generation based on integration of separate episodes. Effects of the number of potentially relevant episodes on integration were examined in 4- to 8-year-olds (N = 121; racially/ethnically heterogeneous sample, English speakers, from large metropolitan area). Information was presented along with unrelated or related episodes; the latter challenged children to identify the relevant subset of episodes for integration. In Experiment 1, 4- and 6-year-olds integrated in the unrelated context. Six-year-olds also succeeded in the related context in forced-choice testing. In Experiment 2, 8-year-olds succeeded in open-ended and forced-choice testing. Results illustrate a developmental progression in productive extension of knowledge due in part to age-related increases in identification of relevant information.

Building a knowledge base is one of the most important tasks in development. Parents and other caregivers support the process with stories and causal explanations of the way the world works (see Gelman, 2009, for a review). As children enter school, the force of facilitators expands to include the teachers who impart information through formal instruction as well as the docents and aids who guide experiences in informal educational settings, such as museums (e.g., Jant, Haden, Uttal, & Babcock, 2014), for example. Critically, the information that children experience directly represents only a subset of their knowledge, owing to the productive quality of semantic memory (see Bauer, 2009, 2012; Bauer & Jackson, 2015, and Bauer & Varga, 2015, for discussions). That is, semantic memory is extended through logical processes such as analogy, deduction, and induction (see Goswami, 2011, for a review). Semantic memory also is productively extended through self-generation based on integration of separate episodes of experience; the process is more robust in older than in younger children (e.g., Bauer & San Souci, 2010). In this productive process, identification of relevant information is a necessary precursor to integration and self-generation of new knowledge and thus is a possible source of age-related variability in self-generation. In the present research, we tested the influence of the amount of potentially relevant information on the probability of integration and subsequent self-generation of new knowledge by 4- and 6-year-old (Experiment 1) and 8-year-old (Experiment 2) children. The effects of the experimental manipulation inform one potential source of observed age-related increases in self-generation, namely, identification of information relevant for integration.

In contrast to research on the productive processes of analogy (e.g., Gentner, 1983, 1989; Goswami, 2011; Goswami, 2013), deduction (e.g., Dias & Harris, 1988), and induction (e.g., Gelman & Markman, 1987), self-generation of new knowledge based on integration of separate learning episodes is only beginning to be explored. How the process might be affected by challenges to identification of information potentially relevant for integration has not been addressed. Bauer and San Souci (2010) conducted an initial investigation of self-generation through integration by 4- and 6-year-old children. In each of two passages of text, children learned a

Support for this research was provided by NICHD HD67359 to Patricia Bauer. We thank Jessica King, Ayla Nolen, and Elizabeth White for assistance with data collection and other members of the Memory at Emory laboratory group for their help with various aspects of the research. The authors extend a special note of gratitude to the children and families who took part in this research.

Correspondence concerning this article should be addressed to Patricia J. Bauer, Department of Psychology, Emory University, 36 Eagle Row, Atlanta, GA 30322. Electronic mail may be sent to patricia.bauer@emory.edu.

DOI: 10.1111/cdev.12584

novel fact (i.e., a “stem” fact), such as how dolphins talk (by clicking and squeaking) and the name of the groups in which they live (*pods*). The pair of related text passages was presented along with two other passages that featured different characters, actions, and concepts that were unrelated to dolphins (one passage each about kangaroos and volcanoes). Children then were asked a question that could only be answered by integrating the separate episodes about dolphins and using them to derive a new fact (i.e., “integration” fact). The integration-fact question first was presented in an open-ended format, requiring that children self-generate the integration fact, presumably through an inference based on the presented information. In the case of the dolphin example, the question was “How does a pod talk?” The information that “pods talk by clicking and squeaking” could be derived via inference from knowing how dolphins talk and that they live in groups called pods. On trials on which children failed to generate the integration fact, the question then was presented in a forced-choice format, requiring that children select the integration fact from among two distractors (three choices total).

The 4-year-olds in Bauer and San Souci (2010) self-generated the integration facts on only 13% of the open-ended trials. Their total performance (the sum of self-generation and forced-choice selection) was higher (62%). The results of the 1-stem control condition—in which the children were presented only one passage in a domain (i.e., one passage about kangaroos and one about volcanoes)—made clear that exposure to both members of the pair of passages was necessary to support production of the integration fact. On the 1-stem control trials, the 4-year-olds produced none of the integration facts and selected them from among distractors 33% of the time (i.e., chance-level performance). The 6-year-olds self-generated the integration facts on 67% of the open-ended trials on which they were presented with both stems (2-stem condition), compared with only 17% of the 1-stem control trials. Their total performance was even higher, reaching 93%, compared with 63%, in the 1-stem control condition. Thus, both 4- and 6-year-olds productively extended their semantic knowledge. Successful performance was dependent on exposure to both members of a stem-fact pair, indicating that integration of the separate episodes was necessary for productive self-generation. Four-year-olds showed this productive ability primarily through forced-choice selection; 6-year-olds self-generated the new information in an open-ended format.

In the present research, we investigated one potential source of observed age-related variability in self-generation through integration, namely, identification of information relevant for integration. To date, studies of self-generation of new knowledge through integration have been conducted under relatively information-constrained circumstances, such that only a subset of the text passages presented to the children were relevant to the question posed. In the case of the dolphin example, two of the four passages mentioned dolphins, pods, or talking; the other two passages featured content that was irrelevant to a question about how pods talk. As a result, even in open-ended testing, the work of selecting which pieces of information to integrate in order to respond to the question was relatively well supported for the child. Under these conditions, 6-year-old children were successful. In contrast, 4-year-olds were successful primarily in forced-choice testing, in which the correct answer was presented among distractors. In this circumstance, the correct response likely served as a cue to the information that was relevant to the question. The information-constrained circumstances of testing and the different conditions of success for 4- and 6-year-olds raise two related questions. The first is how 4- and 6-year-old children would fare when challenged to identify a subset of information relevant to a particular question or task demand under less information-constrained conditions. The second is whether identification of relevant information is a source of age-related variability in self-generation of new knowledge through integration. A direct manipulation of the requirement to identify relevant information from among potential distractors can be expected to inform both of these questions.

Prior research has not provided a direct test of whether identification of the relevance of separate episodes to one another is a source of variance in self-generation of new knowledge, though results are suggestive. One source of suggestive evidence is Bauer, King, Larkina, Varga, and White (2012), in which the level of surface similarity between to-be-integrated text passages was manipulated by changing the main character in the passages. In the high-similarity condition, the same character (e.g., a ladybug) “learned” a novel stem fact in each passage. In the low-similarity condition, different characters (e.g., ladybug and rabbit) were featured in each passage. In this condition, 6-year-old children’s self-generation performance suffered, relative to the high-similarity condition (37% vs. 63%, respectively). A logical explanation of the effect is that the high surface similarity between related passages emphasized the common features across the episodes, making it easier to identify the relevance of the passages to one another and to the question posed. In contrast, low surface similarity between conditions made it more difficult to appreciate the relatedness of separate episodes to one another, thus lowering the probability that they would be integrated. Similar findings have been obtained in other knowledge extension paradigms, such as analogical problem solving (e.g., Brown & Kane, 1988; Brown, Kane, & Echols, 1986; Gentner & Toupin, 1986; Holyoak, Junn, & Billman, 1984).

Hints and suggestions also seemingly change the probability of identification of the relevance of separate episodes to one another and thus successful integration and self-generation of new knowledge (e.g., Bauer et al., 2012). Specifically, when delivered immediately prior to a test prompt, hints to "Think about the stories we read," facilitate self-generation of new knowledge for both 4- and 6-year-old children. In contrast, hints delivered between presentation of pairs of related passages (e.g., "Think about the story we read about dolphins") do not facilitate performance. Bauer, Varga, King, Nolen, and White (2015) interpreted this pattern as evidence that when delivered just before the test prompt, the hints highlighted the relevance of the passages to one another and to the question being posed (i.e., through explicit reference to the stories). Critically, for the 6-year-olds, hints delivered immediately prior to the test prompt facilitated performance even under low surface similarity conditions (see also Bauer et al., 2012). Hints and suggestions also facilitate performance in analogical problem solving (e.g., Brown et al., 1986; Crisafi & Brown, 1986).

In the present research, we addressed the questions of the effects on self-generation through integration of less information-constrained testing circumstances and whether identification of material that is relevant for integration and self-generation is a source of age-related variance in this productive means of knowledge extension. To do so, we experimentally manipulated the amount of potentially relevant information provided to 4- and 6-year-old (Experiment 1) and 8-year-old (Experiment 2) children. Specifically, we tested the children under three conditions, in which they were provided one, two, or four text passages that were potentially relevant to questions posed. The one-passage condition was a control for the necessity for integration of separate passages in order to respond to the question (i.e., similar to the 1-stem control condition described above). In the 2-stem condition, as in prior research (e.g., Bauer & San Souci, 2010), two passages were related to one another and could be integrated to support generation of new knowledge, whereas the other two passages were unrelated. In the 2-stem plus related condition, four text passages were potentially relevant to the questions posed. Specifically, two passages could be integrated to support generation of new knowledge (as in the 2-stem condition). The other two passages concerned the same conceptual domain but did not provide information that was useful for self-generating an inferential response to the question posed. As such, they were potentially distracting to integration.

We expected the larger body of related information to pose a challenge to identification of the subset of information relevant to integration of separate episodes and self-generation of new knowledge from them. We reasoned that if children failed to productively extend their knowledge in the less information-constrained, 2-stem plus related condition—yet succeeded when provided fewer conceptually related items within the target domain (2-stem condition)—the outcome would suggest that the source of limitation on performance was failure to identify the information relevant for integration. In Experiment 1, we tested this hypothesis with children 4 and 6 years of age. We selected these age groups based on prior research that has demonstrated age-related improvements in self-generation of new knowledge through integration in this period (Bauer & San Souci, 2010). To the extent that identification of relevant material is a source of challenge to productive extension of knowledge through integration, we expected lower levels of performance in the less information-constrained, 2-stem plus related condition relative to the more information-constrained 2-stem condition. To the extent that identification of relevant material is a source of age-related variance, we expected age differences in the 2-stem plus related condition in particular, with 6-year-olds providing stronger evidence of integration, relative to 4-year-olds.

To anticipate the results of Experiment 1, we found that among the 4-year-olds, the additional domain information had a detrimental effect on performance. Among the 6-year-olds, the manipulation had a detrimental effect on open-ended task performance. The negative effect was neutralized when children were permitted to select integration facts from among distractors. In light of the developmental difference between 4- and 6-year-olds, and the finding that even the older age group apparently was challenged to identify the relevant subset of information from among a larger body of information, in Experiment 2, we extended the age range of the study to 8-year-old children. Children of this age have not previously been tested for self-generation of new knowledge through integration of separate text passages. We predicted that they would be more likely to identify the subset of information that was relevant to the questions posed, even in the less information-constrained, 2-stem plus related condition.

In summary, in these two experiments, we tested 4- and 6-year-old (Experiment 1) and 8-year-old (Experiment 2) children's self-generation of new knowledge through integration of information presented in separate passages of text. We tested the processes in contexts in which there were more and fewer potential targets for integration. We predicted that additional information within a domain would present a challenge to identification of relevant episodes for integration. The results stand to inform whether identification of relevant episodes for integration is one of the cognitive processes involved in self-generation of new knowledge through integration, and whether it is a source of age-related variability in this productive process. The results also have the potential to inform the applied question of the conditions of formal and informal education under which self-generative processes may be facilitated versus impaired.

Experiment 1

Method

Participants

The participants were forty-one 4-year-olds (21 girls and 20 boys, *M* = 4 years 6 months; range = 4;0–4;10) and forty-one 6-year-olds (18 girls and 23 boys, *M* = 6 years 5 months; range = 6;0–6;11). Children were pseudorandomly assigned (constrained by approximate gender balance) to one of three conditions: two experimental conditions and one control condition (*n*s = 16, 16, and 9, respectively; see below). The children were recruited between 2012 and 2013 from a volunteer pool consisting of English-speaking families from a large metropolitan area in the southeastern United States who had expressed interest in participating in child development research. Based on parental report, the sample was 27% African American, 13% bi- or multiracial, and 60% Caucasian; one parent did not respond to the request for racial self-classification. Six percent of the sample self-identified as Hispanic. Although no specific information on parental income or occupation was collected, the pool consists of families from middle- to upper-middle socioeconomic status. Two additional 4-year-olds were tested but were excluded from data analyses due to severe developmental delay (*n* = 1) and experimenter error (*n* = 1). Parents provided informed written consent and 6-year-olds gave verbal assent at the beginning of the session. For this and the subsequent experiment, a university Institutional Review Board approved the protocol and procedures. At the end of the hour-long session, children received an age-appropriate toy to acknowledge their participation and their parents were given a $5 gift card.

Stimuli

The stimuli were two novel "stem" facts and three novel "nonstem" facts from each of three different domains: dolphins, palm trees, and deserts (see Table 1). Within each domain, the two novel stem facts were related and could be combined to generate a novel integration fact. Within each domain, the three novel nonstem facts were about the same concept (e.g., all facts about dolphins), but they could not be combined to generate the target integration facts. The stem fact and integration fact stimuli from the domain of dolphins had been used in prior related research (Bauer & San Souci, 2010; Bauer et al., 2012; Bauer, Varga, et al., 2015; Varga & Bauer, 2013); the nonstem facts were novel to this study. All of the stimuli from the domains of palm trees and deserts were new to this study. All facts were accurate and determined by pilot testing to be novel for children in the target age range.

Each stem fact and nonstem fact was presented in the context of a text passage read aloud by an experimenter (see Bauer & San Souci, 2010, for an example text passage). The passages were 81–89 words in length, distributed over four pages. Each page consisted of a hand-drawn color illustration depicting the main actions of the text; the text was not featured on the page. The passages were similar in structure: In each passage, a character (e.g., ladybug) learned a new fact in the course of an "adventure." Within a domain, the characters were the same across the passages; across the domains, the characters were different. Only the stem and nonstem facts were included in the passages; the integration facts were not presented.

Because of the highly verbal nature of the task and the between-subjects design, there is the

Table 1
Stem Facts, Integration Facts, and Nonstem Facts Used in Experiments 1 and 2

	Domain		
Item	Dolphins	Palm trees	Deserts
Stem Fact 1	*Dolphins* talk by clicking and squeaking	*Palm tree* leaves are used to make baskets	Largest *desert* in the world is Sahara
Stem Fact 2	*Dolphins* live in groups called pods	*Palm tree* leaves are called fronds	Largest *desert* in the world is located in Africa
Integration fact	Pods talk by clicking and squeaking	Fronds are used to make baskets	Sahara is located in Africa
Nonstem Fact 1	*Dolphin's* tail is called a fluke	*Palm trees* lived at the same time as dinosaurs	*Deserts* get < 10 inches of rain a year
Nonstem Fact 2	*Dolphins* swim while they sleep	*Palm tree* flowers are white	There are hot *deserts* and cold *deserts*
Nonstem Fact 3	*Dolphin* teeth are shaped like cones	Dates come from *palm trees*	*Deserts* could be of five different kinds

Note. See Bauer and San Souci (2010), for an example of the facts as presented in story passage context.

potential for concern that group differences might result from uncontrolled variability in verbal comprehension. As a means of test for uncontrolled differences between the groups of children pseudorandomly assigned to each of the experimental conditions, we administered four subscales of the Woodcock–Johnson III Test of Verbal Comprehension (Woodcock, McGrew, & Mather, 2001): picture vocabulary, synonyms, antonyms, and verbal analogy.

Procedure

The procedure was modeled after that used in the previous related research (Bauer & San Souci, 2010; Bauer et al., 2012; Varga & Bauer, 2013). Children were tested individually in a laboratory room by one of three female experimenters. The sessions were video recorded and were regularly reviewed by the experimenters with one another to ensure protocol fidelity. The 1-hr session was divided into two phases: (a) exposure to stem and nonstem facts and (b) test for self-generation of new knowledge through integration.

Exposure to stem and nonstem facts. Children participated in one of three conditions: 2-stem, 2-stem plus related, and 1-stem plus related control. In each condition, children were exposed to four text passages. The experimenter read each passage twice in succession. After presentation of the first two passages, children were engaged in approximately 15 min of buffer activities unrelated to the purposes of the present research. The experimenter then presented the remaining two passages, followed by approximately 15 min during which we administered the Woodcock–Johnson III Test of Verbal Comprehension (Woodcock et al., 2001).

The 2-stem condition has been used in all prior research. Children were read four passages of text. Two of the passages featured stem facts from the same domain. For example, for the domain of dolphins, the two stems facts were "dolphins talk by clicking and squeaking" and "dolphins live in groups called pods." When integrated with one another, these two stem facts supported generation of the integration fact that "pods talk by clicking and squeaking." In the 2-stem condition, the other two passages of text were unrelated to the target domain: they featured one of the stem facts from each of the other two domains (palm trees and deserts). Because children were exposed to only one stem-fact passage from each of these domains, they were not expected to produce integration facts from these domains.

In the 2-stem plus related condition, children were exposed to a pair of related stem-fact passages, just as in the 2-stem condition. In addition to the two passages of text that featured stem facts that could be integrated with one another to generate a novel integration fact, children were exposed to two other passages of text that featured facts from the same domain, but which could not be combined to generate the integration fact (nonstem facts). For example, in the dolphin domain, the nonstem facts could be "a dolphin's tail is called a fluke" and "dolphins swim while they sleep" (see Table 1). Thus, in the 2-stem plus related condition, all facts were from the same domain (e.g., all four passages of text featured facts about dolphins); two of the facts could be combined to generate the target integration fact.

In the 1-stem plus related control condition, children were exposed to four passages of text, all featuring information from the same domain. One of the passages featured a stem fact and the other three passages featured nonstem facts. The 1-stem plus related condition was included as a control for spontaneous generation of the integration facts. Prior related research has made clear that children do not produce the novel integration facts unless they are exposed to both members of a stem-fact pair (e.g., Bauer & San Souci, 2010). We included this control in the present research to test whether self-generation of novel integration facts was possible when exposed to only one stem fact but in the context of other information about the same concept.

In the 2-stem and 2-stem plus related conditions, one stem-fact passage was read before the 15 min of buffer activities and the other stem-fact passage was read after the buffer activities, thus ensuring a delay of approximately 15 min between presentation of the stem facts. In the 1-stem plus related control condition, presentation of the single stem-fact passage was counterbalanced, such that it was provided before and after the buffer activities approximately equally often across children. Across children, each domain was used approximately equally often in each of the three conditions. Also across children, order of presentation of the stem facts was counterbalanced, such that each stem-fact passage was presented approximately equally often before and after the buffer activities. In addition, in the 2-stem conditions, across children, each nonstem fact was used approximately equally often.

Test for self-generation of integration facts in open-ended and forced-choice selection formats. Approximately 15 min after presentation of the final passage of text (interval filled by the language assessments), children were tested in open-ended and forced-choice formats. The testing procedures were identical across the three conditions. There were four phases of testing. First, using open-ended questions, we tested for production of the integration facts. For all children, regardless of condition, the experimenter asked each of three integration-fact questions, one from each of the three domains: "How does a pod talk?" (dolphin domain), "What are fronds used to make?" (palm tree domain), and "Where is the Sahara located?" (desert domain).

Second, the experimenter asked five fact-recall questions: two stem-fact questions and three nonstem-fact questions (see Table 1). All questions were from the domain in which the children had been exposed to both stem-fact passages (2-stem and 2-stem plus related conditions) or to the single stem-fact passage (1-stem plus related control condition). Note that no child was actually exposed to all five stem and nonstem facts within a domain. Rather, as a function of condition, they were exposed to 2 stem facts and 0 nonstem facts (2-stem condition), 2 stem facts and 2 nonstem facts (2-stem plus related condition), or 1 stem fact and 3 nonstem facts (1-stem plus related control condition). To ensure comparable testing experiences across experimental conditions, all children were tested on 2 stem and 3 nonstem facts.

Third, for each integration-fact question not correctly answered in open-ended format, the experimenter administered forced-choice questions. Each forced-choice question had three response options, one of which was correct. For example, for the question, "how does a pod talk?" the three answer choices were: (a) by clicking and squeaking, (b) by splashing and jumping, and (c) by rubbing noses.

Finally, children were asked forced-choice questions for the stem and nonstem facts that they did not accurately recall in open-ended testing. As for the integration-fact questions, all forced-choice questions had three alternatives, one of which was correct.

To ensure that children had some success during open-ended and forced-choice questioning, in each test phase, we included three questions to which children were likely to know the answers, such as "what sound does a duck make?" and "what color is snow?" These questions were not scored. Testing for the integration facts and stem and nonstem facts was administered in the fixed order just described. For each domain, two orders of presentation of the test questions were created and each was used approximately equally often across participants.

Scoring

The experimenters recorded the children's responses as they were made. In open-ended testing, children received a score of 1 or 0 (correct or incorrect) for production of the integration fact in the domain to which they had been exposed to both stem-fact passages (2-stem and 2-stem plus related conditions) or to the single stem-fact passage (1-stem plus related control condition).

We also derived a weighted total score that was the number of forced-choice selection questions answered correctly added to the number of open-ended questions answered correctly, with 1 point awarded for correct forced-choice selection and 2 points awarded for self-generation in open-ended testing, for a total range of 0–2. In prior related research, a total score that was the simple sum of correct responses in open-ended and forced-choice testing was derived (e.g., Varga & Bauer, 2013). In the present research, we adopted the weighted scoring procedure for two reasons. First, children who generated the novel integration facts in open-ended testing were not permitted the opportunity to earn credit in forced-choice testing yet could reasonably have been expected to make the correct selection in forced-choice testing. Thus, had they participated in both phases of testing, they would have earned 2 points. Second, open-ended performance is generally acknowledged to be more challenging than forced-choice selection, making it reasonable to award a higher score to children who met the greater challenge. Assigning different point values to different phases of testing also explicitly acknowledges that children earned their total score in different ways: through self-generation versus forced-choice (see Ghetti, Goodman, Eisen, Qin, & Davis, 2002; Gluckman, Vlach, & Sandhofer, 2014; Neisser & Harsch, 1992, for similar approaches).

Finally, in both 2-stem conditions, we noted the number of trials on which children recalled both members of the pair of stem facts to which they had been exposed. We used this information to examine the relation between recall of the stem facts and self-generation of the novel integration facts.

Results

Preliminary Analyses

We conducted one-way between-subjects analyses of variance to test whether, within an age group, children pseudorandomly assigned to the different experimental conditions differed in their general language abilities (Woodcock et al., 2001). The within-age-group analyses revealed no significant differences between conditions, for either age group: $F(2, 38) = 0.13$ and 0.66, $ps > .55$, $\eta^2 = .007$ and .034, for 4- and 6-year-olds, respectively.

Main Analyses

Depictions of children's performance in (a) open-ended testing and (b) across open-ended and forced-choice selection (i.e., total performance) are provided in Figure 1, Panels a and b, respectively. In two conditions, namely, the 2-stem and 2-stem plus related conditions, children had both of the stem facts necessary to generate the novel integration facts in the target domain. The 1-stem plus related condition served as a control to test whether integration of the pair of related stem facts was necessary for production, even in the domain-information-rich condition: Children were not expected to successfully produce the integration facts in this condition. For the open-ended phase of testing, because of the dichotomous nature of the data (children either did or did not generate the integration fact in open-ended testing), we conducted between-group comparisons using chi-square or Fisher's exact tests (when one or more of the cells in a frequency table had an expected frequency of 5 or less). For weighted total performance (which had a range of 0–2 for each child), we used parametric statistics.

We first examined group differences in children's open-ended performance in each 2-stem condition (2-stem, 2-stem plus related) relative to the 1-stem plus related control condition, for each age group separately. For the 4-year-olds, levels of open-ended performance were low, in all conditions and did not differ by condition ($ps > .24$, Fisher's exact test). As reflected in Figure 1, Panel a, in the 2-stem condition, only 13% ($n = 2$) of the 4-year-olds produced the novel integration facts (this percentage is identical to the analogous testing condition in Bauer & San Souci, 2010). In the 2-stem plus related and 1-stem plus related control conditions, none of the 4-year-olds produced the novel integration facts in open-ended testing.

As in prior related research (Bauer & San Souci, 2010), greater evidence of 4-year-old children's productive extension of knowledge was apparent in their total performance (weighted sum of correct responses in open-ended [2 points] and forced-choice [1 point] testing formats). In the 2-stem condition, the 4-year-old children's mean weighted total score was 0.63 ($SD = .72$); the percentage of children who either self-generated the novel integration fact or selected it from among distractors in forced-choice testing is reflected in Figure 1, Panel b (50%; $n = 8$). In contrast, in the 1-stem plus related control condition, the 4-year-olds' mean weighted total score was 0.11 ($SD = .33$); only one child (11%) selected the novel integration fact in forced-choice testing. The difference between the conditions was statistically significant, $t(22.497) = 2.43$, $p = .02$, and large, Cohen's $d = 0.928$. In the 2-stem plus related condition, none of the 4-year-olds generated the integration facts in open-ended testing; only 4 (25%) of the children selected the novel integration fact in forced-choice testing ($M = .25$, $SD = .47$). Four-year-old children's performance in the 2-stem plus related condition did not differ from that in the 1-stem plus related control condition, $t(23) = 0.81$, $p = .43$, Cohen's

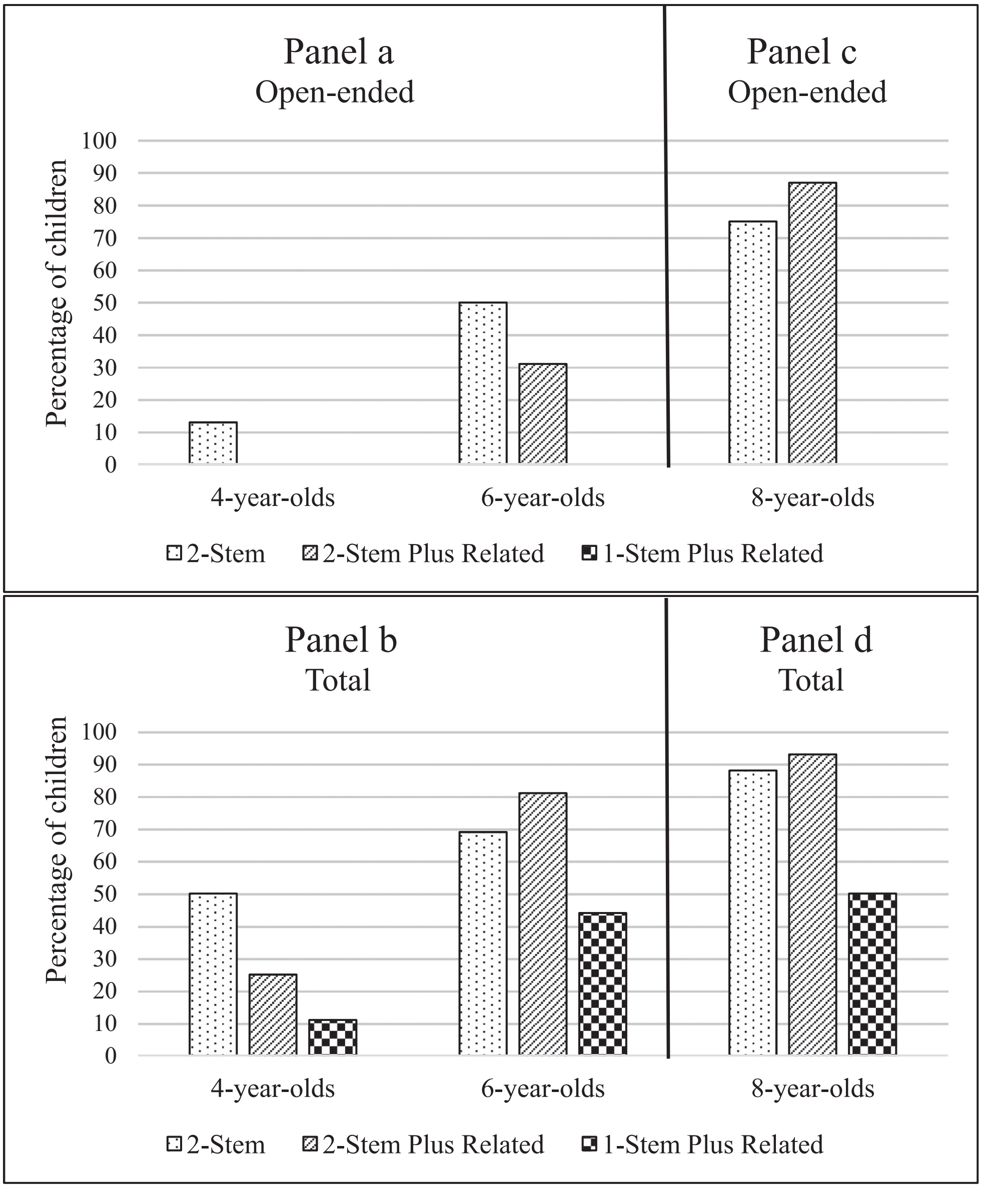

Figure 1. For Experiment 1, percentages of 4- and 6-year-old children who self-generated novel integration facts in open-ended testing (Panel a) and across open-ended and forced-choice testing (i.e., total performance; Panel b). For Experiment 2, percentages of 8-year-old children who self-generated novel integration facts in open-ended testing (Panel c) and across open-ended and forced-choice testing (i.e., total performance; Panel d). In open-ended testing (Panels a and c) the means were 0 for all age groups in the 1-stem plus related condition, and for the 4-year-olds in the 2-stem plus related condition.

d = 0.355, indicating a small effect size. Overall, 4-year-old children showed evidence of self-generation of new knowledge through integration in the 2-stem condition only (the condition used in prior related research; e.g., Bauer & San Souci, 2010). The evidence was derived primarily from their performance in the forced-choice selection phase of testing.

In contrast to the 4-year-olds, 6-year-olds enjoyed more success in open-ended testing. As reflected in Figure 1, Panel a, in the 2-stem condition, 50% (n = 8) of the children self-generated the integration facts. The level of performance was significantly different from performance in the 1-stem plus related control condition, in which none of the children produced the novel integration facts in open-ended testing (p = .02, one-tailed Fisher's exact test, φ = .514, indicating a large effect size). In the 2-stem plus related condition, 31% (n = 5) of 6-year-olds self-generated the integration fact. The level of performance was not statistically different from that in the 1-stem plus related control condition, however ($p < .09$, one-tailed Fisher's exact test), even though the effect size would be considered medium ($\varphi = .375$). Thus, 6-year-olds showed evidence of productive extension of new knowledge through integration in open-ended testing in the 2-stem condition but not in the 2-stem plus related condition.

Six-year-olds' total performance (across open-ended and forced-choice testing) was strong in both 2-stem conditions. Specifically, their mean weighted total scores were 1.19 ($SD = .91$) and 1.13 ($SD = .72$) in the 2-stem and 2-stem plus related conditions, respectively. As reflected in Figure 1, Panel b, 69% ($n = 11$) and 81% ($n = 13$) of the children either self-generated the novel integration fact or selected it from among distractors in forced-choice testing in the 2-stem and 2-stem plus related conditions, respectively. In contrast, in the 1-stem plus related control condition, the mean total weighted score was .44 ($SD = .53$); 44% ($n = 5$) of 6-year-old children selected the novel integration facts. Their level of performance was significantly different from that of children in both the 2-stem and 2-stem plus related conditions: ts(23) = 2.23 and 2.48, ps = .04 and .02, Cohen's ds = 1.007 and 1.091, respectively, indicating large effect sizes. In summary, 6-year-olds showed evidence of self-generation of new knowledge through integration in the 2-stem condition in both open-ended and forced-choice testing. In the 2-stem plus related condition, evidence of self-generation of new knowledge was apparent when both open-ended and forced-choice selections were considered jointly.

The patterns of performance just described imply age-related differences in children's success on the task before them. The differences were confirmed through direct comparisons of performance in the two age groups. For the open-ended phase of testing, the levels of performance differed between the 4- and 6-year-olds in the 2-stem and 2-stem plus related conditions ($p = .03$ and .02, respectively, Fisher's exact test, $\varphi = .405$ and .430, indicating medium effect sizes). In the 1-stem plus related control condition, neither 4- nor 6-year-olds generated any integration facts. Weighted total scores (across open-ended and forced-choice testing) differed significantly between the 4- and 6-year-olds in the 2-stem plus related condition, $t(30) = 4.13$, $p < .001$, Cohen's $d = 1.462$, indicating a large effect size. In the 2-stem condition, the difference between the groups approached statistical significance, $t(30) = 1.94$, $p = .06$, Cohen's $d = 0.685$. Performance did not differ between the groups in the 1-stem plus related control condition $t(16) = 1.60$, $p = .13$, Cohen's $d = 0.755$.

Finally, we examined children's recall of the stem facts and the relation between stem-fact recall and open-ended performance. For this examination, we did not consider the performance of children in the 1-stem plus related control condition, given that they were exposed to only one of the pair of stem facts. We also confined the examination to performance in open-ended testing, because it is most clearly indicative of self-generation based on integration. The examination is descriptive only, given the small number of observed and expected observations per cell. Inspection of the values in Table 2, Panel a, makes clear that for the 4-year-olds, there was not a strong relation between recall of the stem facts and self-generation of the integration facts in open-ended testing. In the 2-stem condition, of the two 4-year-olds who self-generated the novel integration fact, one recalled both members of the stem fact pair in open-ended testing and one did not. Of the two 4-year-olds who recalled both members of the stem-fact pair, one had self-generated the novel integration fact and one had not. Among the 6-year-olds, seven of the eight children who self-generated the novel integration fact also recalled both members of the stem-fact pair in open-ended testing. As well, seven of the eight 6-year-olds who recalled both members of the stem-fact pair also had self-generated the novel integration fact. Thus, for the 6-year-olds in the 2-stem condition, there was a tight coupling between recall of both members of the stem-fact pair and productive self-generation.

In the 2-stem plus related condition (see Table 2, Panel a), none of the 4-year-olds self-generated the novel integration fact in open-ended testing, even though two of them recalled both members of the stem-fact pair. Among the 6-year-olds, four of the five children who self-generated the novel integration fact also recalled both members of the stem-fact pair in open-ended testing. However, only four of the seven 6-year-olds who recalled both members of the stem-fact pair also had self-generated the novel integration fact, whereas three had not. Thus, in the 2-stem plus related condition, there was not a strong relation between recall of both members of the stem-fact pair and productive self-generation for either age group. Importantly, the fact that comparable numbers of children in both age groups recalled both members of the stem-fact pairs in the two 2-stem conditions (2 in the 2-stem and 2-stem plus related conditions for the 4-year-olds; 8 and 7 in the 2-stem and 2-stem plus related conditions for

Table 2
Descriptive Relations Between Recall of Stem Facts and Self-Generation of New Knowledge Through Integration

	Self-generated integration fact			Recalled stem-fact pair		
Condition/Age group	Total *N*	Recalled pair (*N*/%)	Did not recall pair (*N*/%)	Total *N*	Self-generated (*N*/%)	Did not self-generate (*N*/%)
Panel a: Experiment 1						
2-Stem condition						
4-year-olds	2	1 (50)	1 (50)	2	1 (50)	1 (50)
6-year-olds	8	7 (88)	1 (12)	8	7 (87)	1 (13)
2-Stem plus related condition						
4-year-olds	0	NA	NA	2	0	2 (100)
6-year-olds	5	4 (80)	1 (20)	7	4 (57)	3 (43)
Panel b: Experiment 2						
2-Stem condition						
8-year-olds	12	10 (83)	2 (17)	12	10 (83)	2 (17)
2-Stem plus related condition						
8-year-olds	13	12 (92)	1 (8)	13	12 (86)	1 (14)

the 6-year-olds) means that, within an age group, the difference in self-generation performance between the conditions cannot be attributed to differential likelihood of recall of the stem facts.

Discussion

The present experiment provided suggestive evidence that identification of relevant episodes for integration is a source of challenge in self-generation of new knowledge through integration, and that it is a source of age-related variability in this productive process. No 4-year-old child self-generated the novel integration facts in the 2-stem plus related condition, and only 4 children selected the novel integration facts in forced-choice testing. Four-year-olds' performance in the 2-stem plus related condition did not differ from that in the 1-stem plus related control condition. In contrast, a number of 4-year-olds either generated (n = 2) or selected (n = 6, for a total n = 8) the novel integration facts in the 2-stem condition; performance in the 2-stem condition was significantly higher than in the 1-stem plus related control condition. The fact that performance with additional domain information did not differ from performance in the control condition, whereas performance in the "standard" 2-stem condition did differ is consistent with the suggestion that under less information-constrained conditions, the 4-year-old children had difficulty identifying information that was—and was not—relevant to productive extension.

Like the 4-year-olds, the 6-year-olds had difficulty identifying information that was relevant to productive extension of knowledge through integration. Yet in contrast to the younger children, the 6-year-olds "recovered" when given the opportunity to select the integration facts from among distractors in forced-choice testing. Under forced-choice conditions, 6-year-olds' performance in the 2-stem plus related condition was significantly greater than in the 1-stem plus related control condition. It seems that faced with the correct answer to the integration fact question in the forced-choice format, the 6-year-olds no longer were challenged to identify relevant information. Notably, because in both age groups, comparable numbers of children recalled both members of the stem-fact pairs in the 2-stem and 2-stem plus related conditions, the difference in self-generation performance between the conditions cannot be attributed to differential likelihood of recall of the stem facts (see General Discussion, for discussion of the role of memory for the stem facts in explaining age-related differences in performance).

Experiment 1 indicated that especially in open-ended testing, additional domain-specific information was detrimental to performance by 4- and 6-year-old children. The pattern of performance is consistent with the suggestion that in less information-constrained conditions, the 4-year-olds failed to identify the textual material that was relevant to the test probes, even when presented with the strong cues provided by the forced-choice options. With the support provided in the forced-choice context, the 6-year-olds overcame the ambiguity of the numerous potential options for integration and selected the relevant pair of related passages of text. In Experiment

2, we extended the research to 8-year-old children. The study is the first test of self-generation of new knowledge through integration of text passages in this age group. We tested whether the older children would integrate the separate yet related passages of text, such that even in open-ended testing, they would show evidence of self-generation of new knowledge both without and with additional domain information.

Experiment 2

Method

Participants

The participants were 39 eight-year-olds (22 girls and 17 boys, M = 8 years 6 months; range = 8;1–8;11). Children were pseudorandomly assigned (with approximately equal gender balance) to the 2-stem, 2-stem plus related, and 1-stem plus related control conditions (ns = 16, 15, and 8, respectively). The children were recruited in 2014 from the same source and represent the same population as in Experiment 1. None of the children had participated in Experiment 1. Based on parental report, the sample was 23% African American, 3% Asian, 5% biracial, and 67% Caucasian; one parent did not respond to the request for racial self-classification. Ten percent of the sample self-identified as Hispanic. At the end of the hour-long session, children received an age-appropriate toy to acknowledge their participation and their parents were given a $5 gift card. Parents provided informed written consent and children gave verbal assent at the beginning of the session.

Stimuli, Procedure, and Scoring

The stimuli were the same as in Experiment 1. The children were tested by one female experimenter, who also served as an experimenter in Experiment 1. The procedure was identical to that used in Experiment 1. The scoring procedure was identical to that used in Experiment 1.

Results

As in Experiment 1, we conducted a one-way between-subjects analysis of variance to test whether children pseudorandomly assigned to the different experimental conditions differed in their general language abilities (Woodcock et al., 2001). There was not a significant difference between groups, $F(2, 36) = 1.73$, $p = .19$, $\eta^2 = .090$.

Graphic depictions of children's performance (a) in open-ended testing and (b) across open-ended and forced-choice testing are provided in Figure 1, Panels c and d, respectively. Even in open-ended testing, children demonstrated high levels of performance in both conditions in which they received both stem facts from the target domain. In the 2-stem and 2-stem plus related conditions, 75% (n = 12) and 87% (n = 13) of 8-year-olds self-generated the integration facts, respectively. The children did not produce any integration facts in the 1-stem plus related control condition. The difference in performance between the 1-stem plus related control and 2-stem conditions was significant for both two 2-stem conditions (ps < .001, one-tailed Fisher's exact tests) and large in magnitude (φ = .707 and .833).

Eight-year-olds' average weighted total scores (across open-ended and forced-choice testing) reached 1.63 and 1.80 (SDs = .72 and .56) in the 2-stem and 2-stem plus related conditions, respectively; 88% (n = 14) and 93% (n = 14) of the children either self-generated the novel integration fact or selected it from among distractors in forced-choice testing in the 2-stem and 2-stem plus related conditions, respectively. In contrast, in the 1-stem plus related control condition, the average weighted total score was 0.50 (SD = .53); 50% (n = 4) of 8-year-old children selected the novel integration facts from among distractors. Their level of performance was significantly different from that of children in both the 2-stem and 2-stem plus related conditions: $t(22) = 3.90$ and $t(21) = 5.38$, ps < .001, Cohen's ds = 1.787 and 2.384, respectively, indicating large effect sizes.

As in Experiment 1, we examined children's recall of the stem facts and the relation between stem-fact recall and open-ended performance. As reflected in Table 2, Panel b, in the 2-stem condition, 10 of the 12 eight-year-olds who self-generated the novel integration facts also recalled both members of the stem-fact pair in open-ended testing. Of the twelve 8-year-olds who recalled both members of the stem-fact pair, ten had self-generated the novel integration fact and only two had not. In the 2-stem plus related condition, 12 of the 13 eight-year-olds who self-generated the novel integration facts also recalled both members of the stem-fact pair in open-ended testing. Of the 13 children who recalled both members of the stem-fact pair, 12 had self-generated the novel integration fact and only 1 had not. Thus, for the 8-year-olds, there was a tight relation between self-generation of the novel integration facts and recall of both members of the stem-fact pairs.

Discussion

The 8-year-old children in the present experiment were successful in extending their knowledge through integration of information acquired in separate passages of text. They were successful both when the pair of related passages was presented in the context of unrelated passages of text and when the other passages featured additional information within the target domain. The fact that 8-year-old children's performance was not negatively impacted by the less informationally constrained context, even in open-ended testing, suggests that the children successfully identified the pairs of related stem facts necessary for derivation of new knowledge.

General Discussion

One major purpose of the present research was to investigate the influence of an information-rich context on 4-, 6-, and 8-year-old children's productive extension of knowledge through integration of separate episodes of text. By experimentally manipulating the amount of information provided within a target domain, we were able to test how effectively children identified the subset of information that was relevant for integration, a necessary prerequisite to successful self-generation of new knowledge. This in turn advanced the second major purpose of the present research, namely, to test the hypothesis that identification of relevant material is one source of age-related variance in self-generation of new knowledge through integration.

Performance in open-ended testing provided clear evidence of a developmental progression in self-generation of new knowledge through integration of separate yet related episodes, as well as suggestive evidence that the progression was due at least in part to the requirement to identify information relevant for integration. Among the 4-year-olds (Experiment 1), performance was essentially at floor levels, with only two children generating the novel integration facts in open-ended testing. Both children were in the more information-constrained, 2-stem condition in which only two of the four passages of text to which they were exposed were relevant to the integration questions posed. In contrast, 6- (Experiment 1) and 8-year-old (Experiment 2) children productively extended their knowledge through self-generation in open-ended testing. Among the 6-year-olds, in open-ended testing, self-generation was observed only in the more information-constrained 2-stem condition; in open-ended testing, performance in the less information-constrained 2-stem plus related condition did not differ from that in the control condition in which children were provided with only one of the two passages of text necessary for self-generation (though the effect approached significance and was of medium size). Among the 8-year-olds, self-generation was observed in both of the conditions in which they were exposed to both passages necessary for self-generation. The pattern of findings for the 4- and 6-year-olds in the more information-constrained 2-stem condition is a replication of prior research (e.g., Bauer & San Souci, 2010; Varga & Bauer, 2013). The present research represents the first test of productive extension through integration of text passages in 8-year-old children. The finding indicates that 8-year-old children not only self-generate new knowledge through integration of separate episodes, they do so even when challenged to identify the specific information that will lead to productive extension from among potential distractors.

Comparison of open-ended performance in the two conditions in which the children were provided both of the text passages necessary for self-generation sheds light on a source of age-related improvement in productive self-generation of new knowledge in this paradigm. We suggest that a source of change is the extent to which children engage in processes that are necessary (though not sufficient) precursors to self-generation. Specifically, we propose that self-generation of new knowledge through integration requires that children (a) identify the subset of episodes of experience that are relevant to the question or problem at hand, (b) integrate the episodes with one another, and (c) use them as a basis for inferential self-generation of the novel understanding that informs the question. This first step in this process is complicated in information-rich contexts in which there are several episodes that are potentially relevant to self-generation. It seems that 4-year-olds are challenged to identify relevant information even when it is presented among episodes that are poor candidates for informing the question at hand. Six-year-olds seemingly succeed when there is a low level of ambiguity as to what material is relevant to the question posed (as evidenced by high levels of self-generation in the more information-constrained 2-stem condition), but not under conditions of higher levels of ambiguity. Eight-year-old children demonstrate high levels of self-generation (and thus, integration) even when they were required to identify relevant material from among a larger pool of potentially relevant information.

Consideration of performance across open-ended and forced-choice testing indicated the latent potential for integration of separate passages of text among 4-year-olds in the more information-constrained 2-stem condition and among 6-year-olds in the less information-constrained 2-stem plus related condition. When given the opportunity to select the novel integration facts from among distractors, fully 50% of 4-year-olds capitalized on the opportunity in the more information-constrained condition. In contrast, in the less information-constrained condition, forced-choice selection performance was low and did not differ from that in the control condition in which there was no opportunity for integration (since only one of the two relevant text passages was presented). We interpret this pattern to suggest that the 4-year-olds did not recognize the opportunity for integration of related passages of text when they were presented among potential competitors for integration. In contrast, the 6-year-olds capitalized on the opportunity to select the novel integration facts from among other potentially relevant material, such that across open-ended and forced-choice testing, fully 81% of the 6-year-olds self-generated or selected the novel integration facts in the less information-constrained 2-stem plus related condition.

Substantial additional research will be required to fully understand the process involved in correct forced-choice selection of a response alternative that was never previously experienced. Although seemingly the most obvious candidate is the process of recognition, we suggest that characterization is not accurate, owing to the fact that the correct alternative had not previously been experienced and thus could not, technically, be recognized. It is reasonable to speculate that conceptual priming played some role, given that the correct choices had words in common with the stem-fact text passages, whereas the incorrect choices did not. For example, the response options for the question, "how does a pod talk?" were (a) by clicking and squeaking, (b) by splashing and jumping, and (c) by rubbing noses. Yet conceptual priming alone cannot explain performance because the same options were provided in all three conditions, across which performance differed systematically. Moreover, in an event-related potential study with adults, Bauer and Jackson (2015) demonstrated that conceptual priming alone could not account for differential responses to facts derived through integration (which were treated as "well-known") and equally novel facts that were not derived in this self-generative manner (which were treated as "novel").

Based on the results of the present research, we suggest that the correct choice alternative functioned as an implicit hint or cue to the children that they possessed information relevant to the questions being posed. In prior related research, delivering the hint to "think about the stories we just read" just before presenting integration-fact questions significantly increased 4- and 6-year-old children's levels of self-generation performance (Bauer et al., 2012; Bauer, Varga, et al., 2015). In the present research, the forced-choice alternative "hints" apparently suggested different things to the 4- and 6-year-olds. The correct alternative seemingly suggested to 4-year-olds that they had information that could be integrated to answer the question. In the case of the more information-constrained 2-stem condition, there were two—and only two—passages of text to consider. In that case, they were successful. In the case of the less information-constrained 2-stem plus related condition, there were four potential passages to be considered. Under that circumstance, they were not successful—the hint was not sufficient to highlight the relevant passages among the competitors. For 6-year-olds, the correct alternative seemingly suggested which specific information they could integrate to answer the question, thus enabling them to respond with high levels of accuracy.

The suggestion that developmental differences in productive self-generation through integration are due at least in part to increases with age in the probability of identification of episodes of experience that are relevant for integration has potential implications for the way information is presented in formal and informal educational settings. Specifically, the less information-constrained 2-stem plus related condition resembles the manner in which children experience information in their everyday lives in both deliberate and incidental learning situations. Standard nonfiction texts, as well as museum exhibits, for example (Jant et al., 2014), feature many facts about the subject domain. Some of the information can be used to support productive extension beyond the material directly provided. Yet not all information on a given topic can or should be integrated and used as the basis for productive extension—some of it may even be contradictory. Under these circumstances, absent hints, cues, or explicit statements, younger children in particular may be challenged to navigate the body of information in order to make productive connections among individual items. Beyond hints or cues, when multiple facts within a given conceptual domain are presented at once, it may be beneficial to increase the surface similarity between items that are intended to be integrated with one another, and decrease the surface similarity with items that should not be integrated. As in the larger literature on productive extension of knowledge (e.g., Goswami, 2011), in the case of self-generation of new knowledge through integration, high surface similarity facilitates performance relative to low surface similarity (Bauer et al., 2012). Thus, the degree of surface similarity among items can serve to support integration or segregation, as appropriate.

Another means of promoting productive extension of knowledge under domain-information-rich conditions may be to create the demand for integration of appropriate material before moving on to material not intended for integration. In the present and prior related research, the demand for integration came in the form of an integration-fact question. Children who successfully self-generate novel integration facts in response to a demand remember them over delays of at least 1 week (Varga & Bauer, 2013; Varga, Stewart, & Bauer, 2016). The act of productive extension seemingly inoculates against forgetting and interference. These and other suggestions await direct test. Such a test would be maximally informative in an educational setting.

Importantly, variance in the probability of identification of episodes of experience that is relevant for integration is not the only source of variability in self-generation of new knowledge through integration of separate episodes. Variability in performance also can be accounted for in part by differential recall of the to-be-integrated stem facts. In Experiment 2 of Bauer and San Souci (2010), 4-year-old children were brought to a criterion level of learning of the stem facts prior to test for self-generation through integration. With enhanced recall of the stem-fact information, their self-generation performance increased, from 13% without criterion learning (Experiment 1) to 33% with criterion learning (Experiment 2). Although learning stem facts to a criterion resulted in higher levels of performance for 4-year-old children, their performance still was not as high as that of 6-year-old children (67% self-generation). Thus, increase in the accessibility of the stem facts to memory reduced, but did not eliminate, age-related differences. As noted earlier, differences in recall of the stem facts also cannot explain the variability in performance in the more and less information-constrained conditions in Experiment 1 of the present research. Both 4- and 6-year-old children had comparable levels of recall of the stem facts in the less information-constrained conditions relative to the more information-constrained conditions, yet they had lower levels of self-generation. In future research, it would be desirable to evaluate the joint and shared contributions to age-related variance of the requirements to identify relevant information and recall the information. It also would be desirable to explore the potential contributions of other sources of individual variability, such as executive function, for example.

Even as it informs basic process and shows promising implications for applied settings, the present research had certain limitations. One limitation is that we were able to administer only one test trial per condition. The story passage context in which the work was conducted has the benefits of being engaging for young children and of making the processes of integration and self-generation truly incidental. Yet the requirement that the stem facts (and nonstem facts) be presented in the context of stories about characters who have experiences places limits on the number of trials that can be administered in a given unit of time. In future research, it would be beneficial to develop alternative paradigms that permit administration of a larger number of test trials. Steps in this direction were taken in Bauer, Blue, Xu, and Esposito (2016), in which 7- to 10-year-old children played a board “game” during which they learned a number of novel stem facts. They later were tested on up to 10 trials of self-generation based on integration of the stem facts.

A second limitation of the present research was that the 8-year-old children in Experiment 2 were virtually at ceiling levels of performance, even in open-ended testing. Importantly, the 8-year-olds’ performance in the 1-stem plus related control condition made clear that their high levels of performance were not due to prior knowledge of the novel integration facts. In open-ended testing, none of the children produced the novel integration facts when given only one member of the pair of related stem facts. Eight-year-olds also produced very few of the novel integration facts in response to questions from the domains to which they were either not exposed (2-stem plus related and 1-stem plus related conditions) or were minimally exposed (2-stem condition). That is, all children were asked integration fact questions from all three of the domains (dolphins, palm trees, deserts). Thus, in addition to the data reported above, the 8-year-olds had 78 no- or minimal-exposure trials (39 participants $\times$ 2 trials per participant) on which they could, in principle, have produced the novel integration facts. They produced the integration facts on 10 of the 78 possible trials. This represents 13% of the trials on which they could not have generated the novel facts as a result of integration of a pair of stem facts. In contrast, in the 2-stem and 2-stem plus related conditions, the percentages of trials on which the 8-year-olds generated the integration facts were 75% and 87%, respectively. These data make clear that generation of the novel integration facts in the absence of exposure to both members of a stem fact pair was a low-frequency event. They thus provide additional support for the argument that the task provided a valid measure of self-generation of new knowledge through integration, even for the 8-year-olds. In future research, efforts should be made to develop a more sensitive test, thereby permitting examination of factors that may further facilitate older children's performance.

In conclusion, the present research allowed us to elucidate a developmental progression in children's productive extension of new knowledge under more and less domain-information-rich conditions; the information conditions were experimentally manipulated, thus affording substantial control. The work also informed one of the sources of age-related increases in productive knowledge extension, namely, increases with age in the probability of identification of episodes of experience that are relevant for integration. At least under the conditions of the present research, 4-year-old children seemingly require the equivalent of implicit hints or cues to the potential for integration in order to engage in this step of productive extension. Six-year-old children engage in identification and integrative processes under conditions of high certainty as to which materials are relevant for integration. Under less certain conditions, they too depend on the equivalent of implicit hints or cues. By 8 years of age, children readily engage in identification, integration, and self-generation, both when it is relatively clear what should be integrated with what, and when the circumstances are more ambiguous. The work thus informs both the conditions under which children 4–8 years of age extend their knowledge through integration of separate episodes and the cognitive processes involved. It also sheds light on the sources of age-related change in this productive process.

References

Bauer, P. J. (2009). Neurodevelopmental changes in infancy and beyond: Implications for learning and memory. In O. A. Barbarin & B. H. Wasik (Eds.), *Handbook of child development and early education: Research to practice* (pp. 78–102). New York, NY: Guilford.

Bauer, P. J. (2012). Facilitating learning and memory in infants and young children: Mechanisms and methods. In S. L. Odom, E. P. Pungello, & N. Gardner-Neblett (Eds.), *Infants, toddlers, and families in poverty: Research implications for early child care* (pp. 49–74). New York, NY: Guilford.

Bauer, P. J., Blue, S. N., Xu, A., & Esposito, A. G. (2016). *Productive extension of semantic memory in school-age children: Relations with reading comprehension and deployment of cognitive resources*. Manuscript under review. doi:10.1037/dev0000130

Bauer, P. J., & Jackson, F. L. (2015). Semantic elaboration: ERPs reveal rapid transition from novel to known. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 271–282. doi:10.1037/a0037405

Bauer, P. J., King, J. E., Larkina, M., Varga, N. L., & White, E. A. (2012). Characters and clues: Factors affecting children's extension of knowledge through integration of separate episodes. *Journal of Experimental Child Psychology*, *111*, 681–694. doi:10.1016/j.jecp.2011.10.005

Bauer, P. J., & San Souci, P. (2010). Going beyond the facts: Young children extend knowledge by integrating episodes. *Journal of Experimental Child Psychology*, *107*, 452–465. doi:10.1016/j.jecp.2010.05.012

Bauer, P. J., & Varga, N. L. (2015). The developmental cognitive neuroscience of memory: Implications for education. In E. Tardif & P. Doudin (Eds.), *Collective work on the topics of neuroscience, cognition and education* (pp. 1–16). Oxford, UK: De Boeck.

Bauer, P. J., Varga, N. L., King, J. E., Nolen, A. M., & White, E. A. (2015). Semantic elaboration through integration: Hints both facilitate and inform the process. *Journal of Cognition and Development*, *16*, 351–369. doi:10.1080/15248372.2013.849707

Brown, A. L., & Kane, M. J. (1988). Preschool children can learn to transfer: Learning to learn and learning by example. *Cognitive Psychology*, *20*, 493–523. doi:10.1016/0010-0285(88)90014-x

Brown, A. L., Kane, M. J., & Echols, C. H. (1986). Young children's mental models determine analogical transfer across problems with a common goal structure. *Cognitive Development*, *1*, 103–121. doi:10.1016/S0885-2014(86)80014-4

Crisafi, M. A., & Brown, A. L. (1986). Analogical transfer in very young children: Combining two separately learned solutions to reach a goal. *Child Development*, *57*, 953–968. doi:10.2307/1130371

Dias, M. G., & Harris, P. L. (1988). The effect of make-believe play on deductive reasoning. *British Journal of Developmental Psychology*, *6*, 207–221. doi:10.1111/j.2044-835x.1988.tb01095.x

Gelman, S. A. (2009). Learning from others: Children's construction of concepts. *Annual Review of Psychology*, *60*, 115–140.

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development*, *58*, 1532–1541. doi:10.2307/11303693

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*, 155–170. doi:10.1016/b978-1-4832-1446-7.50026-1

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199–241). New York, NY: Cambridge University Press.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science*, *10*, 277–300. doi:10.1207/s15516709cog1003_2

Ghetti, S., Goodman, G. S., Eisen, M. L., Qin, J., & Davis, S. L. (2002). Consistency in children's reports of sexual and physical abuse. *Child Abuse and Neglect*, *26*, 977–995. doi: 10.1016/s0145-2134(02)00367-8

Gluckman, M., Vlach, H. A., & Sandhofer, C. M. (2014). Spacing simultaneously promotes multiple forms of learning in children's science curriculum. *Applied Cognitive Psychology*, *28*, 266–273. doi:10.1002/acp.2997

Goswami, U. (2011). Inductive and deductive reasoning. In U. Goswami (Ed.), *Childhood cognitive development* (pp. 399–419). Oxford, UK: Wiley-Blackwell. doi:10.1016/j.tics.2010.10.001

Holyoak, K., Junn, E. N., & Billman, D. O. (1984). Development of analogical problem-solving skill. *Child Development*, *55*, 2042–2055. doi:10.2307/1129778

Jant, E. A., Haden, C. A., Uttal, D. H., & Babcock, E. (2014). Conversation and object manipulation influence children's learning in a museum. *Child Development*, *85*, 2029–2045. doi:10.1111/cdev.12252

Neisser, U., & Harsch, N. (1992). Phantom flashbulbs: False recollections of hearing the news about Challenger. In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall: Studies of "flashbulb" memories* (pp. 9–31). Cambridge, UK: Cambridge University Press.

Varga, N. L., & Bauer, P. J. (2013). Effects of delay on 6-year-old children's self-generation and retention of knowledge through integration. *Journal of Experimental Child Psychology*, *115*, 326–341. doi:10.1016/j.jecp.2013.01.008

Varga, N. L., Stewart, A., & Bauer, P. J. (2016). Integrating across episodes: Investigating the long-term accessibility of self-derived knowledge in 4-year-old children. *Journal of Experimental Child Psychology*, *145*, 48–63.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson—III test of cognitive abilities*. Itasca, IL: Riverside.