Distribution Agreement

Tristan S. Yates April 10, 2018

Call to Order: Four-Year-Olds Exhibit Difficulty Self-Deriving Knowledge through Integration of Ordered Facts

by

Tristan S. Yates

Patricia J. Bauer

Adviser

Neuroscience and Behavioral Biology

Patricia J. Bauer

Adviser

Leah Roesch

Committee Member

Kim Loudermilk

Committee Member

2018

Call to Order: Four-Year-Olds Exhibit Difficulty Self-Deriving Knowledge through Integration of Ordered Facts

By

Tristan S. Yates

Patricia J. Bauer

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Sciences with Honors

Neuroscience and Behavioral Biology

2018

Abstract

Call to Order: Four-Year-Olds Exhibit Difficulty Self-Deriving Knowledge through Integration of Ordered Facts

By Tristan S. Yates

The productive extension of knowledge through the integration of semantic facts acquired in separate but related episodes of new learning is crucial in the building of one's knowledge base. The association between self-derivation through integration (SDI) and academic achievement in both children and adults calls for further examination of the factors that influence young children's SDI performance. The current investigation was conducted to explore one factor that may affect children's ability to integrate novel facts: the relations between the to-be-integrated facts (hereafter, "stem facts"). Four-year-olds were assigned to one of two conditions. Children were either presented with facts that required hierarchical ordering for integration (linear condition; A>B, B>C, therefore A>C) or facts that did not require hierarchical ordering (nonlinear condition; A = B, B = C, therefore A = C). In both conditions, children demonstrated low SDI performance in an open-ended format, consistent with previous studies (e.g., Bauer & San Souci, 2010). In a forced-choice test, children were significantly more likely to select the correct integration fact in the nonlinear condition. Differences in memory for the stem facts in the linear and nonlinear conditions may help explain the difference in integration performance: whereas stem fact memory was greater than chance in the nonlinear condition, in the linear condition, performance did not differ from chance. Together, these results suggest that young children may have particular difficulty self-deriving new knowledge when integrating across ordered relative to non-ordered facts; the difference may be due to the nature of the relations or to differential stem fact memory.

Keywords: memory integration, knowledge extension, relational reasoning, transitivity

Call to Order: Four-Year-Olds Exhibit Difficulty Self-Deriving Knowledge through Integration of Ordered Facts

By

Tristan S. Yates

Patricia J. Bauer

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Sciences with Honors

Neuroscience and Behavioral Biology

2018

Acknowledgements

I would like to acknowledge the members of the Bauer Memory Development Lab at Emory University for all of their guidance throughout this past year. Thank you to Jillian Lauer for playing a large role in helping me with this research, and for providing me with continuous support throughout the thesis process. Finally, I would also like to acknowledge my committee members for their time and assistance, and the participants and their families for taking part in this study.

Table of Contents

Call to Order: Four-Year-Olds Exhibit Difficulty Self-Deriving Knowledge through Integration of Ordered Facts

The ability to combine novel pieces of information (e.g., "The father of evolution began his career as a creationist") with items stored in memory (e.g., "Charles Darwin was the father of evolution") to derive new knowledge (e.g., "Charles Darwin began his career as a creationist") is advantageous for learning within a complex environment. The self-derivation of new knowledge via the integration of separate but related episodes of new learning allows us to expand our knowledge base through the flexible recombination of information previously stored in memory (see Figure 1). Both children and adults demonstrate the ability to self-derive new knowledge through integration, but this ability shows substantial developmental change across childhood and is particularly challenging for 4-year-olds (e.g., Bauer & San Souci, 2010; Bauer & Larkina, 2016). Manipulations in the presentation of facts are known to influence 4-year-olds' ability to self-derive new knowledge, but how ordered relations between the to-be-integrated facts themselves influence the acquisition of new knowledge has not yet been investigated. In the present research, we examined how ordered relations and the role of logical reasoning may impact young children's knowledge integration by comparing performance between two conditions that differed by the relational terms used to link the facts. Specifically, children were tasked with integrating novel facts when successful self-derivation performance required substituting out a common element in non-ordered relations, or hierarchically ordering the facts—the latter of which may or may not be particularly difficult for young children (e.g., Piaget, 1964; Piaget, 1999; Wright & Smailes, 2015).

Implications of Knowledge Integration for Academic Achievement

The ability to successfully integrate information acquired at different points in time has been shown to facilitate spatial navigation, inference formations, creativity, and imaginative thinking (Schlichting & Preston, 2015). Furthermore, the self-derivation of new knowledge through integration has significant implications for academic achievement, because it enables the possibility of expanding the knowledge base from which individuals draw upon to solve problems and arrive at new conclusions. In the knowledge integration paradigm developed by Bauer and San Souci (2010), participants are exposed to novel, semantic facts—hereafter referred to as "stem facts"—that can be integrated together to create new knowledge. After learning the stem facts, individuals are tested on their ability to self-derive the novel integration fact. The semantic facts used in this paradigm are similar to those acquired in the classroom, and the questions used to assess the self-derivation of novel integration facts parallel the types of questions often probed in educational examinations. Together, these elements support the similarity between self-derivation through knowledge integration and the building of a knowledge base rapidly and efficiently for use in academic environments. Indeed, in young adults, the ability to self-derive new knowledge through integration is positively associated with academic outcomes, namely SAT scores and college GPA (Varga, Esposito, & Bauer, under review). In school-aged children, self-derivation performance similarly relates to standardized end-of-grade reading comprehension and math scores (Esposito & Bauer, 2017; Varga, Esposito, & Bauer, under review). Thus, research advancing our understanding of the factors that enable children to successfully integrate novel information may ultimately inform interventions aimed at promoting educational success.

Knowledge Integration in Preschool-Aged Children

Given the association between knowledge integration and academic achievement, it is important to investigate the factors that influence knowledge integration in children at the age when they begin formal schooling. In the first study examining the self-derivation of new knowledge in preschool-aged children, Bauer and San Souci (2010) found that 4-year-olds, on average, performed significantly above chance on a forced-choice test of self-derivation ability. During the experiment, children were presented with two story passages, each of which contained a novel fact about dolphins, kangaroos, or volcanoes. In the "two-stem" condition, the facts presented in the two different story passages about the same subject (e.g., "Dolphins live in groups called pods" and "Dolphins talk by clicking and squeaking") and could be combined to derive the novel integration fact (e.g. "pods talk by click and squeaking"). In the self-derivation test, children were first asked to provide an open-ended answer to a question that required integrating these two stem facts (e.g., "How does a pod talk?"). Importantly, children who did not hear both of the stem facts were not successful at self-deriving the novel integration fact. Six-year-olds exposed to both stem facts supplied the novel integration fact 67% of the time, whereas 4-year-olds exposed to both stem facts only supplied the novel integration fact 13% of the time. Although 4-year-olds experienced difficulty in providing the novel integration fact in this open-ended format, they were able to select the correct answer ("by clicking and squeaking") in a multiple-choice format over the two distractor items 67% of the time. Four-year-olds therefore required more support in order to successfully self-derive new knowledge. Nevertheless, 4-year-olds' performance remained noticeably below that of 6-year-olds, who selected the correct answer on 80% of trials. Given these results, the authors concluded that children's ability to self-derive new knowledge, when tested in open-ended and forced-choice formats, exhibits significant developmental gains between 4 and 6 years of age.

Although self-derivation performance is lower in 4-year-olds than in older children and adults, a number of manipulations have been shown to enhance 4-year-olds' self-derivation performance. For instance, learning the stem facts to criterion allowed more 4-year-olds (33% vs. 13%) to integrate in open-ended format (Bauer & San Souci, 2010). Although 4-year-olds' open-ended self-derivation performance was still considerably lower than that of 6-year-olds, their overall self-derivation performance, which combined their forced-choice self-derivation performance with their open-ended performance, was comparable to that of 6-year-olds (87% vs. 93% correct, respectively). Learning the stem facts to criterion therefore assisted in the acquisition of new knowledge in younger children, and importantly demonstrated that modifications in the knowledge integration paradigm can facilitate self-derivation performance in younger children.

Subsequent research has indicated that task-specific hints also enhance 4-year-olds' self-derivation through integration performance. In one study, Bauer and colleagues (2015) presented children with to-be-integrated facts that were differentially emphasized as being related to one another. In a "between-stem" hint condition, children were told to think about the first story read to them prior to hearing the second story, whereas children in a "before-test" hint condition were told to think about both stories before answering the open-ended integration question. Four-year-olds in the before-test condition exhibited significantly higher self-derivation performance compared to those in the control condition (who were simply told to "think" before answering the question), and children in the between-stem condition demonstrated an intermediate level of performance between the before-test condition and the control condition. Building on these findings, Varga and colleagues (2016) demonstrated that hinting to the relevance of stem facts also facilitated 4-year-olds' self-derivation performance across a delay period. Specifically, 4-year-olds provided with these task-specific hints successfully self-derived the novel integration facts on 60% of trials following a one-week delay. Moreover, reminding children of the stem facts that they learned prior to testing for integration during the second session allowed some children who were previously unable to self-derive the novel integration facts during the first session to do so during the second (Varga et al., 2016). Therefore, task-specific hints about the relatedness of the stem facts appear to aid in both the initial demonstration and the retention of new knowledge derived through integration in this age group.

The Role of Ordered Relations and Transitive Reasoning

The manipulations described thus far have focused on the conditions of self-derivation testing and not on the facts required for knowledge integration themselves. It is possible that the nature of the presented material may influence children's ability to self-derive new knowledge through the integration of novel facts learned across episodes, and that 4-year-olds in particular are limited by their ability to engage in different forms of logical reasoning based on the ordered relations between novel facts (Halford, 1993). Because semantic facts come in a variety of relational forms, it is worthwhile to investigate how ordered relations between novel facts relate to self-derivation performance in young children. Relational reasoning in particular is positively associated with current and later mathematical proficiency in school-aged children (Singley & Bunge, 2014). Therefore, characterizing the extent to which children are able integrate facts linked by different relational terms may ultimately add to our understanding of the association between self-derivation through knowledge integration and later academic achievement, with particular relevance to success in the sciences and mathematics.

In previous knowledge integration experiments, the subject and object of a novel fact could often be substituted without changing the meaning of the fact (e.g., "Dolphins live in groups called pods / A pod is a group of dolphins"). Sometimes, however, the order of relations is important to derive the appropriate conclusion. For instance, integration of the two facts "Mary is taller than Bob; Joe is taller than Bob" yields the conclusion that "Mary and Joe are taller than Bob." However, re-ordering the two individuals cited in the second fact causes the stem fact pair to be "Mary is taller than Bob; Bob is taller than Joe," which changes the conclusion one could derive. Acquisition of the novel integration fact ("Mary is taller than Joe") may now require the learner to perform a type of logical reasoning known as transitive inference to transfer the relation "taller than" from the first fact to the second fact. Children's capacity to perform this and other underlying forms of reasoning may factor into their ability to successfully self-derive new knowledge.

To date, there is no consensus in the extant literature regarding the ease at which young children perform transitive inference. Under some conditions, children younger than 6 or 7 years of age do appear to demonstrate transitive inference (Bryant & Trabasso, 1971; Goswami, 1995; Mustafchieva, Gotseva, & Kokinov, 2012), but under other conditions, they do not (Smedslund, 1963; Wright, Robertson, & Hadfield, 2011). The earliest studies predicted that 4-year-olds would not be able to perform transitive inference since later-developing formal logical operations were deemed essential for successful transitivity (Piaget, 1964; Piaget, 1999). Yet, nonhuman primates, rats, pigeons, hens, and even fish demonstrate evidence of transitive inference (Vasconcelos, 2008). Furthermore, recent research has indicated that transitive inference in the social domain can be seen in infants less than one year of age (Gazes, Hampton, & Lourenco, 2017). Based on these findings, some researchers have argued that transitive inference may depend on visuospatial reasoning and the mapping of items onto an organized spatial array (Gazes, Lazarea, Bergene, & Hampton, 2014). This would lead to the prediction that young children are able to demonstrate transitive inference, and may therefore be able to employ it when self-deriving new knowledge. Given the conflicting evidence over the ease at which young children perform transitive inference, assessing the impact of ordered relations on the self-derivation of new knowledge will be particularly informative to young children's self-derivation performance.

When the need to perform transitive inference with hierarchically ordered items is implied in facts or premises, individuals may construct mental models that utilize ordered representations of the facts to successfully integrate across pairs. Such hierarchical representations may be distinguishable from those required to perform integration across semantic facts that share common elements but are not hierarchically structured. If spatial reasoning is uniquely necessary for knowledge integration of ordered transitive relations, we would expect a positive association between self-derivation through knowledge integration performance and measures of mathematical achievement (Gazes, Lazareva, Bergene, & Hampton, 2014; Wright & Smailes, 2015). Research has also hinted that verbal abilities are sometimes important for transitive reasoning (Hummel & Holyoak, 2001; Prado, Mutreja, & Booth, 2013). Therefore, it is worth investigating how the integration of transitive versus other forms of relations relates to verbal and mathematical achievement.

Alternatively, rather than relying on a specific cognitive ability, the self-derivation of new knowledge through the integration of transitive relations may involve greater executive functioning compared to other relational forms, meaning that the self-derivation of ordered relations may be more correlated with measures of general intelligence such as working memory

(Halford, Wilson, & Phillips, 2010; Weldelken & Bunge, 2012). Researchers have posited that working memory capacity and complex forms of reasoning share similar capacity limitations, and that transitive reasoning in particular is subject to high processing loads (Halford, Cowan, and Andrews, 2007). Finally, it is also possible that knowledge integration is no more difficult for children when transitive relations are or are not presented, and that by using verbally explicit pairs of relations that are semantically rich, children will demonstrate similar self-derivation performance regardless of the ordered relations linking facts. Because some researchers argue that transitive inference is a later-developing logical process while others posit that it is evolutionarily and ontogenetically conserved, the current research was conducted to assess how this form of reasoning may be one factor that impacts the self-derivation of new knowledge. Thus, we were interested in how self-derivation through knowledge integration differs by stem fact relations, and whether different cognitive abilities relate to the successful self-derivation of new knowledge when hierarchical ordering is and is not necessary.

Importantly, when 4-year-olds do not learn transitive inference premises to criterion, they can still perform transitive inference when provided with a well-known example or a physical representation of transitive relations for comparison—in other words, 4-year-olds's ability to combine ordered transitive relations seems to be aided by stimuli that have real-world applications, in contrast to the abstract stimuli favored in the literature (e.g., Piaget, 1964; Smedslund, 1963; Wright et al., 2011). The knowledge integration paradigm is therefore the ideal learning environment to test for the integration of ordered relations in young children because of its application to real world knowledge acquisition and the context it provides children. By scaffolding abstract concepts within an episodic framework, the knowledge integration paradigm provides an avenue through which we can explore the impact of transitive and non-ordered relations on the acquisition of new knowledge in young children, with relevance for later academic achievement.

Research Objective

We investigated whether the structure of presented facts impacted 4-year-olds' abilities to extend their knowledge by self-derivation through knowledge integration. Children were presented with novel information (e.g., "the heart is smaller than the lungs" or "the heart is pink like the lungs") that was embedded in short picture stories. The primary objective was to determine whether different relations among stem facts facilitate or hinder self-derivation performance. The second objective was to determine whether cognitive abilities, such as working memory, and measures of verbal, mathematical, and relational skills are associated with self-derivation performance when transitive inference may or may not be necessary for success. By characterizing individual differences that may account for successful self-derivation through knowledge integration over different types of relations amongst novel facts, this research will contribute to our current understanding of how the expansion of semantic knowledge may be impacted by different cognitive abilities.

Children were assigned to one of two conditions in which they were tasked with integrating two stem facts by combining facts based on non-ordered relations linked by a similar characteristic (nonlinear condition) or by ordered relations that may require performing transitive inference on a hierarchically ordered set (linear condition). The former served as a comparison condition, since previous research has demonstrated that preschoolers can integrate non-ordered information (e.g., Bauer & San Souci, 2010). Although transitive inference is usually demonstrated using five interlocking terms, the present experiment opted to present only three related terms in order to lessen children's memory load. Children's integration was additionally supported by sequentially presenting the story containing the first stem fact and the story containing the second stem fact without imposing a delay, repeating each stem fact twice in its respective story, and repeating the stories themselves twice. Stem facts were always presented in the appropriate order for integration (i.e. item A > item B was always presented before item B > item C) and visual representations of the terms were present during the encoding of facts to facilitate the creation of a spatially organized representation in the linear condition. A variety of relational terms (e.g., “has moons like,” or “is smarter than”) were investigated to minimize the influence of any one particular relation on performance. Each child in the nonlinear condition listened to three stories that contained a stem fact pair in which the common element linking the stem facts was a physical property of the items, a detail about the movement/growth of the items, or a category from which the items belonged (see Table 1). In the linear condition, these three stories contained facts that consisted of a spatial, temporal, or abstract relation. These steps were taken in order to specifically compare transitive reasoning more generally with integration that may not necessarily require hierarchical ordering during the acquisition of new knowledge through knowledge integration.

Methods

Participants

The sample consisted of 32 children (16 female) who participated at 4 years of age (M= 4.46 years, SD = 0.22). Half of the sample was pseudo-randomly assigned to the nonlinear condition, and the other half was assigned to the linear condition, with the constraint that gender was balanced between conditions (50% female in both conditions), and the mean age was approximately the same for each condition (M_{NL} = 4.51, SD_{NL} = 0.27; M_{LIN} = 4.42, SD_{LIN} = 0.15). Participants were recruited through a database of families who had previously expressed interest in participating in research at the Emory University Child Study Center. Based on caregivers’ reports, participants were 25% African American, 3% Asian or Pacific Islander, 56% White or Caucasian, and 6% multiracial. In addition, 13% of caregivers identified their children's ethnicity as Hispanic and/or Latinx; the remaining 87% indicated that their children were not Hispanic and/or Latinx. One caregiver did not provide information regarding race and ethnicity. Primary caregivers reported being highly educated: 49% had obtained a graduate degree, 36% had obtained a college degree, and 12% had at least some college and/or received a technical or associate's degree. 94% of children were already enrolled in formal schooling, and 16% of children had begun reading independently, based on parental report.

Caregivers provided written informed consent on behalf of their children. Children were given stickers throughout the study to maintain their motivation, and they received a small toy and t-shirt after completing the study. Caregivers were compensated with a $5 gift card at the end of the session. The Emory University Institutional Review Board approved all of the protocol and procedures for this experiment.

Materials and Measures

Stem Fact Stories. Children were presented with previously unknown stem facts, which were determined to be novel to this age group through pilot testing. Each stem fact was embedded in a four-page story (57-60 words total; 13-16 words per page) that involved popular children's characters (see Figure 2 for example). Stem facts were first presented on the third page of each story and repeated on the last page. Each stem fact could be integrated with another stem fact presented in a second story that involved the same children's character. Pilot testing confirmed that the successful self-derivation or selection of the novel integration fact required that children hear both of the stem facts. In the two conditions, story pairs were matched for content and characters with the exception of the actual facts presented on the last two pages of each story. Six total story pairs were divided into two story sets, and children were presented with one set of three stories for their condition (Table 1). Story sets were counterbalanced between participants, and the order of the story pairs within the set was also counterbalanced.

Cognitive Assessments. Children completed three subtests from the Woodcock-Johnson (WJ) III Tests of Achievement and Cognitive Abilities (Woodcock, McGrew, & Mather, 2001) that assesses various cognitive abilities. The WJ-Picture Vocabulary test evaluates verbal abilities and lexical knowledge by requiring that children recognize and produce the correct name for an object viewed on paper. Items become progressively more difficult as the child proceeds. The WJ-Applied Problems test evaluates math abilities and the construction of mental mathematics models through items on counting, basic arithmetic, and the application of math knowledge (e.g., telling time, understanding money). Items are read aloud with related pictures. Finally, the WJ-Auditory Working Memory test was administered to determine the influence of executive function and working memory capacity on knowledge integration. Items required that children repeat the words and numbers they hear with the words first and then the numbers in the same presented order. For all WJ subtests, children's performance was measured using standardized scores.

Children also completed the Test of Relational Concepts (TRC; Edmonston & Litchfield Thane, 1988), a standardized measure of relational abilities. During the task, children are presented with pictures depicting different relational concepts, including temporal (e.g., first, last), quantitative (e.g., many, least), dimensional (e.g., tallest, widest), spatial (e.g., left, right), and miscellaneous (e.g., same, different) relations. Children completed all 56 items in the Test of Relational Concepts, and their performance was measured via standardized scores.

Procedure

Participants visited the laboratory for a single one-hour testing session (see Figure 3 for a schematic of the procedure). Two female experimenters alternated testing. The author tested 66% of the participants. All sessions were recorded with a LogiTech Webcam and were reviewed by the experimenters to ensure protocol fidelity.

After arriving to the laboratory, children completed a coloring worksheet, which served as a warm up activity. The experimenter then began the session by administering the WJ-Picture Vocabulary Test. Children were given a sticker following completion of WJ-Picture Vocabulary and after each subsequent task, to maintain their motivation.

Next, children viewed visual representations of the first pair of stem fact stories on an ASUS touch-screen laptop while listening to an audio recording of the stories. The same female voice read all of the stories, and no words were displayed on the screen as the child was read the story. Story pages were presented via PowerPoint® and recordings began automatically when each page appeared. Experimenters turned each page of the story by swiping across the screen to proceed to the next slide.

In between the first and second story, the experimenter hinted at the relatedness of the two stories by emphasizing that both stories contained the same characters ("Let's read another story. This one is also about [character]!"). The integration test phase then immediately followed the second reading of the second story. Children were asked to tell the experimenter two things that satisfied the relation given in the story set (e.g., "Can you name two things that are smaller than the lungs?") to assess their ability to self-derive new knowledge through integration in an open-ended format. Children were asked to name two things since one item ("the heart") was provided as a stem fact, whereas the other item ("the kidneys") required that children integrate across stem facts. Prior to asking this open-ended question, the experimenter encouraged the child to "think about the two stories we just read."

For the forced-choice portion of the test phase, the experimenter introduced the child to two puppets that were identical except for their color (yellow and red). Children were taught that the two puppets never say the same thing at the same time. For instance, one puppet would correctly call the child her or his name, while the other puppet would call the child the name of the main character from the first story set. Children were asked to point to the puppet that was "telling the truth," and the experimenter emphasized correct answers, while correcting wrong answers. All children were asked two example questions and did not continue to the study questions until they understood the task.

During the actual test phase, one puppet would say the correct integration fact that could be derived from the two stories, while the other puppet would say an incorrect version of the integration fact. For the linear condition, the incorrect answer was an inverse of the correct relation (e.g., "the lungs are smaller than the kidneys" compared to the correct answer "the kidneys are smaller than the lungs"). However, the inverse is still considered correct in the nonlinear condition (e.g., "the kidneys are pink like the lungs" and "the lungs are pink like the kidneys"). Therefore, the incorrect answer in the nonlinear condition used one of the distractor words embedded in the second page of each story (e.g., "the lungs are pink like the spleen" compared to the correct statement "the lungs are pink like the kidneys"). Adding a negation to the incorrect answer (e.g., "the lungs are not pink like the kidneys") was hypothesized to more readily draw children's attention, since the relation of the incorrect answer would differ from the learned stem fact relations. Thus, because children heard the distractor words the same number of times as the stem facts in the stem fact stories, the incorrect integration fact in the nonlinear condition was a reversal of the correct integration fact, with the exception that it contained the distractor word (e.g., “spleen”) instead of the word found in the first stem fact (e.g, “kidneys”).

All children were asked the forced-choice integration question regardless of their success on the open-ended question in order to keep the session consistent for each participant. After the forced-choice integration question, children’s memory for each of the stem facts embedded in the stories was tested using the puppets. Here, the incorrect answer that was pitted against the correct stem fact always included the distractor word (e.g., “spleen”). The puppet that spoke first was counterbalanced across participants, and the “correct” puppet was randomized for each forced-choice question.

After reading and answering the integration questions for the first pair of stories, children completed the WJ-Applied Problems test. The second and third story pair followed the same procedures as the first, and the WJ-Auditory Working Memory test was administered in between the second and third story pairs. After the final story pair, children completed the Test of Relational Concepts. Because of the length of this test, children were given a short break and a sticker halfway through the task to maintain their motivation. At the conclusion of the study session, children were thanked for their participation and received their toy and t-shirt.

Data Analysis and Scoring

Children received a score of 1 for correctly providing the integration answer in an open-ended question format, and a score of 0 if they were unable to successfully provide the integration fact answer. No points were given if the child only provided the second stem fact answer in response to the open-ended integration question. The difference between children’s open-ended self-derivation performance in the two conditions was analyzed via independent sample *t*-tests. Forced-choice self-derivation performance was similarly analyzed via independent sample *t*-tests, as were stem fact recognition scores. Finally, correlation analyses were conducted to examine the associations between different cognitive abilities and children's self-derivation performance.

Results

Self-Derivation Performance

Across conditions, children provided the correct integration fact in an open-ended format an average of 13.54% of trials (*SD* = 22.17%; see Table 2), similar to previous studies examining open-ended self-derivation performance in 4-year-olds (Bauer & San Souci, 2010; Bauer & Larkina, 2017; Varga et al., 2016). Open-ended self-derivation performance in the nonlinear condition (*M* = 14.48%, *SD* = 20.97%) did not differ significantly from open-ended self-derivation performance in the linear condition (*M* = 12.50%, *SD* = 23.95%; $t(29) = -0.26$, $p = .795$ $d = 0.09$; see Figure 4a). These results indicate that 4-year-olds' difficulty in self-deriving new knowledge in an open-ended format was similar for children in the two conditions.

Children's average forced-choice self-derivation performance in the nonlinear condition (*M* = 70.83%; *SD* = 29.50%) was significantly higher than children's self-derivation performance in the linear condition (*M* = 45.83%; *SD* = 29.50%; $t(30) = -2.40$, $p = .023$, $d = 0.85$; Figure 4b). Moreover, children's performance was significantly greater than chance in the nonlinear condition [$t(15) = 2.82$, $p = .013$, $d = 0.71$], but not in the linear condition [$t(15) = -0.56$, p = .581, $d = -0.14$]. Together, these results indicate that children demonstrated the ability to self-derive new knowledge in a forced-choice format in the nonlinear condition, but not in the linear condition (for all relational terms and an explanation of self-derivation performance by fact domain, see Appendix A).

To accurately credit children who successfully self-derived new knowledge in an open-ended format but did not subsequently select the correct integration fact in forced-choice testing, analyses were also conducted using a combined score in which children who self-derived new knowledge in either format were given a score of 1 and children who did not exhibit self-derivation were given a score of 0. Children's forced-choice self-derivation performance in the nonlinear condition remained significantly different from chance (M = 72.92%; SD = 27.81%; $t(15) = 3.30$, $p = .005$, $d = 0.82$) and children's forced-choice self-derivation performance in the linear condition remained not significantly different from chance (M = 52.08%; SD = 32.13%; $t(15) = 0.26$, $p = .799$, $d = 0.06$). When analyzing combined scores, the difference in forced-choice self-derivation performance between the two conditions was no longer statistically significant [$t(29) = -1.96$, $p = .059$, $d = 0.69$]. Thus, although children displayed greater forced-choice self-derivation performance in the nonlinear condition in comparison to the linear condition, this difference was no longer significant when trials on which children self-derived in the open-ended format but did not subsequently select the correct integration answer in the forced-choice format also were included.

Stem Fact Performance

On average, children correctly recognized 66.15% (SD = 27.20%) of the stem facts out of a possible total of 6 stem facts in the forced-choice stem fact portion of the test phase (see Table 3). This value was revealed to be significantly different from the chance rate of 50% [$t(31) = 3.35$, $p = .002$, $d = .59$]. Nominally, children recognized a greater percentage of stem facts in the nonlinear condition (M = 71.88%, SD = 20.83%) than in the linear condition (M = 60.42%, SD = 32.13%; see Figure 5). In comparison to the chance rate of 50%, children's memory for the stem facts was significantly above chance in the nonlinear condition [$t(15) = 4.2$, $p = .001$, $d = 1.05$], but not in the linear condition [$t(15) = 1.30$, $p = .214$, $d = 0.32$]. However, unlike forced-choice self-derivation performance, the difference between the two conditions was not significant [$t(26) = -1.20$, $p = .242$, $d = 0.42$] . Thus, greater forced-choice self-derivation performance in the nonlinear condition may or may not be attributable to differences in children's memory for the individual stem facts (for results regarding first and second stem fact memory, see Appendix B).

Within-Child Self-Derivation and Stem Fact Performance

Children selected the correct integration fact in the forced-choice format in 79.17% (*SE* = 8.31%) of the trials in which they correctly identified both stem facts in the nonlinear condition and 50.00% (*SE* = 10.66%) of the trials in which they correctly identified both stem facts in the linear condition (see Table 4). In comparison, children selected the novel integration fact in 66.67% (*SE* = 10.28%) of the trials for which they correctly identified only one of the two stem facts in the nonlinear condition and 42.86% (*SE* = 13.22%) of those trials in the linear condition. Children only selected the novel integration fact in 33.33% (*SE* = 27.21%) of the trials in which they correctly selected neither of the stem facts in the nonlinear condition and 41.67% (*SE* = 14.23%) in the linear condition. The difference between the proportion of children who correctly selected the novel integration fact when both stem facts were identified in the nonlinear versus the linear condition was significant in a two-sample *z*-test [$Z = 2.08$, $p = .039$]. These results indicate that even when children in the linear condition correctly identified both stem facts, they were less likely to select the correct integration fact compared to children in the nonlinear condition.

Overall, greater stem fact memory was not significantly associated with greater forced-choice self-derivation performance in a Chi-Square Lambda test in the nonlinear [$\chi^2(2, 16) = 3.03$, $p = .220$, $V = 0.25$] and linear [$\chi^2(2, 16) = 0.29$, $p = .865$, $V = 0.08$] conditions. These results indicate that children's memory for the stem facts is not always fully predictive of their ability to self-derive new knowledge in a forced-choice format. Nonetheless, there was a statistically significant positive correlation between forced-choice self-derivation performance and stem fact recognition across the two conditions [$r = .43$, $p = .014$]. Therefore, children with better recognition of the stem facts tended to have better forced-choice self-derivation performance compared to children who did not correctly identify the stem facts.

Correlations between Self-Derivation and Cognitive Abilities

Importantly, children's cognitive ability scores did not significantly differ by condition for any of the cognitive assessments administered (*t*s < 1.18 *p*s > .260; see Table 5). These results confirm our assumption that children in each condition possessed similar cognitive abilities, and that children in one condition were not more highly performing than the other. Both within and across conditions, many of the cognitive assessments were significantly associated with one another (see Table 6). Unsurprisingly, this was especially true for the overall correlation between standardized scores on the WJ-Applied Problems and the Test of Relational Concepts [$r(30) = 0.75$, $p = 2.243e-06$]. The moderate to high correlations among these cognitive assessments may suggest that our measures reflect a common underlying construct or a more general measure of cognitive function. This idea is further explored in the discussion.

Across conditions, correlation analyses revealed that open-ended self-derivation performance was significantly correlated with all of the cognitive assessment measures (*r*(30)s > 0.38, *p*s < .033; see Table 7). In the nonlinear condition, there was a significant correlation between open-ended self-derivation performance and standardized scores on the WJ-Picture Vocabulary test [$r(14) = .50$, $p = .048$]. This association was not found for children in the linear condition [$r(14) = .28$, $p = .302$]. In the linear condition, there was a significant correlation between open-ended self-derivation performance and standardized scores on the WJ-Applied Problems [$r = .55$, $p = .025$], WJ-Auditory Working Memory [$r(14) = .63$, $p = .009$], and the Test of Relational Concepts [$r(14) = .59$, $p = .028$] but no such associations were found for children in the nonlinear condition [$r(14)$s < .48, ps > .067]. The largest difference in the correlation coefficients between conditions was found for the association between open-ended self-derivation performance and working memory, although this was not statistically significant ($Z = 0.75$, $p = .227$; see Figure 6). Together, these results suggest that open-ended self-derivation performance in the nonlinear condition may be more related to verbal abilities, whereas open-ended self-derivation performance in the linear condition may be more related to mathematical abilities, working memory, and relational knowledge.

Across conditions, correlation analyses revealed that forced-choice self-derivation performance was significantly associated with standardized scores on the WJ-Applied Problems ($r = .40$, $p = .023$; Table 7). Surprisingly, forced-choice self-derivation performance was significantly correlated with standardized scores on the WJ-Applied Problems test in the nonlinear condition [$r = .52$, $p = .039$] but not in the linear condition ($r = .18$, $p = .507$). However, there was not a statistically significant difference in the correlation coefficients for forced-choice self-derivation performance and WJ-Applied Problems scores between the two conditions [$Z = -1.01$, $p = .313$]. Furthermore, there were no statistically significant differences in the correlation coefficients for forced-choice self-derivation performance and any of the other cognitive assessments between the two conditions [Zs < 0.88, ps > .189]. Both across conditions and within each condition, there were no significant associations between forced-choice self-derivation performance and the WJ-Picture Vocabulary, WJ -Auditory Working Memory, or The Test of Relational Concepts [rs < .30, ps > .113]. Altogether, these results indicate that forced-choice self-derivation performance may not have been meaningfully associated with the cognitive abilities measured in this study.

Discussion

The present investigation both replicates and expands upon prior research regarding the self-derivation of new factual knowledge through integration in 4-year-olds. First, open-ended self-derivation performance was comparable to previous studies examining knowledge integration in 4-year-olds (14% in the current investigation; 13% in Bauer & San Souci, 2010). This was true for both conditions. Therefore, children's ability to integrate the new stimuli created for this study in an open-ended format was of a similar difficulty as in previous experiments.

Children displayed evidence of knowledge integration in a forced-choice format when integrating non-ordered relations via common element substitution, but performed significantly worse when integrating ordered relations via transitive reasoning. The observation that children's performance did not differ from chance in the linear condition indicates that something may have impaired their ability to self-derive new knowledge through integration when a hierarchical order was implied. Factors that we know hinder young children's self-derivation performance include the presence of irrelevant but related facts, and a diminished memory for the stem facts (Bauer & Larkina, 2016; Bauer et al., 2015). The current investigation did not utilize irrelevant facts, and importantly, children's memory for stem facts was not significantly different between the two conditions. However, it should be noted that in contrast to the nonlinear condition, forced-choice stem fact performance was not significantly above chance in the linear condition. This difference is a potential explanation for the difference in forced-choice self-derivation performance between the two conditions. Since the story content and fact presentations were held constant between the two conditions, one distinguishing factor between the conditions was the specific relational terms used to link the two facts.

Yet, another factor that may have contributed to differences in forced-choice self-derivation was the wording of the incorrect choices. For instance, it may be that children exhibited greater difficulty in choosing the correct stem fact in the linear condition because the incorrect version of the stem fact was a reversal of the correct fact. In the nonlinear condition, this was not an issue, since the incorrect version of the integration fact contained the distractor word, which may have acted as a stronger error signal. Thus, although efforts were made to create a forced-choice self-derivation test that was comparable across the two conditions, the potential effect of the study design cannot be ruled out as factoring into the results. Given the similarities between the two conditions in most respects, the observed difference in forced-choice self-derivation through knowledge integration may be a product of children's difficulty with transitive inference. Alternatively, children in the linear condition may have exhibited greater difficulty in choosing the correct integration fact because the incorrect version was more closely related to the novel integration fact. Thus, further research should be conducted to address this potential issue.

Clues as to the elements that may support successful knowledge integration through transitive reasoning as assessed in the current study are found in the significant positive correlations between open-ended self-derivation performance in the linear condition and scores on cognitive assessments that measured working memory, relational knowledge, and mathematical abilities. Although it may be tempting to interpret these correlations as suggestions of the specific cognitive processes involved in self-derivation through integration, in light of the pattern of significant correlations among children's standardized scores on the cognitive assessments for working memory, mathematical achievement, and relational knowledge, such an exercise would not be warranted. Instead of specific cognitive processes, the pattern of intercorrelation suggests that a more general cognitive function may be related to self-derivation performance. WJ-Applied Problems is assumed to measure children's mathematical concepts and problem-solving abilities, which likely require the relational knowledge and reasoning that is captured in the Test of Relational Concepts. Additionally, many of the WJ-Applied Problems test items can be accomplished mentally by using working memory processes (Wendling, Schrank, & Schmitt, 2007). For instance, children may need to store and maintain instructions from the experimenter while viewing accompanying pictures on the test booklet. The Test of Relational Concepts may similarly rely on the maintenance and manipulation of items in a child's working memory. Furthermore, language comprehension and lexical knowledge may be a common factor influencing scores on The Test of Relational Concepts, WJ-Picture Vocabulary, and the WJ-Applied Problems, where listening ability is necessary for successful performance. Thus, although we hypothesized that the cognitive assessments we utilized in this study would act as independent measures, it is possible that young children's scores on these assessments reflect a more general measure of cognitive function, such as fluid intelligence.

In the current research, we observed that open-ended self-derivation performance was differentially associated with cognitive abilities in the two conditions, yet there was not a significant condition difference in open-ended self-derivation performance. Four-year-olds are not as skilled at strategically utilizing prior knowledge to autonomously display knowledge integration, and our results may have been subject to a floor effect. Condition differences in open-ended self-derivation would require that young children provide the novel integration fact more often than they typically do. This study therefore collected a second measure of self-derivation performance that aided young children by having their answers narrowed down in a forced-choice format. Previous studies utilizing this paradigm have demonstrated that constraining young children's choices and having them select the correct integration fact out of distractors allows them to perform in a way that suggests successful knowledge integration has occurred (Bauer & San Souci, 2010). Because integration facts are internally derived and never actually encountered in the real world, the idea of "recall" of information versus "recognition" is not fully representative of open-ended versus forced-choice measures of self-derivation. Nonetheless, it is helpful to think of open-ended self-derivation as a more robust form of knowledge integration, requiring greater depth of processing and retrieval than forced-choice self-derivation, which may be more supported by familiarity or a "gist" representation (Reyna & Kiernan, 1994). While the majority of 4-year-olds experience difficulty with the former, previous research has shown that the latter is an acceptable measure for quantifying self-derivation performance in young children (Bauer & San Souci, 2010; Bauer & Larkina, 2016; Bauer et al., 2015; Varga et al., 2016).

Our results indicate that young children were less likely to self-derive new knowledge when transitive inference may have been used to integrate ordered relations, despite prior evidence of successful transitive inference in 4-year-olds (Bryant & Trabasso, 1971; Goswami, 1995; Mustafchieva, Gotseva, & Kokinov, 2012). It is possible that some measures of transitive reasoning tap into different representations than what was necessary for transitive reasoning in the current study (non-analytical and analytical; see Halford, Wilson, & Phillips, 2010). Rather than requiring more logical reasoning operations, transitive relations may be successfully integrated through associative learning, or the process by which two stimuli are learned to occur together. This alternative explanation assumes that transitive inference does not require explicit knowledge of the individual relations that are to be transferred, but instead relies on the construction of a unified representation that facilitates the retrieval of items in a sequence. Indeed, items farther apart in a transitive inference sequence elicit shorter response times and more accurate responses than items closer together in the sequence—a phenomenon known as the symbolic distance effect (Acuna, Sanes, & Donoghue, 2002). Ordering items into a linear series is argued to be the default solution to problems of transitivity because it is an efficient mechanism by which humans and nonhuman animals can extract novel relations (Riley & Trabasso, 1974). In the current study, individual relations were presented in different learning contexts, requiring that children first encode the relations separately before integrating across them. Therefore, there may be times when the integration of ordered relations may be accomplished through varying associative strengths, as in the case of certain nonhuman animal studies (Dusek & Eichenbaum, 1997; Vasconcelos, 2008), and other times when it may rely on logical transitive reasoning, as is speculated to be the case for the current study.

Although the current research suggests that young children are not as adept at applying transitive reasoning to the self-derivation of new knowledge through the integration of ordered relations when tested in this particular way, there is evidence that training preschoolers to understand serial logical operations encourages them to apply these operations in more complex ways (Pogozhina, 2014). Specifically, teaching children to solve problems that involved asymmetric transitive relations helped them apply their knowledge to more complex concepts such as the conservation principle, which is the ability to understand invariant properties such as mass, volume, or length. Logical training could similarly be applied to increase self-derivation through knowledge integration performance when certain types of logical reasoning are required. Young children with strengthened logical skills as boosted through training would be expected to experience an increased ability to assimilate novel information into their knowledge base, which in turn may have positive effects on their academic outcomes, particularly in subjects known to rely more heavily on relational and spatial reasoning.

Limitations

As noted, conclusions from this study are limited by the inability to completely match the two conditions. The nature of the relations in the nonlinear condition prevented the use of a reversal of the correct fact to create the incorrect version of the integration fact, although a reversal of the correct integration fact was used in the linear condition. It is possible that children in the linear condition were at a disadvantage given that the choices they had for the forced-choice integration question were more difficult to discriminate from one another. Therefore, in order to more definitively conclude whether or not a condition difference exists in forced-choice self-derivation performance, future research should take care to more closely match the choices used in the forced-choice portion of the test phase.

Another limitation of this study was the small sample size, which led to a diminished power, potentially inflating the observed effect. Additionally, the small sample size may have led to difficulty in detecting true correlations between self-derivation performance and cognitive abilities, particularly for forced-choice self-derivation performance. A larger sample size will therefore be necessary to form more reliable conclusions about differences between the two conditions. Additionally, although our sample came from a diverse database, a great majority of participants' caregivers were highly educated and identified with a common background. Therefore, these results may not extend to the wider population.

Future Directions

The current investigation utilized a between-subjects design, but future research should consider testing each child on stories from the two conditions in order to determine if the same abilities are recruited for successful self-derivation through knowledge integration performance when relations differ. This research could help elucidate what it is about ordered relations that seem to hinder young children's performance on the knowledge integration paradigm. Furthermore, extending this study beyond this particular age group will be necessary to see if condition differences persist across development. Once children become more skilled at logical reasoning in later childhood, they may demonstrate similar self-derivation performance regardless of the type of logical reasoning that may be required for the integration of ordered versus non-ordered relations (Halford, 1993). Alternatively, if older children still demonstrate poorer performance on the knowledge integration task when ordered relations are presented, the underlying mechanisms that enable knowledge integration may be highly dependent on the relations between the facts, and potentially, the complexity of reasoning required for successful self-derivation through knowledge integration.

Another line of work may investigate how learning logical premises to criterion impacts self-derivation performance. From the literature, we know that learning the individual stem facts to criterion increases self-derivation performance, and that training children to fully grasp transitive relations premises allows for transitivity (Bauer & San Souci, 2010; Bryant & Trabasso, 1971). Therefore, encouraging children to learn the stem facts to criterion may boost their self-derivation performance when ordered relations are presented. It would be interesting to note whether condition differences are still present when the premises are more fully encoded. If children still underperform in the condition that consists of ordered relations, this may signal that the underlying cognitive abilities play a significant role in children's ability to integrate ordered relations. On the other hand, if children perform equally well in both conditions, this may indicate that additional support is necessary for children to successfully self-derive new knowledge integration when ordered relations that may require transitive inference are utilized. However, if children's performance during the condition consisting of ordered transitive relations increases beyond that of a comparison condition, this could indicate that children may have utilized a different mechanism, such as evaluating associative strengths, to tackle the transitive inference problem and successfully self-derive new knowledge of ordered relations (Dusek & Eichenbaum, 1997). In any case, these results would add to our understanding of the extent to which ordered relations act as a factor influencing children's ability self-derive new knowledge.

Conclusions

The current research contributed to our understanding of 4-year-olds' ability to self-derive new knowledge when different relations link novel information. Knowledge integration of ordered relations that may require transitive inference appears to be more difficult for 4-year-olds than non-ordered relations that may require common-element substitution, signifying that ordered relations may be one factor that affects successful self-derivation performance in young children. From the current research, it is unclear whether 4-year-olds' difficulty with the self-derivation of new knowledge through integration of ordered facts is a result of the nature of the relations, the wording of the forced-choice integration question, or children's memory for the stem facts. Future research should expand upon this as a potential place for intervention in order to enable greater self-derivation through knowledge integration in young children and ultimately contribute to greater learning outcomes for children over time.

References

Acuna, B. D., Sanes, J. N., & Donoghue, J. P. (2002). Cognitive mechanisms of transitive inference. *Experimental Brain Research, 146*(1), 1-10. doi:10.1007/s00221-002-1092-y

Bauer, P. J., & Souci, P. S. (2010). Going beyond the facts: Young children extend knowledge by integrating episodes. *Journal of Experimental Child Psychology, 107*(4), 452-465. doi:10.1016/j.jecp.2010.05.012

Bauer, P. J., King, J. E., Larkina, M., Varga, N. L., & White, E. A. (2012). Characters and clues: Factors affecting children's extension of knowledge through integration of separate episodes. *Journal of Experimental Child Psychology, 111*(4), 68-694. doi:10.1016/j.jecp.2011.10.005

Bauer, P. J., Varga, N. L., King, J. E., Nolen, A. M., & White, E. A. (2015). Semantic elaboration through integration: Hints both facilitate and inform the process. *Journal of Cognition and Development, 16*(2), 351-369. doi:10.1080/15248372.2013.849707

Bauer, P. J., Blue, S. N., Xu, A., & Esposito, A. G. (2016). Productive extension of semantic memory in school-aged children: Relations with reading comprehension and deployment of cognitive resources. *Developmental Psychology, 52*(7), 1024-1037. doi:10.1037/dev0000130

Bauer, P. J., & Larkina, M. (2016). Realizing relevance: The influence of domain-specific information on generation of new knowledge through integration in 4- to 8-year-old children. *Child Development, 88*(1), 247-262. doi:10.1111/cdev.12584

Bornstein, M. H., & Arterberry, M. E. (2010). The development of object categorization in young children: Hierarchical inclusiveness, age, perceptual attribute, and group versus individual analyses. *Developmental Psychology, 46*(2), 350-365. doi:10.1037/a0018411

Brainerd, C. J., & Reyna, V. F. (1993). Memory independence and memory interference in cognitive development. *Psychological Review, 100*(1), 42-67. doi:10.1037/0033-295X.100.1.42

Bryant, P. E., & Trabasso, T. (1971). Transitive inferences and memory in young children. *Nature, 232*(5311) 456-458. doi:10.1038/232456a0.

Dusek, J. A., & Eichenbaum, H. (1997). The hippocampus and memory for orderly stimulus relations. *Proceedings of the National Academy of Sciences of the United States of America, 94*(13), 7109-7114. doi:10.1073/pnas.94.13.7109

Edmonston, N. K., & Litchfield Thane, N. (1988). *TRC: Test of Relational Concepts.* Austin, TX: Pro-Ed.

Esposito, A. G., & Bauer, P. J. (2017). Going beyond the lesson: Self-generating new factual knowledge in the classroom. *Journal of Experimental Child Psychology, 153*(1), 110-125. doi:10.1016/j.jecp.2016.09.003

Gazes, R. P., Lazareva, O. F., Bergene, C. N., & Hampton, R. R. (2014). Effects of spatial training on transitive inference performance in humans and rhesus monkeys. *Journal of Experimental Psychology: Animal Learning and Cognition, 40*(4), 477-489. doi:10.1037/xan0000038

Gazes, R. P., Hampton, R. R., & Lourenco, S. F. (2017). Transitive inference of social dominance by human infants. *Developmental science, 20*(2). doi:10.1111/desc.12367

Goswami, U. (1995). Transitive relational mappings in three- and four-year-olds: The analogy of goldilocks and the three bears. *Child Development, 66*(3), 877-892. doi:10.2307/1131956.

Halford, G. S. (1993). *Children's understanding: The development of mental models*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating Cognitive Capacity from Knowledge: A New Hypothesis. *Trends in Cognitive Sciences, 11*(6), 236–242. doi:10.1016/j.tics.2007.04.001

Halford, G. S., Wilson, W. H., & Phillips, S. (2010). Relational knowledge: The foundation of higher cognition. *Trends in Cognitive Sciences, 14*(11), 497-505. doi:10.1016/j.tics.2010.08.005

Hummel, J. E., & Holyoak, K. J. (2001). A process model of human transitive inference. In M. Gattis (Ed.). *Spatial Schemas in Abstract Thought* (pp. 279-305). Cambridge, MA: MIT Press.

Mutafchieva, M., Gotseva, K., & Kokinov, B. (2012, January). Does analogy facilitate transitive inference in young children? In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 34, No. 34). https://escholarship.org/uc/item/67c4g8v9

Piaget, J. (1964). Part I: Cognitive development in children: Piaget development and learning. *Journal of Research in Science Teaching 2*(1), 176-186. doi:10.1002/tea.3660020306

Piaget, J. (1999). *Judgement and reasoning in the child*. London, England: Routledge. doi:10.4324/9780203207260

Pogozhina, I. (2014). Development of the logical operations in preschool children. *Procedia-Social and Behavioral Sciences, 146*, 290-295. doi:10.1016/j.sbspro.2014.08.132

Prado, J., Mutreja, R., & Booth, R. J. (2013). Fractionating the neural substrates of transitive reasoning: Task-dependent contributions of spatial and verbal representations. *Cerebral Cortex, 23*(3), 499-507. doi:10.1093/cercor/bhr389

Reyna, V. F., & Kiernan, B. (1994). Development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval. *Developmental Psychology, 30*(2), 178-191. doi:10.1037/0012-1649.30.2.178

Riley, C. A., & Trabasso, T. (1974) Comparatives, logical structures, and encoding in a transitive inference task. *Journal of Experimental Psychology 17*(2), 187-203. doi:10.1016/0022-0965(74)90065-4

Smedslund, J. (1963). Development of concrete transitivity of length in children. *Child Development, 34*(2), 389-405. doi:10.2307/1126735

Schlichting, M. L., & Preston, A. R. (2015). Memory integration: Neural mechanisms and implications for behavior. *Current Opinion in Behavioral Sciences, 1*, 1-8. doi:10.1016/j.cobeha.2014.07.005

Siegler, R. S. (1989). Mechanisms of cognitive development. *Annual review of psychology, 40*(1), 353-379. doi:10.1146/annurev.ps.40.020189.002033

Singley, A. T. M., & Bunge, S. A. (2014). Neurodevelopment of relational reasoning: Implications for mathematical pedagogy. *Trends in Neuroscience and Education, 3*(2), 33-37. doi:10.1016/j.tine.2014.03.001

Sloutsky, V. M., & Fisher, A. V. (2004). Induction and Categorization in Young Children: A Similarity-Based Model. *Journal of Experimental Psychology: General, 133*(2), 166-188. doi:10.1037/0096-3445.133.2.166

Varga, N. L., Stewart, R. A., & Bauer, P. J. (2016). Integrating across episodes: Investigating the long-term accessibility of self-derived knowledge in 4-year-old children. *Journal of Experimental Child Psychology, 145*(1), 48-63. doi:10.1016/j.jecp.2015.11.015

Varga, N. L., & Bauer, P. J. (2017). Young adults self-derive and retain new factual knowledge through memory integration. *Memory & Cognition, 45*(6), 1014-1027. doi:10.3758/s13421-017-0711-6

Varga, N. L., Esposito, A. G., & Bauer, P. J. (under review). Cognitive correlates of memory integration across development: Explaining variability in an educationally relevant phenomenon.

Vasconcelos, M. (2008). Transitive inference in non-human animals: An empirical and theoretical analysis. *Behavioural Processes, 78*(3), 313-334. doi:10.1016/j.beproc.2008.02.017

Wendelken, C., & Bunge, S. A. (2010). Transitive inference: Distinct contributions of rostrolateral prefrontal cortex and the hippocampus. *Journal of Cognitive Neuroscience, 22*(5), 837-847. doi:10.1162/jocn.2009.21226

Wendling, B. J., Schrank, F. A., & Schmitt, A. J. (2007). *Educational Interventions Related to the Woodcock-Johnson III Tests of Achievement* (Assessment Service Bulletin No. 8). Rolling Meadows, IL: Riverside Publishing.

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

Wright, B. C., & Smailes, J. (2015). Factors and processes in children's transitive deductions. *Journal of Cognitive Psychology, 27*(8), 967-978. doi:10.1080/20445911.2015.1063641

Table 1

Example of Stem Facts and Integration Facts in Nonlinear Condition

Nonlinear Condition							
Story Set	**Domain**	**Relation**	**Stem Fact A**	**Stem Fact B**	**Integration Fact**	**Distractor word**	**Incorrect Integration**
X	Space	Physical property detail (Possessive quality)	*Jupiter has moons like Saturn*	*Saturn has moons like Neptune*	*Jupiter has moons like Neptune*	The sun	*Neptune has moons like the Sun*
X	Aviation	Motion detail (Movement process)	*Rockets fly up like hot-air balloons*	*Hot-air balloons fly up like helicopters*	*Rockets fly up like helicopters*	Fireworks	*Helicopters fly up like fireworks*
X	Birds	Category detail (Biological category)	*Ravens are aves like parrots*	*Parrots are aves like eagles*	*Ravens are aves like eagles*	Bats	*Eagles are aves like bats*
Y	Organs	Physical property detail (Color)	*The kidneys are pink like the heart*	*The heart is pink like the lungs*	*The kidneys are pink like the lungs*	The spleen	*The lungs are pink like the spleen*
Y	Fruits	Motion detail (Growth process)	*Peach trees grow from pits like cherry trees*	*Cherry trees grow from pits like coconut trees*	*Peach trees grow from pits like coconut trees*	Banana trees	*Coconut trees grow from pits like banana trees*
Y	Color Gems	Category detail (Non-biological category)	*Green gems are compounds like red gems*	*Red gems are compounds like blue gems*	*Green gems are compounds like blue gems*	Clear gems	*Blue gems are compounds like clear gems*

Example of Stem Facts and Integration Facts in Linear Condition

Linear Condition							
Story Set	**Domain**	**Relation**	**Stem Fact A**	**Stem Fact B**	**Integration Fact**	**Distractor word**	**Incorrect Integration**
X	Space	Spatial (Distance)	*Jupiter is closer than Saturn*	*Saturn is closer than Neptune*	*Jupiter is closer than Neptune*	The sun	*Neptune is closer than Jupiter*
X	Aviation	Temporal (Chronology)	*Rockets were made before hot-air balloons*	*Hot-air balloons were made before helicopters*	*Rockets were made before helicopters*	Fireworks	*Helicopters were made before rockets*
X	Birds	Abstract (Social)	*Ravens are smarter than parrots*	*Parrots are smarter than eagles*	*Ravens are smarter than eagles*	Bats	*Eagles are smarter than ravens*
Y	Organs	Spatial (Size)	*The kidneys are smaller than the heart*	*The heart is smaller than the lungs*	*The kidneys are smaller than the lungs*	The spleen	*The lungs are smaller than the kidneys*
Y	Fruits	Temporal (Speed)	*Peach trees grow faster than cherry trees*	*Cherry trees grow faster than coconut trees*	*Peach trees grow faster than coconut trees*	Banana trees	*Coconut trees grow faster than peach trees*
Y	Color Gems	Abstract (Value/worth)	*Green gems cost more than red gems*	*Red gems cost more than blue gems*	*Green gems cost more than blue gems*	Clear gems	*Blue gems cost more than green gems*

Table 2

Descriptive Statistics for Self-Derivation through Integration Performance by Condition

Condition	Open-Ended Performance *M (SD)*	Forced-Choice Performance *M (SD)*
Nonlinear	0.14 (0.21)	0.71 (0.30)
Linear	0.13 (0.24)	0.46 (0.30)
Overall	0.14 (0.22)	0.58 (0.32)

Note. Values represent proportion of integration items correctly provided or selected out of the total number of integration test items (max = 3).

Table 3

Descriptive Statistics for Forced-Choice Stem Fact Performance

	Overall	Stem Fact A	Stem Fact B
Condition	*M (SD)*	*M (SD)*	*M (SD)*
Nonlinear	0.72 (0.21)	0.69 (0.31)	0.75 (0.26)
Linear	0.60 (0.32)	0.56 (0.36)	0.65 (0.35)
Overall	0.66 (0.27)	0.63 (0.34)	0.70 (0.31)

Note. Values represent proportion of stem fact items correctly identified (six items overall; three for stem fact A and three for stem fact B).

Table 4

Children's Successful Forced-Choice Self-Derivation Performance by Stem Fact Recognition

	Stem Facts Recalled		
Condition	0 Facts	1 Fact	Both Facts
Nonlinear	0.33 (0.27)	0.67 (0.10)	0.79 (0.08)
Linear	0.42 (0.14)	0.43 (0.13)	0.50 (0.11)
Overall	0.40 (0.13)	0.57 (0.08)	0.65 (0.07)

Note. Values represent the proportion of children who successfully self-derived the novel integration fact in a forced-choice format based on the number of stem facts recalled for that particular integration fact. Standard error is presented in parentheses.

Table 5

Descriptive Statistics for Cognitive Assessments by Condition

	WJ-Picture Vocabulary	WJ-Applied Problems	WJ-Auditory Working Memory	Test of Relational Concepts (TRC)
Condition	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
Nonlinear	105.8 (8.30)	112.7 (15.29)	110.2 (13.90)	53.87 (9.66)
Linear	102.4 (8.88)	107.1 (13.55)	103.4 (16.76)	49.79 (9.42)
Overall	104.1 (8.63)	109.9 (14.49)	117.5 (15.51)	58.00 (9.60)

Note. Values represent standardized scores on the cognitive assessments. Two children did not complete the TRC, so TRC data was derived from only 28 of the children. For all other cognitive assessments, $N = 30$.

Table 6

Intercorrelations Among Cognitive Assessments Overall and By Condition

Cognitive Assessment	1. PicVocab	2. AppProb	3. AudWM	4. TRC
Overall (Both Conditions)				
1. Picture Vocabulary	-	0.49**	0.06	0.45*
2. Applied Problems		-	0.54**	0.75***
3. Auditory Working Memory			-	0.47**
4. Test of Relational Concepts				-

Cognitive Assessment	1. PicVocab	2. AppProb	3. AudWM	4. TRC
Nonlinear				
1. Picture Vocabulary	-	0.42	0.21	0.26
2. Applied Problems		-	0.64*	0.73**
3. Auditory Working Memory			-	0.59*
4. Test of Relational Concepts				-
Linear				
1. Picture Vocabulary	-	0.52*	-0.17	0.62*
2. Applied Problems		-	0.42	0.76**
3. Auditory Working Memory			-	0.32
4. Test of Relational Concepts				-

* = $p < .05$, ** = $p < .01$, *** = $p < .001$

Note. There were no statistically significant differences in the correlation coefficients for the associations between any of the cognitive assessments between the two conditions (*Z*s < 1.17, *p*s > 0.121).

Table 7

Correlation Coefficients for Associations between Self-Derivation through Integration Performance and Cognitive Assessments

Measure of Self-Derivation	Picture Vocabulary	Applied Problems	Auditory Working Memory	Test of Relational Concepts
Open-Ended				
Overall	0.38*	0.51**	0.55**	0.53**
Nonlinear	0.50*	0.47	0.42	0.48
Linear	0.28	0.56*	0.63**	0.59*
Forced-Choice				
Overall	0.15	0.40*	0.30	0.12
Nonlinear	-0.03	0.52*	0.26	0.20
Linear	0.19	0.18	0.22	-0.14

* = $p < .05$, ** = $p < .01$, *** = $p < .001$

Note. There were no statistically significant differences in the correlation coefficients for either measure of self-derivation performance and the cognitive assessments between the two conditions (Zs < 1.01, ps > 0.156).

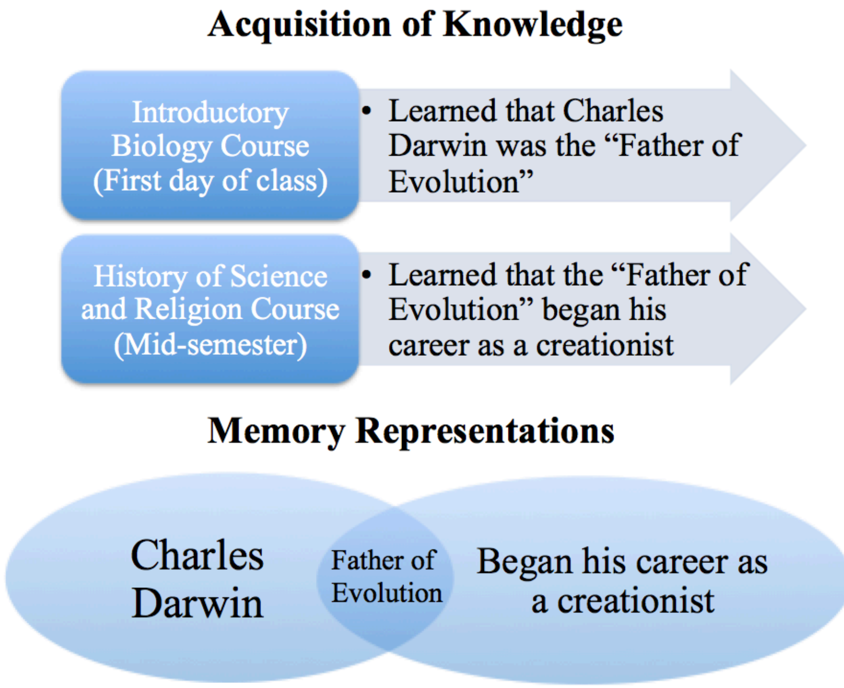

Figure 1. Diagram illustrating the process of knowledge integration, whereby an individual acquires related factual knowledge across different episodes of learning and integrates them to derive new knowledge. Overlapping representations of concepts are used to extract a new relation that was never explicitly learned (e.g., "The father of evolution began his life as a creationist").

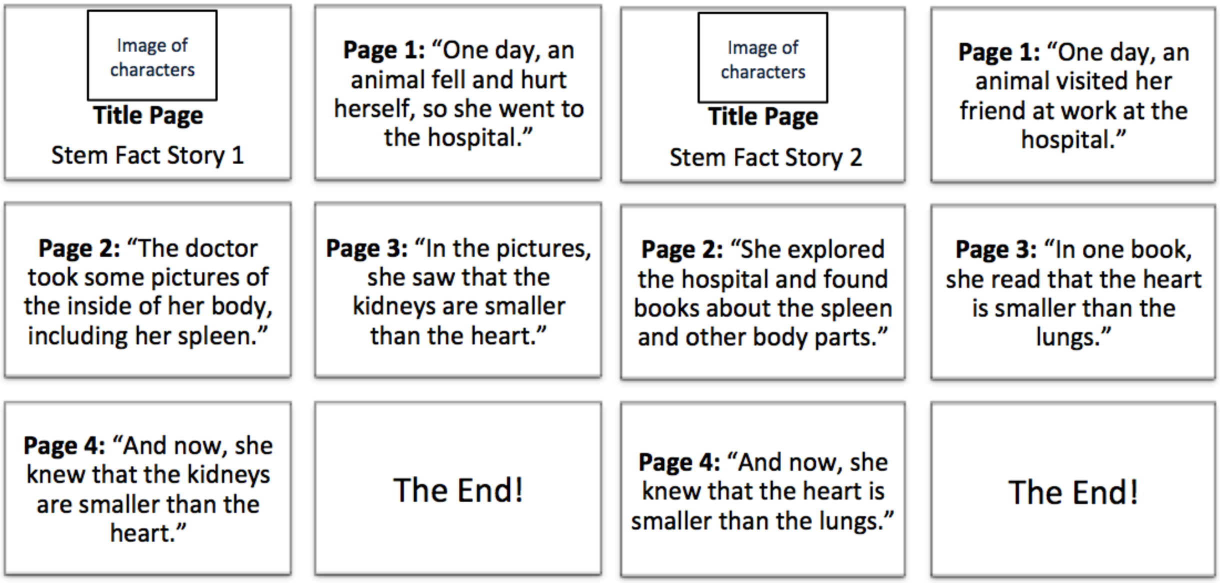

Figure 2. Example pages from stem-fact stories. Children viewed images of the story characters while listening to a female voice recording of the story text. The story containing the first stem fact always preceded the story containing the second fact. The first story was read twice before the second story was read twice. The same stories were used in both the linear and nonlinear conditions, with the exception of the types of facts learned on pages 3 and 4 of the story.

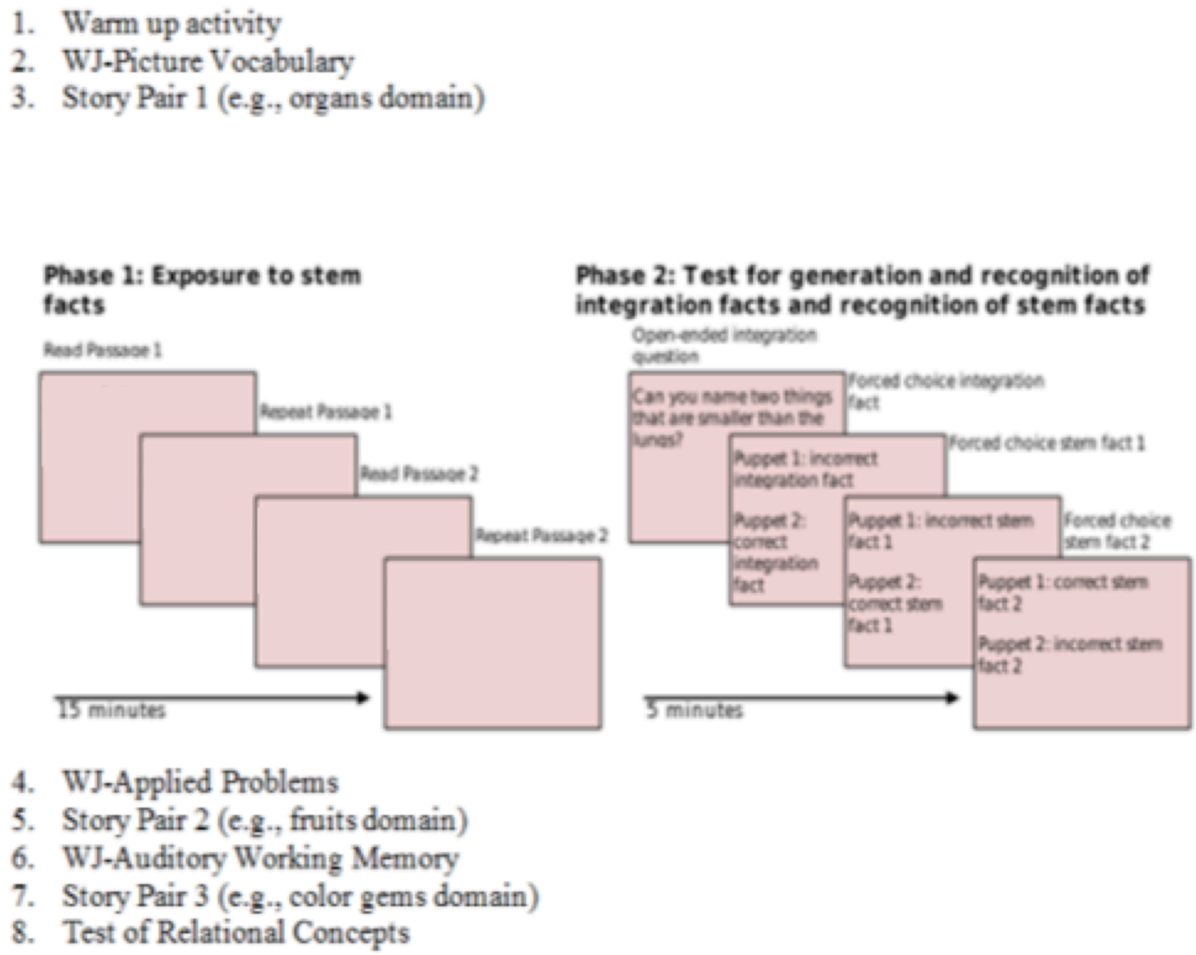

Figure 3. Schematic of the study procedure. Children were administered different cognitive assessments in between different story pairs. The order of the stem-fact story pairs (i.e., first, second, or third story pair) was counterbalanced across participants, while the cognitive assessments were administered in the same order for every participant. In total, the session took around one hour to complete and families received compensation upon completion of the session.

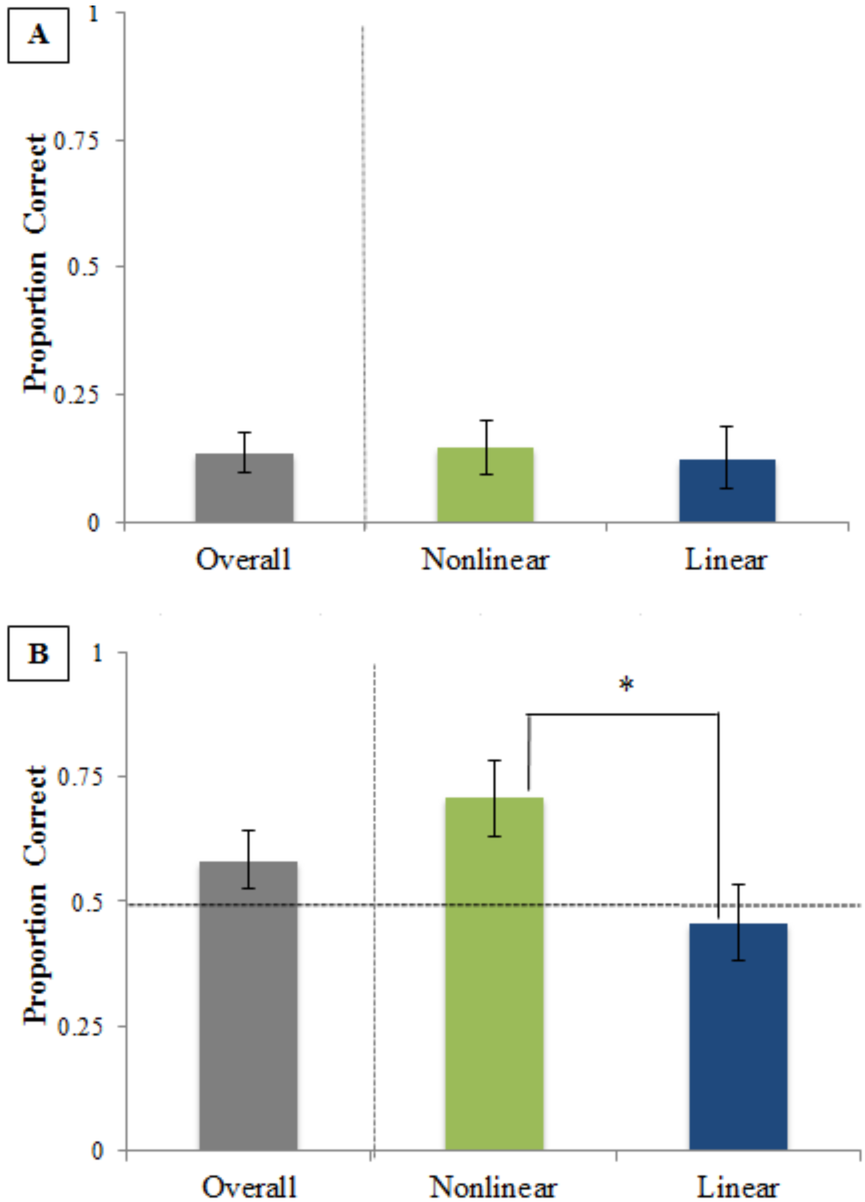

Figure 4. (A) Proportion of children who correctly answered the open-ended integration question overall and by condition. (B) Proportion of children who correctly selected the novel integration fact in the forced-choice portion of the integration test phase. There was a statistically significant difference in forced-choice self-derivation performance between the nonlinear and linear conditions; $t(30) = -2.40$, $p = .023$, $d = 0.85$. Error bars represent standard error.

*$p < 0.05$

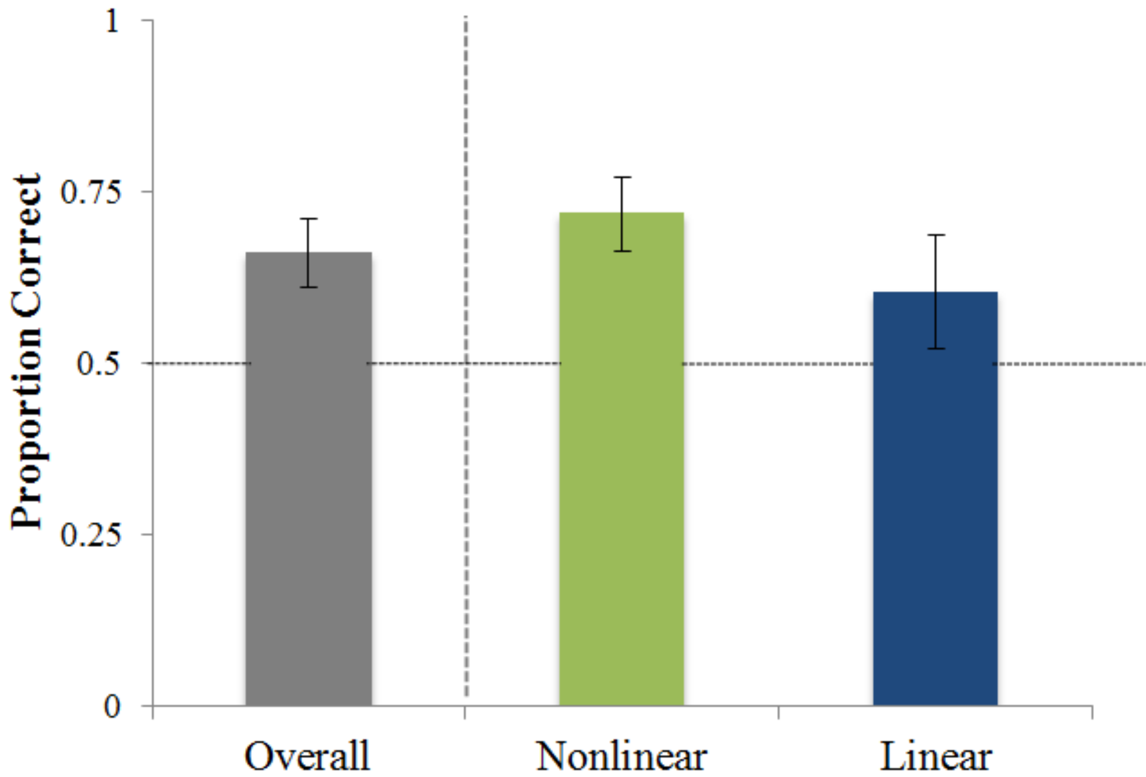

Figure 5. Proportion of children who identified the correct answer for at least one of the stem facts (out of a total of 6). There was no significant difference between the linear and nonlinear conditions $t(26) = -1.20$, $p = .242$, $d = 0.42$. Error bars represent standard error.

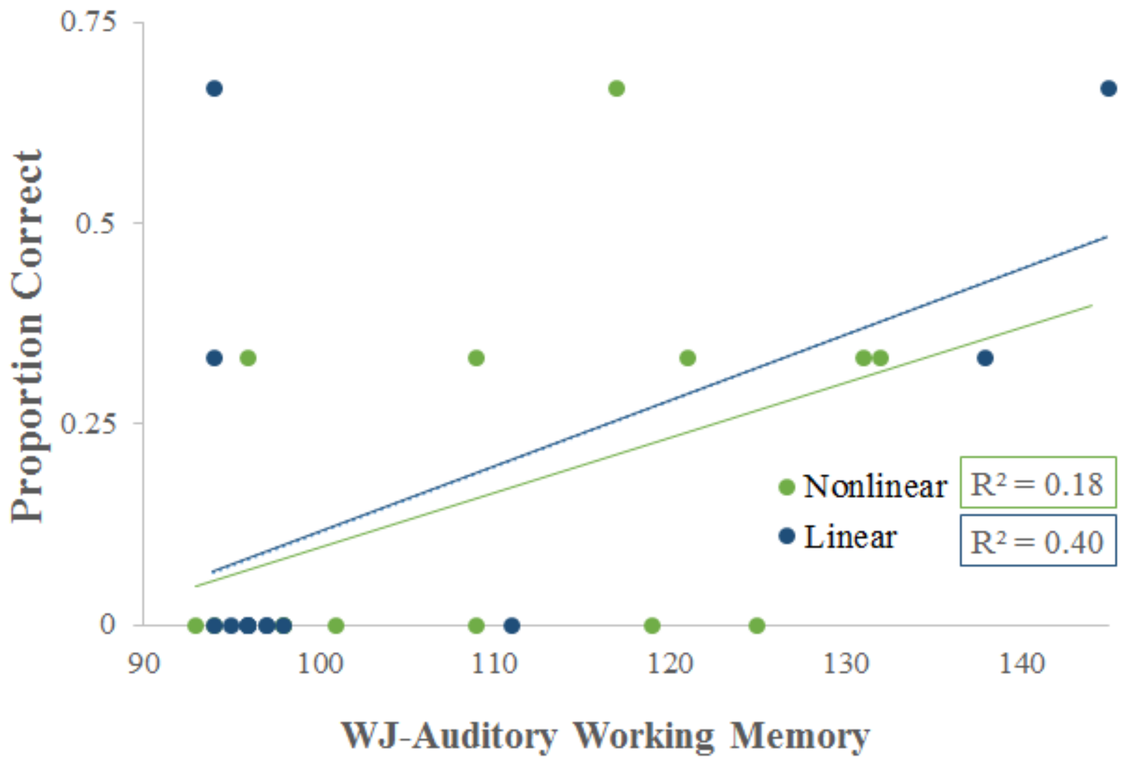

Figure 6. Open-ended self-derivation performance as a function of standardized scores on the WJ-Auditory Working Memory test. Overall, there was a significant positive correlation between open-ended self-derivation performance and scores on WJ-Auditory Working Memory ($p < .05$). This significant positive correlation was found for children in the linear condition ($p < .05$) but not for children in the nonlinear condition.

Appendix A

Stem Fact Performance by Fact Domain

Forced-choice self-derivation performance in both conditions differed based on the specific relational terms that were used to link the facts (for the list of all relational terms, see Table 1). In the nonlinear condition, children's forced-choice self-derivation performance was significantly above chance in the Gems domain (M = 87.50%, SE = 11.69%; Z = 2.10, p = .034), in which the integration fact was a non-biological category detail ("are compounds like"; see Table A1). Object category learning emerges early in development, and labeling greatly increases children's ability to form categories (Bornstein & Arterberry, 2010; Sloutsky & Fisher, 2004). Therefore, children in our task may have found grouping three like objects (a green gem, red gem, and blue gem) into a labeled category easier than assessing the similarities between objects that share a mechanism of growth or the same perceptual qualities. All other nonlinear relational terms were not significantly greater than chance (Zs < 1.40 , ps > .157). Nonetheless, children were at or above the chance level 50% for every relational term used in the nonlinear condition. Because the sample size for children exposed to each relational term was so small (N = 8), interpreting these null results is cautioned.

In the linear condition, children's forced-choice self-derivation performance was not significantly different from chance for any of the different relational terms [Zs < 1.40 , ps > .157]. Children's forced-choice self-derivation performance in the linear condition ranged from an average of 25.00% (SE = 15.31%) in the Space domain to an average of 62.50% (SE = 17.11%) in the Birds domain. Interestingly, children demonstrated the highest forced-choice self-derivation performance on the transitive fact pair that contained a more abstract concept ("smarter than")—although again, their performance was not significantly above the chance level of 50%.

There was a significant difference in forced-choice self-derivation performance between the two conditions for the Space domain, in which the relational term was one regarding a possessive quality ("has moons like"; $M = 75.00\%$, $SE = 15.31\%$) in the nonlinear condition and spatial distance ("closer than"; $M = 25.00\%$, $SE = 15.31\%$) in the linear condition [$Z = -2.00$, $p = .046$]. The observation that the lowest self-derivation performance in the linear condition was found for the Space domain is somewhat surprising, given that the relational term used in the Space domain was the one most grounded in physical reality. This may suggest that the integration of spatial relations is more difficult than the integration of abstract relations when facts are embedded in a rich, episodic framework. For all 8 domains, forced-choice self-derivation performance was equal to or higher in the nonlinear condition than in the linear condition, although small sample sizes may have precluded significant results from emerging [Zs < 1.62, ps $> .105$].

Table A1

Forced-Choice Self-Derivation Performance by Domain

Story set X

Condition	Space	Aviation	Birds
Nonlinear	0.75 (0.15)	0.75 (0.15)	0.75 (0.15)
Linear	0.25 (0.15)	0.50 (0.18)	0.63 (0.17)
Overall	0.50 (0.15)	0.63 (0.16)	0.69 (0.16)

Story Set Y

Condition	Organs	Fruits	Gems
Nonlinear	0.63 (0.17)	0.50 (0.17)	0.88 (0.12)
Linear	0.38 (0.17)	0.50 (0.17)	0.50 (0.17)
Overall	0.50 (0.18)	0.50 (0.17)	0.69 (0.15)

Note. Values represent the proportion of children who successfully self-derived the novel integration fact in a forced-choice format for that given stem fact pair domain. Standard error is presented in parentheses.

Appendix B

Stem Fact Performance by Stem Fact Order

Across conditions, children's memory for the second stem fact (*M* = 69.79%, *SD* = 30.94%) was greater than their memory for the first stem fact (*M* = 62.50%, *SD* = 33.60%). Nonetheless, children performed above the chance rate of 50% in the recognition of both the second stem fact [$t(31) = 3.62$, $p = .001$, $d = 0.64$] and the first stem fact [$t(31) = -2.10$, $p = .044$, $d = 0.37$]. Therefore, children were still able to recognize both stem facts despite showing evidence of a recency effect in their marginally better recognition of the second stem fact. Recognition of the second stem fact was lower in the linear condition than in the nonlinear condition (*M* = 64.58%, *SD* = 35.42%, and *M* = 75%, *SD* = 25.81%, respectively; see Figure B1), but this difference was not statistically significant [$t(27) = -0.95$, $p = .350$, $d = 0.34$]. Similarly, recognition of the first stem fact was lower in the linear condition than in nonlinear condition (*M* = 56.25%, *SD* = 35.93%, and *M* = 68.75%, *SD* = 30.95%, respectively). Again, this difference was not statistically significant [$t(29) = -1.05$, $p = .300$, $d = 0.37$]. Together, these results indicate that stem fact memory for both the first stem fact and the second stem fact was comparable across the two conditions.

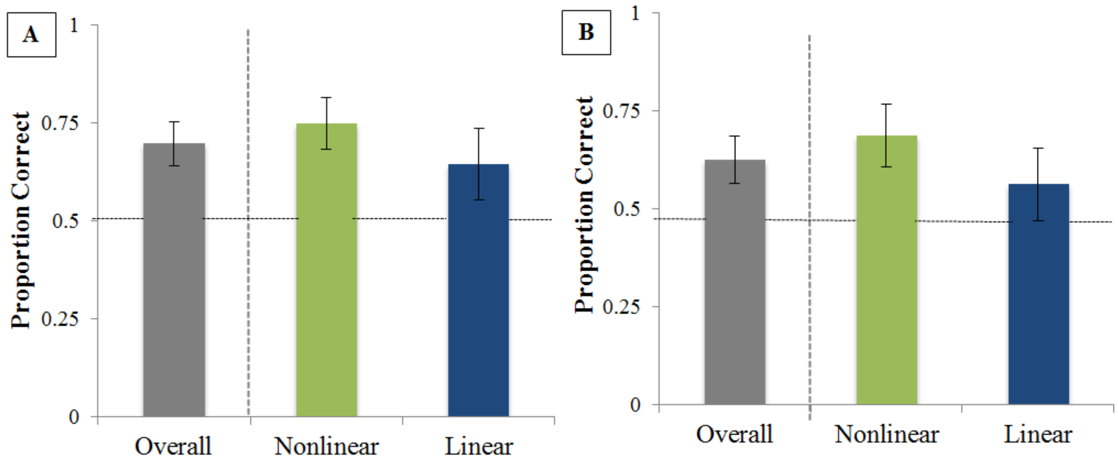

Figure B1. (A). Proportion of children who selected the correct answer for the second stem fact (out of a total of 3). There was no significant difference between the two conditions; $t(27) = -0.95$, $p = .350$, $d = 0.34$. (B). Proportion of children who selected the correct answer for the first stem fact. Again, there was no significant difference between the two conditions; $t(29) = -1.05$, $p = .300$, $d = 0.37$. Error bars represent standard error.