CHILD DEVELOPMENT

Child Development, January/February 2014, Volume 85, Number 1, Pages 160–175

The Effect of Executive Function on Biological Reasoning in Young Children: An Individual Differences Study

Deborah Zaitchik, Yeshim Iqbal, and Susan Carey
Harvard University

There is substantial variance in the age at which children construct and deploy their first explicit theory of biology. This study tests the hypothesis that this variance is due, at least in part, to individual differences in their executive function (EF) abilities. A group of 79 boys and girls aged 5–7 years (with a mean age of 6½ years) were presented with two test batteries: (a) a biology battery that probed their understanding of life, death, and body functions and (b) an EF battery that tested working memory, inhibition, and set-shifting skills. Individuals' EF scores significantly predict their biology scores, even after controlling for age and verbal IQ.

Learning science is hard. In the last quarter of the 20th century, research has uncovered one reason for students' difficulties in mastering both *explicit* theories—theories that are verbalizable and often the target of instruction in school—and *implicit* theories—theories that articulate everyday thought, often supported by representations that are not verbally accessible. Students do not come to the learning problem as blank slates; rather, they come already in the grip of systematic intuitive theories of many scientific domains. These intuitive theories differ from—and are often even incommensurable with—the scientific theories taught in school and those held by the adults in the child's culture. The important point is this: Learning science demands more than just the accumulation of new information; it demands conceptual change (see Carey, 2000, 2009, for reviews).

Vitalist Biology

This article focuses on the acquisition of v*italist* biology, the cross-culturally universal theory that underlies thinking about life, death, and health (see Carey, 1985; Inagaki & Hatano, 2002, for reviews). According to this theory, air, food, and water are sources of *vital energy* or *vital substance*, which must be obtained from the outside world to be sent to all parts of the body in order to maintain bodily function. Vitalism provides a functionalist understanding of bodily processes: Their goals are to sustain life, health, and growth. Body parts are specialized and work together. Some body parts serve as containers: Lungs hold air, the stomach holds food and water—and each body part also has specific causally relevant processes to carry out (e.g., food is broken into tiny pieces in the stomach; the heart pumps blood through the body and blood carries food, air, and water to all parts of the body). Death, in this view, is due to the breakdown of bodily function. The process of constructing a vitalist biology begins as young as age 4 or 5 for some children, with an average age of first emergence around age 6 or 7 (Inagaki & Hatano, 2002; Slaughter, Jaakkola, & Carey, 1999).

Before the Acquisition of Vitalist Biology: Preschoolers' Understanding of Animals

Research on preschool children's biological knowledge began with Piaget's (1929) classic studies of childhood animism, the tendency to attribute life—clearly a key concept in a theory of biology—to inanimate objects. When asked what it means to be alive, preschoolers often respond that it means to move or to be active. They often attribute life to the sun and the wind, as well as animals and people, but deny life to plants. Moreover, young children fail to differentiate *alive* from *real*, *visible*, *present*, or even just *existing* (Carey, 1985). Similarly, they fail to differentiate the contrast between *alive* and *dead* from the contrast between *animate* and *inanimate* (Carey, 1985). In line with the claim that

This work was supported by a grant from the National Science Foundation (NSF INSPIRE 1247396).

Correspondence concerning this article should be addressed to Deborah Zaitchik, Department of Psychology, Harvard University, Cambridge, MA 02138. Electronic mail may be sent to dzaitchiksamet@partners.org.

 0009-3920/2014/8501-0011
DOI: 10.1111/cdev.12145

young children have not worked out the biological meaning of *alive*, they show a similar confusion when asked what it means to *die*. Common responses are that it means to stop *doing* things or to become *invisible* or simply to *go away, to live on in vastly altered circumstance, such as under the ground* (Carey, 1985). Later work showed that when asked about the function of the body organs, they tend to report a single independent function for each body part (e.g., the heart is *for beating*), showing no understanding of the body as a biological *system* whose parts work together to sustain life (Slaughter et al., 1999). Moreover, though they understand that bodily processes are not under *intentional* control (Hatano & Inagaki, 1994; Hickling & Wellman, 2001), they appear to have no alternative conception—no conception of *biological* control—with which to understand them. Similarly, young preschool children have been shown to have a concept of death as the cessation of existence, a concept they draw on when thinking about relations between predators and prey (Barrett, 1999). This, however, is a far cry from vitalist biology, the biological theory that sees death not as the end of existence, but as the end of life.

Underlying these immature responses, we believe, is a system of concepts that identifies life with animals—and animals, in this system, are conceptualized as fundamentally causal and intentional *agents.* This explains why the sun and the wind (as well as people and animals) are considered *alive* while plants, which certainly do not *do* much or *go* anywhere, are not. It is only when children undergo the conceptual change from this agent-centered theory of animals to the more mature vitalist biology that they can differentiate *alive* from *existent*, *real*, or *active* and *dead* from *inanimate.* Because of these conceptual differentiations, along with the coalescence of *animal* and *plant* into the single ontological kind *living thing*, a vitalist biology is locally incommensurable with preschoolers' concepts of animals and plants. In constructing a vitalist biology, children create an interrelated system of knowledge couched in a whole set of concepts not available to young preschoolers.

Accretion of Facts and Conceptual Change

Clearly, it is only with age that children acquire the culturally widespread vitalist theory, but what is it about age that leads to this acquisition? Is it merely that older children, having lived longer, have acquired more facts about animals? The accretion of facts certainly plays an important role in learning any science. Nevertheless, accretion of facts is different from conceptual change in several critical ways. First, in learning new facts, children simply alter beliefs stated in the very same vocabulary they already have. On the basis of the above analysis, however, preschool children do not have the concepts to represent the fact "animals and plants are alive." Second, in cases of conceptual change, a system of concepts is mastered all together, as a coherent, interdefined whole, and the new concepts are partly defined by their place in that new coherent whole.

The claim that the acquisition of vitalist biology is an instance of such conceptual change is based on evidence for the incommensurability of concepts of biological phenomena in preschoolers and young elementary school children (Carey, 1985; Johnson & Carey, 1998). In addition, two findings in the developmental literature (Slaughter & Lyons, 2003; Slaughter et al., 1999) argue for the strong coherence of biological concepts. First, performance on an interview tapping the understanding of body parts and their functions was shown to predict performance on an interview about death. Second, a curricular intervention aimed at teaching the biological functions of internal organs increased performance on the death interview as much as the body parts interview, despite the fact that the concept of death was not the focus of instruction. This study will put the existence of such coherence to a strong test.

Mechanisms Underlying Conceptual Change

Conceptual change requires intellectual work on the part of the learner, but just which cognitive mechanisms underlie such work? Recently, research has begun to focus on the acquisition of a set of domain general higher cognitive capacities known as executive function (EF) skills. This suite of EF skills includes working memory (WM), the capacity to inhibit competing responses, and the capacity to monitor and flexibly select among potentially relevant sources of information (set shifting). EF is implicated in the formation of abstract representations, which in turn aid flexibility in reasoning and rule following (Kharitonova & Munakata, 2011; Snyder & Munakata, 2010), as well as in the hierarchical control of thought (Zelazo & Frye, 1997).

EF develops dramatically during infancy and the preschool years, and continues to do so through the school years (Cepeda, Kramer, & Gonzalez de Sather, 2001; Davidson, Amso, Anderson, & Diamond, 2006; Luciana & Nelson, 1998). Recent studies attest to the importance of EF in academic success: It is more strongly associated with school readiness than is IQ or entry-level reading or math skills (e.g., Diamond, Barnett, Thomas, & Munro, 2007). Moreover, EF has been shown to maintain its importance throughout the school years; indeed, WM and inhibition independently predict math and reading scores in every grade from preschool through high school (e.g., Blair & Razza, 2007; Gathercole, Tiffany, Briscoe, Thorn, & ALSPAC Team, 2005).

The Role of EF in Conceptual Change

In episodes of conceptual change, the learner builds a new explanatory framework for understanding some phenomena, and comes to inhibit previously useful explanatory frameworks for those same phenomena. Consider this example of the conceptual work engaged in by a preschool child in the midst of constructing a vitalist biology: She said, "That's funny, statues are not alive, but we can still see them!" The adult replied, "What's funny about that?" She said, "Well, Grandpa's not alive and that's sad because we can't see him anymore." She had noticed a contradiction in her concept *alive* (due to her failure to differentiate *dead* from *inanimate*) and was seeking help in resolving it (Carey, 1985). Being only 4, she could not understand the adult's response (i.e., she did not have a conception of death as the end of the life cycle, due to breakdown of the bodily machine), but this example illustrates the comprehension monitoring, WM, and set-shifting skills needed even to *engage* these issues. Conceptual change requires *detecting* conceptual problems, WM is necessary to keep all the information in mind (statues are not alive but are visible; Grandpa is not alive and not visible), and set shifting is necessary to eventually discriminate the two senses of "not alive" ("dead" and "inanimate") and reason appropriately depending on the context (statues or Grandpa). The bootstrapping processes involved in episodes of conceptual change often involve analogical reasoning (see Carey, 2009). All of these processes make heavy demands on EF. This analysis motivates the present study, the aim of which is to begin an exploration of the role of EF in the construction of, or expression of, vitalist biology.

Several empirical findings also motivate the research reported here. First is a study of adults with Williams syndrome (WS), a condition that leads to severe mental retardation yet spares the capacity to acquire both language and factual knowledge (Johnson & Carey, 1998). Strikingly, though their factual knowledge was relatively intact, equivalent to the factual knowledge of normally developing 12-year-olds, not a single WS adult demonstrated the biological understanding achieved by normally developing 5-year-olds. Lacking the framework of a vitalist biology, they failed to understand life in terms of bodily function and death as the breakdown of the bodily machine; similarly, lacking understanding of any biological mechanisms involved in reproduction or biological inheritance, they failed to distinguish between biological and social conceptions of family. These data provide powerful evidence that a huge database of factual knowledge does not, in and of itself, lead to conceptual change. We now ask *why* WS adults never create or express a vitalist or inheritance biology. Our hypothesis is that the mechanisms of EF—which are impaired in WS (e.g., Rhodes, Riby, Park, Fraser, & Campbell, 2010)—are the very mechanisms necessary for the construction or expression of theoretical knowledge.

Second, if the Animism Interview is administered in a speeded presentation to college undergraduates, they attribute life to the sun and the wind while denying it to plants (Goldberg & Thompson-Schill, 2009). That is, they make the same errors that young children make under nonspeeded conditions. In line with this is our own recent finding (Zaitchik & Solomon, 2008a, 2008b) that healthy elderly adults (as well as patients with Alzheimer's disease [AD]) attributed life to those very same items—the sun, the wind, and other seemingly active inanimates. These findings suggest that even healthy adults, with their firmly entrenched biological understanding, still suffer interference from their immature agency-based conceptions of life. Overcoming such interference, we suggest, may draw heavily on EFs such as inhibition and response selection.

Finally, this study builds on a large body of work showing that EF skills and theory of mind milestones are highly correlated in preschoolers, even controlling for age and verbal IQ (e.g., Carlson & Moses, 2001; Sabbagh, Xu, Carlson, Moses, & Lee, 2006). While these studies are important in demonstrating a role for EF in theory development or deployment, the case for vitalist biology may be quite different. In the theory of mind case, there is controversy over the role of infant representations of mental states (for arguments that the development of theory of mind does not require conceptual change, see Fodor, 1992; Scholl & Leslie, 1999). In other words, it is not clear how much of theory of mind is actually *constructed* and to what extent it involves genuine conceptual change. Vitalist biology, on the other hand, is undoubtedly a construction.

This study tests the hypothesis that EF scores in children from 5 to 7 years will correlate with their performance on conceptual tasks tapping vitalist biology, even controlling for age and verbal IQ. Such a finding, if obtained, would be the first result suggesting a role of EF in a wholly constructed theory, one whose development undoubtedly requires conceptual change. In addition, it would be the first demonstration of a relation between EF and a *scientific* theory, a scientific theory that is, moreover, a central target of instruction in the elementary school curriculum.

The Study

This study tests two hypotheses: (1) that an individual's scores across a range of vitalist biology tests will be significantly correlated, demonstrating the coherence associated with cases of conceptual change, and (2) that an individual's scores on vitalist biology tests and EF tests will be significantly correlated, providing support for a role of EF in the construction of vitalist biology, its deployment in reasoning, or both. To test these hypotheses, we constructed and administered two test batteries, a biology battery composed of tasks that diagnose mastery of a vitalist biology, and an EF battery. Hypothesis 1 predicts significant within-subject correlations among the three biology tests even after controlling for individual differences in age and IQ. To assess IQ, we administered the Peabody Picture Vocabulary Test (PPVT), on which standard scores provide a measure of verbal IQ. Hypothesis 2 predicts that individuals' biology scores will be correlated with their EF scores, even after controlling for the variance attributable to both age and PPVT. This is a stringent test of Hypothesis 2 because among the cognitive abilities that change with age are the very same EFs that we hypothesize play a role in the construction and deployment of vitalist biology.

Method

Participants

Participants were 79 children (48 girls, 31 boys) aged 5–7 years (M_{age} = 6 years 6 months), an age span in which, according to the literature, we should find a large degree of variance in their conceptual understanding of vitalist biology. The children were drawn from a relatively homogeneous middle-class population (92% White, 3% Black, 5% Hispanic). Roughly half were tested in a quiet room in their schools. Parents were informed about the study by means of a detailed written description and consent form sent home from school with the children. If parents consented to having their child participate, they signed the form and sent it back. The rest of the children were tested in the Harvard Laboratory for Developmental Studies. In the lab, the study was described in detail to the parents and usually parents were in the room with the child for both sessions. All children were given a small stuffed animal as a thank-you gift for participating.

Procedures

The Biology Battery

As noted above, the vitalist biology constructed between ages 5 and 8 sees air, food, and water as sources of *vital energy* or *vital substance* that must be obtained from the outside world to be sent to all the parts of the body; the lungs hold air, the stomach holds food and water—and the heart moves them through the body in blood. The teleological goals of bodily function include maintaining life (and preventing death, which is now seen as due to the breakdown of bodily function), maintaining health, and ensuring growth. According to vitalism, then, it is the body and its functions—not merely causal and intentional agency—that determine whether an entity is a living thing. We administered three interviews to assess understanding of the key concepts of *life*, *death*, and *bodily function*.

The Animism Interview. The Animism Interview (Carey, 1985; Piaget, 1929) probes the child's understanding of what it means to be alive. In this task, participants are asked several open-ended questions: (a) What does it mean to be *alive*, to be *a living thing*? (b) Can you name some things that are *alive*, that are *living things*? and (c) Can you name some things that are *not alive*, that are *not living things*? These questions are followed by a list of animate and inanimate objects including people, various animals, plants, natural kinds, and artifacts. For each object named, the participant is asked, "Is it *alive*? Is it a *living thing*?" Justifications for yes–no judgments of a subset of the items are also collected.

The Death Interview. The Death Interview (Carey, 1985; Slaughter & Lyons, 2003; Slaughter et al., 1999) probes the child's understanding of what it means to die. As in the animism task, children are asked several open-ended questions: (a) What does it mean to *die*? (b) What happens to a person when he dies? and (c) What happens to a person's *body* when he dies? These questions are followed by a series of yes–no questions tapping the understanding that dead people no longer have *any* bodily or mental functions (e.g., "Does a dead person: eat? pee? feel bad that he's dead?") Children are then asked what causes someone to die, whether everyone has to die, and whether someone who is dead can come back to life.

The Body Parts Interview. The Body Parts Interview (Carey, 1985; Slaughter & Lyons, 2003; Slaughter et al., 1999) asks about the location and function of a series of body parts (brain, heart, lungs, stomach, blood) and, for each body part, what would happen if a person did not have it. Children were then asked why we eat, why we breathe, and what happens to the food and air after they are taken in to the body.

Scoring of the Biology Interviews

The interviews were transcribed and coded blind with respect to participant (for complete interviews and details of the scoring criteria, see the online Appendix S1). In the Animism Interview, points were awarded for appropriately constraining attributions of life to plants and animals and justifying those attributions by appealing to biological processes such as birth, growth, and death. The total score was the sum of two subscores: (a) the degree of animism indicated in the yes–no *judgments* and (b) a number assigned, using a qualitative coding of the *introductory questions* and *justifications*, which reflected the degree to which an explicit mastery of a vitalist biology has been achieved. In addition, positive evidence of failing to distinguish *existence* from *life* characteristic of the responses of 3- and young 4-year-olds) could result in a score of −1. Scores could range from −1 to 19.

In the Death Interview, points were awarded for indicating that death is the end of all body function, that the body decays after death, that all living things, both animals and plants, eventually die, and that death is irreversible. Points were subtracted if the child provided positive evidence that they failed to distinguish the religious and biological aspects of death, or if they took death to be living on in some altered state (e.g., asleep under the ground). Scores could range from −12 to 10.

In the Body Parts Interview, points were awarded for mapping bodily functions that support vitalist goals onto body organs, and for showing an understanding that the body is a system whose parts work together to support life. Scores could range from 0 to 25.

Two independent coders scored the battery: one of the authors, who administered the interviews, and a research assistant who administered no tasks and was blind to study hypotheses. For each test, agreement between coders exceeded 90%. All disagreements were resolved by discussion.

The EF Battery

Our EF battery included three tests.

Hearts & flowers (H&F) and Flanker Fish (FF). These two computer tests, developed by Adele Diamond, are normed for children as young as age 5. Each test taps three EFs on which there is strong consensus in the literature: inhibition, set shifting, and WM (Davidson et al., 2006; Diamond et al., 2007). The structure of the two computer tests is identical, with each test including three condition blocks presented in the following order: congruent, incongruent, and mixed. Each block contains 20 trials. We illustrate the two computer tests here using H&F.

H&F: Congruent. In this condition, the child is presented with a series of hearts that appear, one at a time, on either the left or the right side of the screen. The child must press the button on the *same* side of the screen that the heart appears on. This is the easiest of the three blocks as it requires remembering just one rule, and little or no inhibition or set shifting.

H&F: Incongruent. The child is presented with a series of flowers that appear, one at a time, on either the left or the right side of the screen. The child is instructed to press the button on the *opposite* side of the screen that the flower appears on. Like the congruent condition, this condition demands remembering only one rule. However, since this rule differs from the rule just followed, the task is somewhat more demanding of EF insofar as it requires inhibiting the practiced response. Moreover, this condition demands the inhibition of the "Simon effect," the prepotent tendency to press the button on the same side of the screen that the stimulus appears on.

H&F: Mixed. The child is presented with a randomly ordered series of congruent and incongruent trials. When presented with a heart, the child must press the button on the same side of the screen, as in the congruent condition. When presented with a flower, the child must press the button on the opposite side of the screen, as in the incongruent condition. Since this condition is a mix of congruent and incongruent trials, correct responding requires the child to keep both rules in mind (for heart, press same side; for flower, press opposite side) on each trial, so the WM load is far more demanding than in either single block condition. Moreover, the child must flexibly shift between rules depending on whether the stimulus presented is a heart or a flower, and must inhibit any tendency to push the button right below a stimulus or to respond without fully identifying the stimulus.

The FF test is similar in structure to the H&F test so we describe it only briefly. In each array, a row of five fish appears on the screen. In the first block (corresponding to the congruent condition of H&F), the fish are pink. The child is told to press the button specifying the direction in which the central fish is facing. In the next block (corresponding to the incongruent condition of H&F), the fish are blue. Here the child is told to press the button specifying the direction in which the FF are facing. In the mixed condition, there are both congruent and incongruent trials.

For young children, these tests yield two types of informative score: (a) accuracy and (b) failures to respond to stimulus (FRS). Accuracy is a measure of correct responding; as such, it reflects the ability to hold two rules in WM, to inhibit a salient but inappropriate response, and to switch rules (or *set*) depending on the stimulus presented. FRSs include failures to press a button at all, as well as button presses made before the stimulus could possibly have been perceived (< 200 ms). Since such rapid button presses are not responses *to the stimulus*, they are not counted in the subject's accuracy score. As Diamond and her colleagues argue, such rapid button presses reflect the failure to appropriately inhibit a response until a stimulus is processed (Davidson et al., 2006). Of course, what seemed a rapid response to one trial might really have been a slow response to the previous trial. If this were so, however, the previous trial should have had no response. To test this, we analyzed whether rapid responses were likely to follow nonresponses. Of roughly 100 rapid response trials, only two followed trials on which there was no response. We take this as good evidence in favor of Diamond's interpretation that rapid responses reflect inhibitory failures, and as justification for treating both nonresponses and rapid responses as failures to respond (FRS).

Each of the two computer tests (H&F and FF) yields the same two types of scores (FRS and accuracy) for a total of four scores—while reaction times are also recorded, they have been reported to be unreliable in young children on these tasks (Diamond et al., 2007) so they are not included in the analyses.

Color words: Forward and backward. In addition, we administered a classic test of WM, the Color Words Test. In the simpler forward condition, which measures WM *span*, the experimenter asked the child to repeat a list of spoken color words. The length of the list grew by one word on each successive trial. If children failed on a particular list, they were given another list of the same length, and if they succeeded on the second trial they were given credit for that list length and allowed to continue. The final score, the longest list length on which they succeeded, reflects the span of their WM. The backward condition, which measures *flexibility* of WM, was identical in both administration and scoring to the forward condition—except that the child had to repeat the words in backward order. This condition demands that the child maintain the spoken word list in memory and successfully manipulate the word representations held in WM.

Peabody Picture Vocabulary Test

PPVT is a receptive vocabulary test, in which children must indicate which of four pictures depicts a word (e.g., *wrench*, choosing among pictures of a wrench, hammer, screwdriver, and anvil). Vocabulary grows throughout the life span and the PPVT is normed for mental age from roughly age 3 to 91. Vocabulary is a major part of verbal IQ tests and indeed PPVT standard scores are highly correlated with verbal IQ scores as assessed by full tests like the Wechsler Intelligence Scale for Children (Hodapp & Gerken, 1999). We chose PPVT because the items on this task do not require conceptual change for their mastery (e.g., young preschoolers have the concept of a tool, so whether they know what a wrench is and what it looks like is a matter of input and efficiency of uptake of information).

Testing occurred over the course of two 25-min sessions. In accordance with design requirements for individual differences studies, the testing proceeded in the same fixed order for all participants. In Session 1, each child was administered the vitalist biology task battery as follows: Animism Interview, Death Interview, Body Parts Interview. This was immediately followed by the PPVT. In Session 2, the EF task battery was administered in the following order: H&F, color words *forward*, FF, color words *backward*. Session 2 was administered between 1 and 3 weeks after Session 1.

Results

The Biology Battery: Descriptive Statistics and Within-Child Consistency

On the animism task (possible scores from −1 to 19), scores ranged from 0 to 16, with an average score of 7.7 and a standard deviation of 3.8. On the death task (possible scores from −12 to 10), scores ranged from −4 to 8, with an average score of 3.4 and a standard deviation of 4.3. On the body parts task (possible scores from 0 to 25), scores ranged from 0 to 21, with an average score of 10.3 and a standard deviation of 3.2. Thus, children in our sample showed considerable variance on each task, with performance ranging from the complete absence of an expressed vitalist biology to a nuanced explicitly articulated vitalist understanding of life, death, and bodily function.

To assess relations among our different tasks, all measures were converted to *z* scores. Data analysis began with bivariate correlation analyses among the three biology tests. As expected, all three tests were significantly correlated: animism_body parts, $r(79) = .64$; animism_death, $r(79) = .54$; and body parts_death, $r(79) = .58$, in all cases $p < .001$ (all *p* values reported in this article are two-tailed unless otherwise specified).

These correlations justify aggregating the three scores into a single composite measure of vitalist biology for the purpose of exploring the relations between EF and understanding of vitalism (Hypothesis 2). To calculate the aggregated score, raw scores were transformed to *z* scores and then summed, yielding a single vitalist biology score for each subject.

The correlations among the biology tests provide the starting point for exploring the coherence of intuitive vitalism. The three tasks differ from one another a great deal, each tapping different aspects of vitalist understanding. If vitalism develops as an integrated system of knowledge, the concepts of which are interdefined, then the three tasks should remain correlated when standard scores of PPVT (our measure of verbal IQ) and age are partialed out (see Table 1). With PPVT standard scores removed, correlations among the three tests of vitalist biology remained significant: animism_body parts, $r(73) = .64$; animism_death, $r(73) = .49$; body parts_death, $r(73) = .52$, all $ps < .001$. Next, age (in months) was partialed out. Correlations still remained significant for animism_body parts, $r(76) = .43$, $p < .001$, and body parts_death, $r(76) = .33$, $p = .004$. The correlation for animism_death, $r(76) = .19$, $p = .09$, remained marginally significant.

Table 1
Partial Correlations After Controlling for Age and Peabody Picture Vocabulary Test (Standard Scores)

Variable	Bio composite	Animism	Body parts	Death
EF composite	0.26*	0.22*	0.12	0.23*
Animism			0.45***	0.14
Body parts				0.26*

Note. $df = 72$. EF = executive function.
*$p < .05$, one-tailed. ***$p < .001$, one-tailed.

Even with both age and PPVT partialed out at once, correlations remained significant for animism_body parts, $r(72) = .45$, $p < .001$, as well as body parts_death, $r(72) = .26$, $p = .024$, while animism_death: $r(72) = .14$, $p = .246$, fell below significance.

These data provide strong evidence that vitalist biology is acquired as an integrated system of knowledge, as predicted by the analysis that the child's understanding of the functioning of the bodily machine supports understanding death as the breakdown of bodily function, which in turn plays a role in differentiating *alive* from *real*, *existent*, or *agentive* and distinguishing *dead* from *inanimate.*

The EF Battery: Descriptive Statistics and Within-Child

Each of the two computer tests (H&F and FF) yielded two outcome measures—accuracy and FRS—in each of the three conditions (congruent, incongruent, and mixed). Accuracy was near ceiling on the congruent condition (95% averaged across H&F and FF) and there were few FRSs in either the congruent or incongruent conditions (0.23% averaged across H&F and FF in each condition). Apparently, these conditions were too easy to yield the variance we would need to test the relation between EF and biology. The mixed condition, in contrast, yielded adequate variance: Mean accuracy was still high (H&F: $M = .83$, $SD = .16$; FF: $M = .88$, $SD = .11$), but not at ceiling; scores ranged from approximately 43% to 100% (on both H&F and FF), so there were marked individual differences.

The FRSs showed the same pattern as did errors in the accuracy analysis reported above. Figure 1 shows the steep increase in the percent of FRSs in the mixed condition compared to the congruent and incongruent conditions. Only the mixed condition (H&F: $M = .85$, $SD = 1.45$; FF: $M = 1.5$, $SD = 2.2$) yielded sufficient variance (ranging from 0 to 6 FRS errors on H&F and from 0 to 9 on FF)

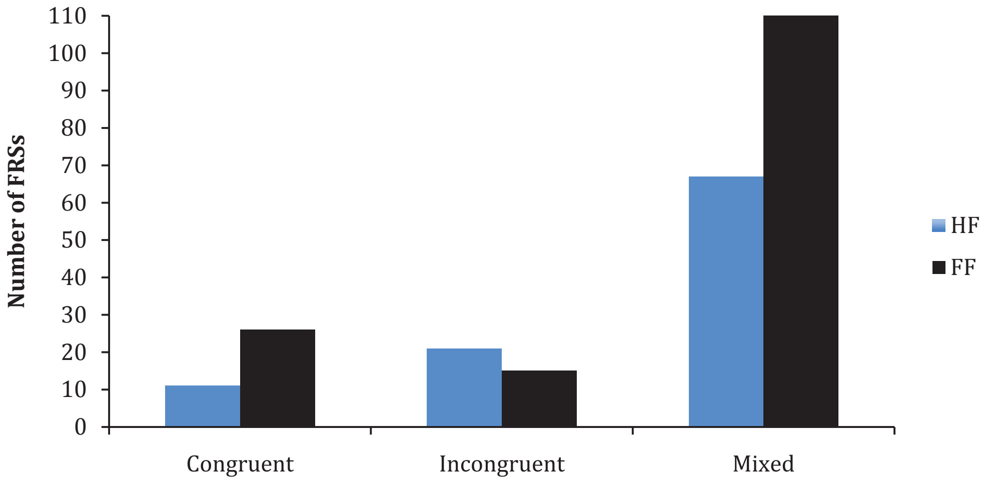

Figure 1. Number of failures to respond to stimulus, by test and condition.

for our purposes. The mixed condition was difficult enough to provide a sensitive measure of the child's abilities, so we took the mixed block scores as our measure of EF.

The Aggregated EF Score

Children's accuracy scores on the mixed condition of H&F and FF were significantly correlated, $r(77) = .40$, $p < .001$, justifying their aggregation (by summing the z scores) into a single accuracy score for each child. FRS scores in the mixed condition were also significantly correlated across the two tests, $r(77) = .47$, $p < .001$, so they were aggregated (again, by summing the z scores) into a single FRS score for each child. To facilitate comparisons and aggregations with accuracy scores, where higher scores reflect better performance, FRS z scores were reversed in direction (i.e., positive z scores became negative and vice versa) so that higher scores reflect better performance, as they do in the accuracy scores. The two aggregated scores—one for accuracy and one for FRSs—were also significantly correlated, $r(79) = .25$, $p = .026$, justifying further aggregation (again, by summing the z scores) of these measures into a single composite EF score for each child.

The Color Word Span Test scores did not show the variance among children necessary for meaningful use in a correlation analysis. For that reason, they are not included in the aggregated EF scores and we analyze them no further here. Presumably, we had too few trials to assess meaningful differences in WM, a problem that will be rectified in future work.

Gender

Preliminary analyses examined the effects of gender on each task. Boys and girls did not differ on any measure of EF or vitalist biology (simple or aggregate scores), or PPVT. Thus, gender was not further analyzed.

Relation Between EF and Vitalist Biology

The correlation between the aggregated EF and biology scores was significant, $r(79) = .52$, $p < .001$. To test our main hypothesis—that individual children's EF scores predict the degree to which they have constructed a vitalist biology—a multiple regression analysis was run with age, PPVT standard scores, and aggregate EF score as predictor variables and with aggregate biology score as the dependent measure (see Table 2). Not surprisingly, age and PPVT were significant predictors of the biology scores: age, $t(75) = 7.356$, $p < .001$, and PPVT, $t(75) = 2.988$, $p = .004$. Our question, however, is this: After removing the variance associated with age and PPVT, do EF scores *still* predict biology scores? They do—EF, $t(75) = 2.279$, $p = .026$. As mentioned above, this is a conservative test of the statistical relations between EF and vitalist biology, since the variance associated with age, which is controlled for in the regression analysis, is likely carried to some extent by age-related changes in EF.

Thus, the regression analysis confirms that children's EF scores are significant predictors of their mastery of basic concepts within a vitalist biology. This pattern of results is consistent with two interpretations: WM, inhibition, and set shifting are

Table 2
Linear Multiple Regressions of Variables Predicting Theory of Biology

Criterion	Variables	β	*t*	p^a
Bio composite	Age	.603	7.356	< .001
	PPVT	.222	2.988	.002
	EF	.185	2.279	.013
Animism	Age	.553	5.702	< .001
	PPVT	.100	1.139	.130
	EF	.188	1.953	.028
Body parts	Age	.489	4.574	< .001
	PPVT	.159	1.643	.053
	EF	.110	1.043	.151
Death	Age	.504	5.620	< .001
	PPVT	.308	3.794	< .001
	EF	.176	1.987	.026

Note. PPVT = Peabody Picture Vocabulary Test; EF = executive function. [a]One-tailed.

powerful factors affecting theory acquisition and conceptual change, or they are powerful factors affecting children's capacity to recruit and express their theoretical understanding on our clinical interviews. These two possibilities are not mutually exclusive; both could be in play. One way to begin to explore these two possibilities is to consider the relations between EF and each biology task. Although we have no formal analysis of the executive demands associated with each biology task, on their face they seem to provide quite different online challenges to EF.

Relation Between EF and Each Individual Biology Test

The Animism Interview is known to require inhibition of competing responses. The immature responses on this interview include judgments that moving entities, especially natural ones like the sun or the wind, are alive, and apparently inert entities like plants are not. Such responses most likely stem from the presence of a very salient, developmentally earlier, competing response in the conceptual system. Such a response might be the effect of a strong early developing attribution of *agency* to moving or active things, including animals, and an initial identification of "life" with animals, with agency at the core of reasoning about animals. Indeed, it was the presence of such errors in adults with WS, as well as the recurrence of such errors in the elderly and in in AD patients, that first led us to consider a role for inhibitory processing in the explanation of animist judgments.

The Death Interview, on its face, does not seem to tap any default response requiring strong inhibition. There are, however, two different ways to think about death and these are clearly reflected in interviews with both children and adults (Brent & Speece, 1993; Harris & Giménez, 2005). Specifically, subjects think about death in a religious or spiritual context as well as a biological one, and it is likely that the response selection and consistency monitoring aspects of EF will be needed to maintain a consistent stance. Unlike the animism phenomenon, there is no response that must be flatly *inhibited* as an *error*. That is, there is nothing wrong with the response that after death the soul goes to Heaven (as opposed to the response that after death, the body rots). Thus it is likely that the responses on the Death Interview might draw particularly heavily on the *set-shifting* aspects of EF.

The Body Parts Interview includes questions that probe the understanding that air and food provide the energy needed to sustain life. However, there are also questions that can be answered on the basis of factual knowledge alone, so the score on this test, more than the other biology tests, may depend to a greater degree on learned facts—that the *lungs are for breathing*, *the heart is for pumping blood*, *the stomach is for storing and digesting food*, and so on. These facts are critical in the child's construction of a notion of the bodily machine and, as we have noted, the task is correlated with other measures of the constructed vitalist theory (the Animism and Death Interviews). Still, the factual database itself could simply be learned. Moreover, and perhaps more significantly, there seem to *be no conflicting representations* or *alternative contexts* at play in these mappings of form to function. In short, unlike the Animism and Death Interviews, the Body Parts Interview does not appear to make high demands on either inhibition or set shifting. For that reason, we speculate, performance on this task itself does not draw on central executive resources as much as the Animism and Death Interviews do.

If the above informal analyses are correct—if the different EF demands of our biology tasks affect the ability to express vitalist biology, and these different expression demands partially underlie the relation between performance on our EF tasks and our biology tasks, then EF should predict performance on the three biology tasks to different degrees. As a test of this hypothesis, we assessed the degree to which individuals' scores on each biology task were independently correlated with their aggregated EF scores.

The strongest correlations were between EF and death scores, $r(77) = .48$, $p < .001$, and between EF and animism scores, $r(77) = .47$, $p < .001$. The weakest of the three correlations, though still significant, was between EF and body parts scores, $r(77) = .38$, $p < .002$. These correlations do not differ from each other statistically (Steiger's Z). With age partialed out, the order of the strengths of the correlations is preserved, and the correlation between animism scores and EF is significantly greater than that between body parts scores and EF, $Z(3) = 1.68$, $p < .05$. The correlation between death scores and EF fall between and do not differ significantly from either of the two other correlations. Finally, as Table 1 shows, with both age and PPVT standard scores partialed out, animism and death are each significantly—and nearly equally—correlated with EF, $r(72) = .22$, $p = .054$, and $r(72) = .23$, $p = .05$, respectively, while body parts is not, $r(72) = .12$, $p = .300$.

In sum, these analyses provide suggestive evidence that some of the relation between EF and performance on the biology battery does indeed reflect the online demands of our biology tasks and the expression of children's understanding of vitalism. We return, in the Discussion, to the likelihood that the construction of a vitalist biology also makes demands on EF, such that the relation we found between EF and performance on the biology battery may also reflect the role of EF in theory construction and conceptual change.

PPVT Raw Score

Constructing a vitalist biology clearly involves amassing factual knowledge: children must simply learn that food is necessary for growth and health and life, that the heart pumps blood, that the body rots after death, and a host of other facts that are integrated in the course of creating a vitalist biology. If, however, creating a vitalist biology requires *more* than the accretion of factual knowledge, namely, if it requires conceptual change—and if conceptual change requires EF—then EF should continue to predict the composite biology scores in a regression analysis with age and measures of factual knowledge in the model. For this purpose, we take PPVT raw scores as such a measure.

Standard score PPVT is a measure of the child's performance given his or her age. Thus, the standard score is appropriate for use as a measure of crystallized IQ, a construct that includes general factual knowledge as well as vocabulary. The raw score, in contrast, provides an absolute measure, one not standardized by age, of the word meanings and factual knowledge captured in the construct of crystallized IQ. As noted above, items mastered later on the PPVT (e.g., wrench vs. hammer) are merely less frequent; no conceptual change is needed to learn the names of rarer tools. The specific tools one can name, like many simple facts, is a function of input and the efficiency with which one learns new factual knowledge. Consistent with this analysis are correlations between PPVT scores and scores on tests of general information (e.g., Markwardt, 1997). Although we here use raw PPVT scores as a rough measure of accumulated factual knowledge, in future work we will include a standardized measure of factual knowledge (the Woodcock-Johnson) to further address this issue.

Of course, age predicts raw scores PPVT, $r(79) = .71$, $p < .001$, as well as the composite biology score, $r(79) = .74$, $p < .001$—from ages 5 to 7 children amass much factual knowledge, whether it be about tools or animals' bodies.

Regression analyses showed that EF does indeed predict our composite measure of performance on biological reasoning, even after controlling for the variance associated with age and raw PPVT, $t(75) = 2.206$, $p = .032$. Understanding the ambient culture's vitalist biology—the theory that sees death as the end of the life cycle and the result of the breakdown of the bodily machine, the theory that supports reasoning about what happens after death, about what things are alive and why, and about the biological functions of eating, breathing, and having a heart—requires biological concepts that are co-constructed during the years of 5–7. While accumulating factual knowledge is the input to this process, the process goes beyond the mere accumulation of knowledge.

Discussion

There are *no* positive demonstrations of the concepts of a vitalist biology in children under age 5. Our biology battery is drawn from the literature that documents the construction of this knowledge structure, and indeed our sample of children included some who had barely begun and some who had made substantial progress. The former group maintained that the sun is alive because it keeps us warm, or because it exists, that dead people need to go the bathroom, and that the heart is for love and we would be lonely if we did not have one. The latter group maintained that the sun is not alive because it does not grow or have babies, that dead people do not need to go to the bathroom because they are just bones (although their souls may have gone to heaven), and that the heart is for pumping blood so that food and air reaches all parts of the body. Replicating earlier findings (Carey, 1985), the responses of the two groups reflect the differentiations (e.g., of *dead* from *inanimate*) and coalescences (e.g., of *animals* and *plants* into the single ontological kind, *living thing*) that constitute conceptual changes.

The first important finding from this study is stronger evidence than previously available that vitalist biology is constructed as a coherent whole. Previous work had already shown that responses on animism and death interviews are correlated (see Carey, 1985, for a review), as are responses of death and body parts interviews (Slaughter et al., 1999). Furthermore, training on the biological functions of internal body parts improves responses on a death interview very similar to that we included here, even though "death" was never mentioned in the training (Slaughter & Lyons, 2003). That the target concepts of life, death, and the bodily machine are interdefined reflects their construction as a coherent whole. Here we have confirmed the interrelations of these concepts, extending them to all three measures at once, and shown that they reflect developmental processes that for the most part are not exhausted by age, variations in verbal IQ, or PPVT raw scores, which we have taken as a rough measure of accumulated factual knowledge.

The most important contribution of this study is to establish a relation between EF and vitalist biology during the years when the latter is just beginning to be constructed. As in the classic studies of EF and theory of mind (e.g., Carlson & Moses, 2001), we created an aggregate measure of EF, in our case one that reveals considerable variance among 6-year-olds. We created an aggregate measure of mastery of the concepts and explanatory structure of vitalist biology, on which there also is considerable variance at this age. Regression analyses showed that EF predicted performance on the biology battery, even after controlling for the variance due to age, verbal IQ, and PPVT raw scores. This is the first demonstration that the same kind of relation that holds between EF and theory of mind (in 3- and 4-year-olds) also holds between EF and an early developing intuitive theory important to science education (in 5- to 7-year-olds).

Online EF Demands of the Biology Tasks

This project opens a parallel set of questions to those debated in the literature on theory of mind and EF. Clearly our biology battery, especially the Animism Interview and the Death Interview, draw online on EF, and indeed, the correlations between EF and performance on the tasks were ordered according to plausible orderings of these demands: (animism = death) > body parts. Judgments about which entities are living require inhibition of still available alternative responses, and judgments about death require selection between two different viable conceptual frameworks (biological and spiritual).

There is ample further evidence, reviewed above, that EF is needed for the expression or demonstration of biological understanding. Relevant findings include that animist responses are found in healthy young adults under speeded conditions (Goldberg & Thompson-Schill, 2009), that is, when executive processes such as inhibition have insufficient time to operate. In addition, several studies of healthy elderly adults and mildly impaired patients with AD have revealed that the AD patients resemble preschoolers on the same battery of biology tasks used here (Zaitchik & Solomon, 2008a, 2008b, 2009). Even the healthy elderly controls made some errors never seen in young adults. To appreciate the surprising nature of this finding, consider that elderly adults have not only amassed a long lifetime of factual knowledge about living things; they have also spent many decades of their lives with a firm hold on the structure of vitalist biological theories (and even later developing biological theories, deriving from evolutionary and chemical understanding). Moreover, the reemergence of immature biological responses can occur quite early in the course of AD. Importantly, so does a marked decline in EF (e.g., Belleville, Chertkow, & Gauthier, 2007). Indeed, even healthy elderly experience a milder, but still noticeable, decline in EF (Dempster, 1992; Hasher & Zacks, 1988). It is reasonable to consider that these declines are related, that impairment in EF interferes with the expression of theoretical biological understanding. Studies with healthy elderly and early AD patients to further explore this hypothesis are ongoing in our lab.

EF and Construction

We conclude there is a strong case for the role of EF in the online expression of vitalist biology. But is EF needed for the construction of a vitalist biology as well? The correlations between EF and the biology battery that remain after removing the variance associated with age, verbal IQ, and PPVT raw scores are consistent with that hypothesis, and failing to find them would have falsified it. An alternative hypothesis is that associative learning processes —processes that do not require WM, inhibition, and set shifting—build vitalist biology from a network of factual knowledge. EF, in this view, is required merely for the expression of that knowledge in our tasks. Although the data from this study cannot adjudicate between these hypotheses, other data, as well as theoretical considerations, suggest to us that EF is likely to be drawn upon in the construction process as well.

The Evidence From Adults With WS

Johnson and Carey (1998) tested 10 WS adults on a much more extensive battery of biological reasoning tasks than used here, some of which appeared to make extremely minimal online EF demands. All 10 performed like our young, low-EF 5-year-olds on the tasks that reflect vitalist theory, in spite of having factual knowledge of animals and PPVT raw scores far beyond those of our 5-year-olds (indeed, those typical of 12-year-olds). One was able to discuss with erudition the funerary practices in the worlds' religions yet still did not understand that death is the inevitable breakdown of the bodily machine or that animals and plants, but not the sun and the wind, are alive.

The Johnson and Carey (1998) results clearly show that associative networks of factual knowledge are not sufficient to construct a vitalist biology. Conceptual change requires the construction of new conceptual primitives and thus *cannot* be the output of any learning mechanism that merely associates or combines already available concepts (cf. Carey, 2009; Kuhn, 1983, for analysis of, and evidence for, episodes of conceptual construction that involve incommensurability between successive conceptual repertoires). Johnson and Carey did not ask *why* WS impacts the capacity for conceptual change. Here we suggest that at least part of the explanation is the impairment of EF in WS, impairment that has been well established in recent studies (e.g., Rhodes et al., 2010).

To understand *why* conceptual change might draw on EF resources, one needs both an understanding of the mechanisms underlying conceptual change and an analysis of the components and structure of EF. Our current understanding of both of these is incomplete, and working out the relations between them will contribute to our understanding of each.

How Might EF Affect Conceptual Change?

On some influential analyses (e.g., Davidson et al., 2006; Miyake et al., 2000), the basic "core" functions that make up EF include WM, inhibition, and set shifting. These were the EFs tapped in our composite EF measure. There are many ways in which these core functions—and others as well—could play a role in conceptual change. Consider the following:

1. *Comprehension monitoring and explanation seeking*, processes that provide an important impetus for conceptual change, are likely to draw heavily on EF. Noticing contradictions in one's thinking—predictions from one's current explanatory frameworks that are falsified by new data, or inconsistencies that arise from undifferentiated concepts (as in the child who wondered how statues could be visible yet not alive)—motivates the conceptual work of building new theories, new explanatory structures,. Interestingly, children with WS (who fail to construct a vitalist biology) do not ask explanation-seeking *why* questions. They would never wonder why statues can be visible in spite of not being alive. Moreover, there is positive evidence of the connection between explanatory reasoning and conceptual change in healthy children: Conceptual change curricula in school settings are more successful when they engage students in articulating and evaluating explicit explanations, including noting where current explanations fail (Carey, 2009). Monitoring one's concepts and resolving interference (such as that created when predictions fail) are key goals of the EF system (Hasher & Zacks, 1988). Engaging in such reasoning requires holding beliefs in WM, flexibly evaluating them with respect to shifting contexts, and then inhibiting the inappropriate response, precisely the processes identified as core EFs.

2. *New data structures*: Some analyses of EF development emphasize the role of increasingly powerful data structures in allowing for flexible behavior. For Zelazo (e.g., Zelazo, Muller, Frye, & Marcovitch, 2003), whose analyses are particularly relevant to logical and rule-following capabilities, EF development occurs as children become able to represent and act on increasingly complex hierarchical rules (e.g., if it is the shape game, then the square ones go in the left bucket and the round ones go in the right bucket). This development is driven by children's increasing ability to reflect on the rules they use and thus formulate increasingly higher order rule structures. In turn, better rule following allows children to better govern their behavior.

A related framework—one particularly relevant to analyses of conceptual change—emphasizes the role of developing EF in the creation of *active, abstract, symbolic, categorical, and relational representa-*

tions (Morton & Munakata, 2002; Munakata, Snyder, & Chatham, 2012). In this view, development of the prefrontal cortex allows increasingly abstract representations to be maintained actively in WM. These active, abstract representations allow children to overcome responses based on latent representations of habitual or previously learned behaviors. Abstract representations also allow for flexible, self-directed behavior by reducing selection demands (e.g., using abstract subcategories such as "pets" to organize responses in an animal-naming task).

Evidence that EF development allows new types of data structures to be represented in WM supports the claim that EF may play a role in conceptual change. Theory changes certainly can involve the representation of increasingly abstract categories (in vitalist biology, for example, the representation *living thing* is surely more abstract than *animal* or *plant*). Moreover, the bootstrapping processes that we believe to be central to conceptual change often involve *relational structures*, such as analogical mappings.

It is important to emphasize, however, that conceptual change goes beyond such changes in data structures available for mental representation. Some important types of conceptual change do not appear to involve abstraction at all. The concept *animal* changes its essence from *agency* to *life*—but the concept *agent* is not less abstract than the concept *life*. As noted above, such theory changes rely on comprehension monitoring, explanation seeking, and bootstrapping—processes we believe require EF. If this is right, then the ability to represent abstract and hierarchical data structures cannot exhaust the relation between EF and conceptual change. Our focus, then, leads us to analyses that are complementary to those of Zelazo and Munakata. For us, what is at issue are not the formal properties of the data structures themselves but the processes through which particular ones arise, and the processes through which they get their meaning.

3. *Bootstrapping* processes are implicated in every episode of conceptual change yet studied. Carey (2009) describes the role of what she calls "Quinian bootstrapping" by children as they construct number concepts (integers and rational number) and an intuitive theory of matter in which material entities are differentiated from nonmaterial physical entities and in which weight is differentiated from density. Carey also documents the role of Quinian bootstrapping in the history of science, illustrating by describing its role in the conceptual breakthroughs of Darwin and Maxwell.

In Quinian bootstrapping, a placeholder structure is formulated in natural or mathematical language, but not yet mapped onto the concepts these external symbols will come to represent, because the learner does not yet have those concepts. The symbols in the placeholder structure gain meaning through active modeling processes (analogies, limiting case analyses, inductive inferences actively evaluated with respect to other information). There is good reason to suppose that this type of bootstrapping—involving the use of active meaning-making modeling processes that result in conceptual change—makes very high demands on EF. Inagaki and Hatano (2002) provide an example of such bootstrapping during the construction of a vitalist biology. They propose that unifying animals and plants into the single category of living things—a conceptual change at the very heart of vitalist biology—involves drawing an analogy between animals and plants. This analogy rests on children's coming to know that water and food are necessary for growth and health of animals, on the one hand, and water and fertilizer are necessary for growth and health of plants, on the other. This is a far analogy: Obviously plants do not have mouths and do not eat in the same sense animals do, and animals are not grown in the same sense plants are. Such analogies play an important role in all bootstrapping episodes, and making them requires constructing new relational structures, abstraction, and active comprehension monitoring. These, in turn, draw on the WM, response selection, and set-shifting resources of EF.

Other Possible Interpretations of the Correlations Found Here

We have argued above that the processes underlying conceptual change are likely to draw heavily on EF and that this is one reason that EF predicts school success at every age tested. That is, we find it likely that EF is needed for both the construction and expression of theoretical knowledge. However, all we have so far is a correlation between measures of EF capacity and vitalist biology. Is it possible that building a vitalist biology drives EF development? Alternatively, might some common factor underlie the variation we found in both vitalist biology and EF? Let us take these two possibilities in turn.

With respect to the first possibility: EF is surprisingly malleable; indeed, it has been described a trainable organ (e.g., Diamond et al., 2007), much like a muscle. It is possible that every episode of conceptual change the child undergoes strengthens

EF as the child exercises it. Conceptual change occurs throughout the lifetime, and children are engaged in many episodes of it during the preschool and school years (e.g., in mathematical concepts, moral concepts, social concepts, and scientific concepts; see Carey, 2009, for a review). Thus, while we find it unlikely that creating a vitalist biology is *uniquely* important to developing EF (as may be the case with the developing theory of mind), engaging in theory change might well play a role in doing so.

With respect to the second possibility, that some common factor is responsible for the variation in both EF and vitalist biology, the most likely possibility is some measure of IQ. This is difficult to evaluate, however, because there is currently no agreement on whether it is even possible to draw a line between EF and other aspects of fluid IQ. On some analyses (e.g., Duncan, Emslie, Williams, Johnson, & Freer, 1996), both constructs reflect common frontal lobe function, whereas others point out that measures of EF can be statistically separated from measures of IQ (Friedman et al., 2006). In this preliminary study, where we measured only a few core components of EF, we do not know which aspects of EF (or even perhaps fluid IQ correlated with, but separable from, EF) underlie the relation we find here.

Next Steps

Further studies are called for that will address the many open questions raised by these initial findings. One way to explore the direction of influence, as well whether EF plays a role in the construction of theoretical understanding, is to carry out vitalist biology training studies. These studies will investigate whether individual differences in EF affect success of the biology training (after controlling for age, factual knowledge, and verbal IQ) and whether biology training affects EF performance.

In addition, we will take advantage of the fact that EF has recently been discovered to be a trainable resource. A year-long preschool and elementary school intervention, Tools of the Mind, has been shown to greatly increase performance on the very measures we used in our EF battery, even though the curriculum includes no practice on EF measurement tasks themselves (Diamond et al., 2007). If EF plays an important role in the construction of a vitalist biology, then children in an EF-enhanced curriculum should benefit more than children in a standard curriculum from explanation-based, bootstrapping curricula designed to foster the development of a vitalist biology.

EF is also a temporarily depletable resource in adults (e.g., Baumeister, Bratslavsky, Muraven, & Tice, 1998), and research from our lab has recently shown this to be so for 4- and 5-year-old children (Powell & Carey, 2013). Biology training studies will compare the degree of theory building in a group of children who underwent EF depletion before each biology training session with a group of controls who did not, providing a direct test of the hypothesis that EF resources are drawn upon in the construction of theoretical knowledge.

Our WM tasks did not yield informative data, although the mixed conditions of H&F and FF make demands on WM. Future studies will use more extensive measures of WM, especially measures of its active, *updating* function. Confirmatory factor analysis of EF in 5- to 7-year-old children has found that such a factor is separable from a combined set-shifting and inhibition factors (Van der Ven, Kroesbergen, Boom, & Leseman, 2012). Our goal will be to explore which aspects of EF are particularly implicated in the construction or expression of biological knowledge.

As noted above, conceptual change is hard work, likely to draw on multiple aspects of domain general processing. Specifying these processes, the precise roles they play, and the ways in which they work together to effect conceptual change will require a great deal of further research. Such research will require a much larger battery of EF tasks, especially WM tasks, as well as tasks tapping aspects of IQ that are at least partly separable from EF, and better measures of factual knowledge. Clearly, we are at the very beginning of our exploration.

References

Barrett, C. (1999) *Human cognitive adaptations to predators and prey*. Doctoral dissertation, Department of Anthropology, University of California, Santa Barbara, CA.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*, 1252–1265.

Belleville, S., Chertkow, H., & Gauthier, S. (2007). Working memory and control of attention in persons with Alzheimer's disease and mild cognitive impairment. *Neuropsychology*, *21*, 458–469.

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*, 647–663.

Brent, S. B., & Speece, M. W. (1993). “Adult” conceptualization of the irreversibility: Implications for the development of the concept of death. *Death Studies*, *17*, 203–224.
Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.
Carey, S. (2000). Science education as conceptual change. *Journal of Applied Developmental Psychology*, *21*, 13–19.
Carey, S. (2009). *The origin of concepts*. New York: Oxford University Press.
Carlson, S., & Moses, L. (2001). Individual differences in inhibitory control and children’s theory of mind. *Child Development*, *72*, 1032–1053.
Cepeda, N. J., Kramer, A. F., & Gonzalez de Sather, J. C. M. (2001). Changes in executive control across the life span: Examination of task-switching performance. *Developmental Psychology*, *37*, 715–730.
Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task-switching. *Neuropsychologia*, *44*, 2037–2078.
Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review*, *12*, 45–75.
Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science*, *318*, 1387–1388.
Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology*, *30*, 257–303.
Fodor, J. A. (1992). A theory of the child’s theory of mind. *Cognition*, *44*, 283–296.
Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*, 172–179.
Gathercole, S. E., Tiffany, C., Briscoe, J., & Thorn, A. S. C., & ALSPAC Team. (2005). Developmental consequences of poor phonological short-term memory function in childhood: A longitudinal study. *Journal of Child Psychology and Psychiatry*, *46*, 598–611.
Goldberg, R. F., & Thompson-Schill, S. L. (2009). Developmental “roots” in mature biological knowledge. *Psychological Science*, *20*, 480–487.
Harris, P. L., & Giménez, M. (2005). Children’s acceptance of conflicting testimony: The case of death. *Journal of Cognition and Culture*, *5*, 143–164.
Hasher, L., & Zacks, R. (1988). Working memory, comprehension and aging: A review and a new view. *The Psychology of Learning and Motivation*, *22*, 193–225.
Hatano, G., & Inagaki, K. (1994). Young children’s naive theory of biology. *Cognition*, *50*, 171–188.
Hickling, A. K., & Wellman, H. M. (2001). The emergence of children’s causal explanations and theories: Evidence from everyday conversation. *Developmental Psychology*, *37*, 668–683.
Hodapp, A. F., & Gerken, K. C. (1999). Correlations between scores for Peabody Picture Vocabulary Test–III and the Wechsler Intelligence Scale for Children–III. *Psychological Reports*, *84*, 1139–1142.
Inagaki, K., & Hatano, G. (2002). *Young children’s naïve thinking about the biological world*. New York: Psychology Press.
Johnson, S. C., & Carey, S. (1998). Knowledge enrichment and conceptual change in folk biology: Evidence from people with Williams syndrome. *Cognitive Psychology*, *37*, 156–200.
Kharitonova, M., & Munakata, Y. (2011). The role of representations in executive function: Investigating a developmental link between flexibility and abstraction. *Frontiers in Psychology*, *2*, 347.
Kuhn, T. S. (1983). Commensurability, communicability, comparability. In P. D. Asquith & T. Nickles (Eds.), *PSA 1982* (Vol. 2, pp. 669–688). East Lansing, MI: Philosophy of Science Association.
Luciana, M., & Nelson, C. A. (1998). The functional emergence of prefrontally-guided working memory systems in four- to eight- year old children. *Neuropsychologia*, *36*, 273–293.
Markwardt, F. C., Jr. (1997). *Peabody Individual Achievement Test–Revised/Normative Update*. Bloomington, MN: Pearson Assessments.
Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.
Morton, J. B., & Munakata, Y. (2002). Active versus latent representations: A neural network model of perseveration, dissociation, and decalage. *Developmental Psychobiology*, *40*, 255–265.
Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012). Developing cognitive control: Three key transitions. *Current Directions in Psychological Science*, *21*, 71–77.
Piaget, J. (1929). *The child’s conception of the world*. London: Routledge and Kegan Paul.
Powell, L., & Carey, S. (2013). *Executive function depletion in children and its impact on theory of mind*. Manuscript under review.
Rhodes, S. M., Riby, D. M., Park, J., Fraser, E., & Campbell, L. E. (2010). Executive neuropsychological functioning in individuals with Williams syndrome. *Neuropsychologia*, *48*, 1216–1226.
Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind. *Psychological Science*, *17*, 74–81.
Scholl, B. J., & Leslie, A. M. (1999). Modularity, development, and “theory of mind.” *Mind and Language*, *14*, 131–153.
Slaughter, V., Jaakkola, K., & Carey, S. (1999). Constructing a coherent theory: Children’s biological understanding

of life and death. In M. Siegel & C. Peterson (Eds.), *Children's understanding of biology and health* (pp. 71–98). Cambridge, UK: Cambridge University Press.

Slaughter, V., & Lyons, M. (2003). Learning about life and death in early childhood. *Cognitive Psychology*, *46*, 1–30.

Snyder, H. R., & Munakata, Y. (2010). Becoming self-directed: Abstract representations support endogenous flexibility in children. *Cognition*, *116*, 155–167.

Van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2012). The structure of executive functions in children: A closer examination of inhibition, shifting, and updating. *British Journal of Developmental Psychology*. *31*, 70–87. doi:10.1111/j.2044-835X.2012.02079.x

Zaitchik, D., & Solomon, G. (2008a). Animist thinking in the elderly and in patients with Alzheimer's disease. *Cognitive Neuropsychology*, *25*, 27–37.

Zaitchik, D., & Solomon, G. (2008b). Inhibitory mechanisms and impairment in domain-specific reasoning: Studies of healthy elderly adults and patients with Alzheimer's disease. *Proceedings of the Cognitive Science Society*, *30*, 141.

Zaitchik, D., & Solomon, G. (2009). Conservation of species, volume, and belief in patients with Alzheimer's disease: The issue of domain specificity and conceptual impairment. *Cognitive Neuropsychology*, *26*, 511–526.

Zelazo, P. D., & Frye, D. (1997). Cognitive complexity and control: A theory of the development of deliberate reasoning and intentional action. In M. Stamenov (Ed.), *Language structure, discourse, and the access to consciousness* (pp. 113–153). Amsterdam: John Benjamins.

Zelazo, P. D., Muller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function in early childhood. *Monographs of the Society for Research on Child Development*, *68*(Serial No. 274).

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

Appendix S1. Biology Interviews With Detailed Scoring Criteria the Animism Interview.