Cognition 195 (2020) 104090

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Specifying the domain-general resources that contribute to conceptual construction: Evidence from the child's acquisition of vitalist biology

Nathan Tardiff[a,*], Igor Bascandziev[b], Susan Carey[c], Deborah Zaitchik[c,d,e]

[a] Department of Psychology, University of Pennsylvania, Philadelphia, PA, United States
[b] Department of Psychology, Reed College, Portland, OR, United States
[c] Department of Psychology, Harvard University, Cambridge, MA, United States
[d] Department of Psychiatry, Harvard Medical School, Boston, MA, United States
[e] Department of Psychiatry, Massachusetts General Hospital, Boston, MA, United States

ARTICLE INFO

Keywords:
Knowledge enrichment
Conceptual change
Conceptual construction
Folkbiology
Executive function

ABSTRACT

There are two dissociable processes that underlie knowledge acquisition: *knowledge enrichment*, which involves learning information that can be represented with one's current conceptual repertoire; and *conceptual construction,* which involves acquiring knowledge that can only be represented in terms of concepts one does not yet possess. Theory changes involving conceptual change require conceptual construction. The cognitive mechanisms underlying conceptual change are still poorly understood, though executive function capacities have been implicated. The present study concerns the domain-general resources drawn upon in one well-studied case of the construction of a new framework theory in early childhood: the framework theory of vitalist biology, the ontogenetically earliest theory in which the concepts *life* and *death* come to have biological content shared with adults. Eighty-three five- and six-year-old children were tested on a battery of tasks that probe central concepts of the vitalist theory, as well as on a battery of tests of domain-general capacities that may be implicated in development in this domain, including measures of knowledge enrichment, executive function, and fluid IQ. With variance in accumulated knowledge and in knowledge enrichment capacity controlled, two specific executive functions, shifting and inhibition, predicted children's progress in constructing the vitalist theory. In contrast, working memory and fluid IQ were not associated with the acquisition of vitalist biology. These results provide further evidence for the distinction between knowledge enrichment and conceptual construction and impose new constraints on accounts of the mechanisms underlying conceptual construction in this domain.

1. Introduction

Progress in understanding conceptual development requires characterizing the learning mechanisms and cognitive resources involved in knowledge acquisition. One suite of domain-general cognitive resources implicated in conceptual development are the executive functions (EF), which have been shown in many studies to predict academic achievement (e.g., Best, Miller, & Naglieri, 2011; Blair & Diamond, 2008; Blair & Razza, 2007; Bull, Espy, & Wiebe, 2008; Gathercole, Tiffany, Briscoe, Thorn, & ALSPAC Team, 2005). There are at least three different, but not mutually exclusive, accounts for this relationship (see Carey, Zaitchik, & Bascandziev, 2015). First, children with stronger EFs can better maintain focused attention; this leads to higher-quality input, which provides a better basis for conceptual progress. Second, EFs are important for online information processing, so children with stronger EFs can better express their existing conceptual knowledge. Finally, EFs support the learning mechanisms involved in building new theoretical and conceptual structures, so children with stronger EFs acquire conceptual knowledge more effectively.

Several recent studies have focused on the role of EF in cases of conceptual construction specifically. In preschool children, EF is involved in the acquisition of concepts of integers (Geary, vanMarle, Chu, Hoard, & Nugent, 2019), theory of mind (Benson, Sabbagh, Carlson, & Zelazo, 2013), and intuitive physics (Bascandziev, Powell, Harris, & Carey, 2016). In middle school children, it supports conceptual change across several domains of science and math (chemistry, physics, biology, rational numbers; Vosniadou et al., 2015, 2018), as well as linear representations of number line estimates (Laski & Dulaney, 2015). The present study concerns the role of EF in the child's acquisition of a vitalist theory of biology (Bascandziev, Tardiff, Zaitchik, &

* Corresponding author at: Department of Psychology, University of Pennsylvania, Stephen A. Levin Building, 425 S. University Ave., Philadelphia, PA 19104, United States.

E-mail address: ntardiff@sas.upenn.edu (N. Tardiff).

https://doi.org/10.1016/j.cognition.2019.104090
Received 24 July 2018; Received in revised form 1 October 2019; Accepted 3 October 2019
Available online 18 November 2019

Carey, 2018; Tardiff, Bascandziev, Sandor, Carey, & Zaitchik, 2017; Zaitchik, Iqbal, & Carey, 2013).

Despite progress in linking EF to conceptual construction, there is an important problem in the current state of the literature: in all the case studies above, the specific role of EFs is not well understood. To make real progress in this area, researchers must identify the specific EFs—specific relative to other EFs and specific relative to other domain general constructs—that play a role in conceptual construction. Moreover, the observed relations between predictor and outcome variables must be shown to be replicable, using the same tasks to measure the same constructs in the same age groups. There are several reasons why isolating domain-general mechanisms will be extremely difficult without well-replicated measures and careful controls for alternative constructs. First, these mechanisms partially share cognitive and neural substrates (Friedman & Miyake, 2017). Second, no task purely measures a single construct (Miyake et al., 2000). Third, the EFs develop and differentiate over time (Davidson, Amso, Anderson, & Diamond, 2006; Lee, Bull, & Ho, 2013), so measures appropriate in older children may be inappropriate in preschoolers and vice versa. Identifying the resources implicated in conceptual construction must be done with great care if it is to inform our understanding of the learning mechanisms involved.

The present study has two overarching goals. First, we seek to replicate the findings reported in Bascandziev et al. (2018) and Zaitchik et al. (2013) that measures of exogenous shifting/inhibition predict progress in constructing vitalist biology, controlling for age and receptive vocabulary. We also extend these basic findings by controlling for factual knowledge as well as receptive vocabulary. This will provide further evidence that the effects of shifting/inhibition on the construction of vitalism are robust and additional data relevant to establishing the effect size of the relation. Second, the main goal of this study is to test the hypothesis that shifting/inhibition is *differentially* involved in this construction, rather than other domain-general constructs with which it covaries, namely working memory and fluid IQ. No previous studies have tested whether working memory and fluid IQ, both of which we might expect to be implicated in conceptual change, are predictive of progress in constructing vitalist biology. It is therefore an open question whether controlling for them would diminish or even eliminate the effect of shifting/inhibition. Additionally, we better fractionate the shifting mechanisms themselves by testing whether exogenous shifting/inhibition and endogenous shifting each predicts unique variance in vitalism. We do not, however, attempt to separate exogenous shifting and inhibition, as they may not be differentiated in young children (Lee et al., 2013; van der Ven, Kroesbergen, Boom, & Leseman, 2013).

In what follows, we first characterize the distinction between conceptual construction and knowledge enrichment. Next, we describe the conceptual construction of a biological theory of vitalism and previous evidence for a role of EFs in this construction. Finally, we further detail the specific goals of the present study.

1.1. *Conceptual change versus knowledge enrichment: the case of vitalist biology*

A wealth of research suggests there are two very different types of knowledge acquisition, distinguished by whether they require knowledge enrichment alone, or conceptual construction as well (see Carey, 2009, for review). *Knowledge enrichment* consists of adding new words or propositions to long-term memory that are representable in terms of one's current conceptual repertoire. For example, learning the generic fact "crickets' ears are on their legs" is an example of knowledge enrichment, so long as one already knows the words for the constituent concepts and can parse the sentence correctly. So too is adding a lexical entry for "wrench" upon first seeing one and being told its label, so long as one already has the superordinate concept *tool*. This type of knowledge is generally "fast-mapped"—easily learned from one or two encounters and remembered weeks later (Bascandziev et al., 2018; Carey, 1978, 2015; Markson & Bloom, 1997).

In contrast to knowledge enrichment, *conceptual construction* requires mastering knowledge that cannot be understood in terms of one's current concepts (Carey, 2009). Telling a child, "Gold is the element with atomic number 79" cannot lead to knowledge shared with a chemist, for the child does not have the concepts *element* and *atomic number*; indeed, in the history of science, these concepts were not constructed until the 19th century. Theory changes involving incommensurabilities (i.e., involving conceptual change) always require conceptual construction. Conceptual change displays a very different time course from knowledge enrichment—it is generally quite difficult and slow. In the development of a single child, the construction of a new theory (Theory 2) that is incommensurable with a currently held theory (Theory 1) might take years of exposure or training[1]. Moreover, there are no guarantees. Conceptual change often fails to occur despite years of explicit tutorial in school (Carey, 2009; Clement, 1982; McCloskey, Caramazza, & Green, 1980; McCloskey, 1983; Shtulman, 2006).

One well-documented case of conceptual change is the construction of the vitalist theory of biology, which is manifest in the biological reasoning of adults in many cultures. Vitalism is a framework theory, and as such supports many different specific instantiations consistent with it. The central tenet of all vitalist theories is that substances in the outside world—air, water, and food—must enter the body and move through it, bringing vital energy/substances to all its parts, in order to maintain life and health (Carey, 1985; Contento, 1981; Hatano & Inagaki, 1994; Inagaki & Hatano, 1993, 2002). The construction of this theory is difficult and protracted: it begins at roughly age 5 or 6 and is not complete until age 10–12, even though it is a central focus of the elementary school curriculum (Carey, 1985; Inagaki & Hatano, 2002). Once constructed, the vitalist framework continues to underlie reasoning well into old age, even in the face of cognitive declines associated with normal aging (Tardiff et al., 2017).

As in other instances of conceptual change, the difficulty of acquiring vitalist biology derives in part from the fact that the theory one is trying to learn is articulated in a conceptual vocabulary that is incommensurable with one's current conceptual repertoire. Furthermore, the acquisition of vitalist concepts occurs in the face of substantial interference from an earlier and still potent theory of animals as intentional and causal agents, rather than as living beings, which has its roots in infancy (Gergely & Csibra, 2003; Luo & Baillargeon, 2005; Onishi & Baillargeon, 2005; Saxe, Tenenbaum, & Carey, 2005; Woodward, 1998; Woodward & Somerville, 2000; see Carey, 2009, for review). The construction of the vitalist theory thus demands conceptual change at the level of individual concepts. Specifically, the concepts *alive* and *dead*, as understood in the agency theory to refer to the activity or existence of intentional agents, have no role to play in vitalist biology. To have a role in vitalist biology, *alive* must be differentiated from *moving, active, real, existing*; similarly, *dead* must be differentiated from *nonmoving, inactive, not real.* Moreover, plants and animals must be coalesced into a new ontological category, *living things*. With the construction of vitalist biology, the conceptual core of the concept *animal* is changed as well, from *causal/intentional being* to *biological organism*. In constructing vitalist biology, then, children have created a new conceptual structure, a new theory—one that is couched

[1] The theories in play here are what Wellman and Gelman (1992) called *framework theories* (see also Vosniadou, 2019), knowledge structures that arise in childhood that embody ontological commitments and explanatory machinery. They are abstract and skeletal relative to fully explicit, mathematized, detailed scientific theories, and as we show below, they do not exhaust the factual knowledge children represent. Nonetheless, they embody the inferential machinery that allows us to characterize the concepts embedded within them, which further allows us to identify episodes of conceptual construction.

in an interrelated set of concepts that were not previously available.

1.2. Domain-general learning mechanisms

Since knowledge enrichment and conceptual change effect distinct types of change to the conceptual system, it stands to reason that distinct learning mechanisms must underlie each. That is, we should expect that 1) mechanisms supporting conceptual change should be at least partially dissociable from mechanisms supporting knowledge enrichment; and 2) if constructing a new theory requires mechanisms beyond those that support knowledge enrichment, then merely learning new facts will not be sufficient for building the new theory. This is true no matter how necessary these facts may be to inform and motivate conceptual change.

As noted above, the EFs are a candidate set of mechanisms implicated in this process. This association is sensible on its face, given that conceptual change often involves resolving conflict between competing conceptual structures. In the domain of vitalist biology in particular, there is mounting evidence that construction of this theory draws on shifting and inhibition (Bascandziev et al., 2018; Zaitchik et al., 2013). Furthermore, these mechanisms are dissociable from those supporting knowledge enrichment. The relationship between progress in constructing vitalist biology and shifting/inhibition holds even after controlling for both age and receptive vocabulary (Zaitchik et al., 2013), the latter of which reflects variance in both accumulated fast-mapped knowledge and in the capacity for knowledge enrichment (see Bascandziev et al., 2018; Johnson & Carey, 1998). Even stronger evidence for a dissociation between these mechanisms is provided by a recent training study (Bascandziev et al., 2018). After training, children's progress in learning isolated factual knowledge about animals was predicted by receptive vocabulary but not shifting/inhibition, while progress in constructing vitalist biology was predicted by shifting/inhibition but not receptive vocabulary. This double-dissociation between measures predictive of knowledge enrichment and those predictive of conceptual change strongly supports the distinction between these two types of knowledge acquisition.

Knowledge enrichment alone is also not sufficient to induce conceptual change in vitalist biology. Adults with Williams syndrome have a genetic form of intellectual disability, but relatively spared language and fact-learning abilities. Specifically, these adults have a relatively large lexicon and relatively large stock of generic factual knowledge (e.g., that kangaroos have pouches), compared to what would be expected from their full-scale IQ (Osório et al., 2012; Rhodes, Riby, Fraser, & Campbell, 2011). Though they clearly maintain the use of language-based domain-general mechanisms that underlie fast-mapped knowledge, they nevertheless fail to construct a vitalist biology even by adulthood, performing no better than six-year-olds on interviews tapping the animate/inanimate and alive/dead distinctions (Johnson & Carey, 1998). Apparently, the mechanisms underlying fast-mapped generic fact learning and fast-mapped lexical learning, both the output of knowledge enrichment mechanisms, cannot themselves accomplish conceptual change. Tellingly, Williams syndrome also presents with EF deficits (Osório et al., 2012; Rhodes et al., 2011).

While these prior studies strongly argue for a distinction between the mechanisms underlying knowledge enrichment and conceptual change in the domain of vitalist biology, theoretically unresolved questions remain regarding the specificity of the role of shifting/inhibition. First, the narrow range of domain-general processes measured in these studies leaves open the possibility that alternative constructs could better account for their findings. Consider that conceptual change often demands extensive conceptual modeling—including the deployment of structure mapping, limiting case analyses, thought experimentation, and inductive inference—in service of the creation of new representational primitives (Carey, 2009; Gentner, 2002; Nersessian, 1992). Indeed, in the case of vitalist biology—at least in industrialized societies—analogies between people and other animals, and between animals and plants, have been implicated in the process (Carey, 1985; Inagaki & Hatano, 2002). Given that such conceptual modeling draws on analogical structure mapping, it might well draw on the cognitive resources tapped by measures of fluid IQ such as progressive matrices (e.g., Crone et al., 2009). Furthermore, it has been proposed that the development of children's analogical reasoning is in turn subserved by the development of working memory and inhibitory control (Crone et al., 2009; Halford, 1993; Richland, Morrison, & Holyoak, 2006). On these grounds, it is plausible that variance in the construction of vitalist biology previously attributed to shifting/inhibition could be better explained by variance in working memory or fluid IQ.

A second question concerns the relative contributions of exogenous (cued) and endogenous (self-directed) shifting. Zaitchik et al. (2013) used measures of exogeneous shifting/inhibition only, while Bascandziev et al. (2018) used a composite measure of endogenous shifting and exogenous shifting/inhibition. While both measures were significant predictors of progress in constructing vitalist biology, it is possible that these two types of shifting may contribute independent variance. Exogenous and endogenous shifting are related but separable in young children (Barker et al., 2014; Snyder & Munakata, 2010), and it is conceivable that the resolution of conceptual inconsistencies may require endogenous shifting to a greater extent. Consider that a child building a new conceptual structure must often select representations or hypotheses from among multiple competing alternatives without recourse to specific instruction, and that such instruction when available may still require considerable endogenously-generated processing if the language of that instruction is incommensurable with the child's current conceptual understanding.

An additional limitation of prior work is that it relied solely on tests of receptive vocabulary to control for accumulated factual knowledge and knowledge enrichment capacities. Vocabulary items, however, are just one type of general factual knowledge acquired during the course of schooling and beyond, and may not accurately reflect the child's broader knowledge base/knowledge acquisition abilities (Ackerman, 2000; Schipolowski, Wilhelm, & Schroeders, 2014). It is therefore possible that the relationship between shifting/inhibition and vitalist biology would not survive the inclusion of more robust measures of factual knowledge. If this were true, it would call into question the distinction between knowledge enrichment and conceptual change in typically developing children.

The present study addresses these limitations. In addition to seeking to replicate the previous finding that exogenous shifting/inhibition predicts progress in constructing vitalist biology (Zaitchik et al., 2013), this study will assess the effects of other resources that may plausibly be implicated. These include: endogenous shifting, working memory, analogical reasoning/fluid IQ, and general factual knowledge. This expanded set of measures will allow us to better identify the *specific* domain-general mechanisms implicated in this conceptual construction.

2. Methods

2.1. Participants

Participants were 83 five- and six-year-old children (40 girls; M_{age} = 76.6; SD = 4.44; range = 66–83 months). Children of this age group were chosen because they are at the very beginning of constructing a vitalist biology and show a great deal of variance in the progress they have made. Volunteers were recruited using public birth records from Cambridge, MA and surrounding towns. They were primarily from middle-class families with a stay-at-home parent. Ethnicity in our participant population as a whole is approximately 70% non-Hispanic White and 9% Hispanic, with the remaining 21% of the population comprising African American, Asian, Native American, and Native Hawaiian participants. All children had English as their primary language. Children were individually tested in the Harvard Laboratory for Developmental Studies. Each child received a small toy as compensation and parents were given $5 toward travel expenses.

2.2. Procedures

The study consisted of two sessions, administered no more than one month apart[2] . In keeping with our individual differences design, the tasks were presented in the same order to all participants—Session 1: Animism interview, Death interview, Body Parts interview, Factual Knowledge; Session 2: Hearts & Flowers, Flanker, Counting Recall, Verbal Fluency, Matrices, Receptive Vocabulary.

2.2.1. Outcome measures: the Vitalism battery

The Vitalism battery consists of three interviews that assess the child's understanding of several central concepts within vitalist biology, including concepts of life, death, and bodily function. Aside from a few minor tweaks to increase clarity and brevity, it is identical to vitalism batteries used previously (Bascandziev et al., 2018; Tardiff et al., 2017; Zaitchik et al., 2013).

The Animism interview (Carey, 1985; Piaget, 1929) probes children's understanding of what it means to be alive and the degree to which they restrict attributions of life to entities deemed alive according to vitalist biology (animals and plants). In this task, participants are asked several open-ended questions: 1) What does it mean to be alive, to be a living thing? 2) Can you name some things that are alive, that are living things? 3) 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 items, the former including various animals and plants, and the latter ranging from apparently self-moving entities (e.g., the sun), to entities that move (e.g., a car), to entities that do something or have a function (e.g., a lamp), to inert, largely functionless entities, (e.g., a mountain). For each item 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 (Carey, 1985; Slaughter, Jaakkola, & Carey, 1999; Slaughter & Lyons, 2003) probes the understanding of death as the complete breakdown of the bodily machine and as inevitable and irreversible. It probes the differentiation of *dead* from *inanimate* (i.e., that death is end of life and not merely the opposite of life) and the differentiation of the spiritual concept of death from the biological concept. As in the Animism interview, children are asked several open-ended questions: 1) What does it mean to die? 2) Can you name some things that die? 3) What happens to a person's body when they die? 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 (Does a dead person need to: eat? pee? sleep? Does a dead person: feel bad that he died? miss his friends? think about things?). Finally, children are asked what might cause someone to die and why.

The Body Parts interview (Carey, 1985; Slaughter et al., 1999; Slaughter & Lyons, 2003) begins with questions 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 didn't have it. Children are then asked why we eat food, what happens to the food we eat, and whether you need a brain to be able to eat. They are then asked why we breathe air, what happens to the air we breathe, and whether you need a brain to be able to breathe. Much of what is tested is simply factual knowledge that requires no conceptual change for its acquisition. However, this task also taps the child's understanding that interrelated bodily functions serve the vitalist goal of maintaining life.

2.2.2. Scoring of the Vitalism battery

Interviews were transcribed and coded blind with respect to participant, as well as to participants' responses on the other interviews. For full transcripts of tasks and scoring procedures, see the appendices of Bascandziev et al. (2018).

In the Animism interview, points are awarded for appropriately constraining attributions of life to people, animals, and plants, and for justifying those attributions by appealing to biological processes such as birth, growth, and death. The total score is the sum of two sub-scores: 1) the degree of animism indicated in the yes-no judgments; and 2) a qualitative scoring of responses to the introductory questions and justifications. This latter sub-score reflects the degree to which an explicit mastery of vitalist biology is expressed. In addition, positive evidence of the failure to distinguish *existence* from *life,* evidence often found in the responses of young preschoolers, results in negative points.

In the Death interview, points are awarded for indicating that death is the end of life and of all bodily and mental function, that the body decays after death, that all living things eventually die, that death is caused by the breakdown of bodily function, and that it is irreversible. Points are subtracted if the child affirms that dead people continue to have any bodily or mental processes or fail to constrain attributions of death to living things.

In the Body Parts interview, points are awarded for mapping bodily functions (e.g., digestion) that support vitalist goals (e.g., distributing vital energy/substances throughout the body in support of growth) onto body organs (e.g., stomach, intestines, circulatory system). Thus it probes for an understanding that the body is a system whose parts work together to support biological goals, including life itself.

Two independent coders scored the interviews. For each test, inter-rater agreement, calculated using the intraclass correlation coefficient, was high—Animism: ICC = .97; Death: ICC = .94; Body Parts: ICC = .98. All disagreements were resolved by discussion.

2.2.3. Predictor variables: domain-general mechanisms

2.2.3.1. The EF battery. No EF task is pure, but these tasks have been shown to differentially tap the 3 core EFs—*inhibition, shifting,* and *working memory.*

2.2.3.2. Endogenous (cued) shifting and inhibition. **Hearts & Flowers** (H &F; Davidson et al., 2006; Diamond, Barnett, Thomas, & Munro, 2007) involves two rules: 1) If a heart appears on the screen, press the button on the same side of the screen as the heart; 2) If a flower appears, press the button on the opposite side of the screen as the flower. This is a Simon task, where the child must inhibit a prepotent tendency to press the button on the same side of the screen as the stimulus. After a congruent block (all stimuli are hearts) and an incongruent block (all stimuli are flowers), the mixed block presents intermixed trials of hearts and flowers. In addition to the inhibitory demands, this task draws on shifting, as the child must switch rules flexibly in response to the cue (heart or flower). At these ages, accuracy scores, rather than reaction times, are the measure derived from the task. We analyzed the 32 mixed trials only, as Zaitchik et al. (2013) found that only the mixed block had sufficient variance in accuracy to be used in correlation analyses. The outcome measure is percentage correct (out of 32).

Flanker (FL; Diamond et al., 2007; Rueda et al., 2004) also has two rules. If the stimulus (a row of fish) is blue, press the arrow button pointing in the direction, right or left, that the middle fish is facing. If the row is pink, press the arrow pointing in the direction the outside fish are facing. In both the blue and pink conditions, there are congruent trials (all fish point in the same direction), incongruent trials (middle and outside fish point in opposite directions), and neutral trials (non-target fish point up or down, rather than right or left). The mixed condition, with intermixed pink and blue trials, is harder than the mixed condition of H&F. The greater difficulty here is due to the demand to shift attention between the middle and the outside fish in addition to managing interference from non-target fish. As in H&F, the shifting required is cued, in this case by the color of the fish. The outcome measure is percentage correct (out of 44) on the mixed trials, those trials that demand both inhibition and shifting.

[2] There were four exceptions due to scheduling issues.

2.2.3.3. Endogenous shifting. **Verbal Fluency (VF).** In this task, children are given 60 s to name as many members of a category (in the first test, animals; in the second test, foods) as they can, without repetition. For each test, the score is the total number of items named, minus errors and repetitions. VF is a complex EF task that appears to draw on multiple EFs as well as possessing unique variance (Gustavson et al., 2019). Good performance is generally thought to rely on two primary components: *clustering* of semantically related words by subcategory (e.g., for animals: barn animals, pets, jungle animals, fish, etc.) and *switching* between clusters when the prior cluster is no longer productive (Troyer, Moscovitch, & Winocur, 1997). Switching in particular is thought to rely on prefrontally-mediated executive processes. Supporting this claim, fluency performance is impaired in adult frontal lobe patients (Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998; see Henry & Crawford, 2004, for review) and in young children with frontal lobe epilepsy (Chieffo et al., 2011; Hernandez et al., 2002; cf. Riva et al., 2005). As with other EF measures, performance on fluency tasks continues to improve throughout childhood (Kavé, Kigel, & Kochva, 2008; Korkman, Kemp, & Kirk, 2001; Riva, Nichelli, & Devoti, 2000; Sauzéon, Lestage, Raboutet, N'Kaoua, & Claverie, 2004). Indeed, children in the age range of the present study (5–7 years old) already exhibit evidence of clustering and switching, further suggesting that this task also serves as an EF measure in childhood (Barker et al., 2014; Snyder & Munakata, 2010, 2013).

Importantly, in young children VF has been shown to uniquely measure endogenous shifting (Barker et al., 2014; Snyder & Munakata, 2010). As compared to the exogenous shifting assessed in most shifting tasks, including H&F and FL in the present study, endogenous shifting may make particular demands on executive resources due to the need to resolve competition among multiple conflicting alternatives (Snyder & Munakata, 2010). For this reason, we consider VF as primarily a measure of endogenous flexibility.

2.2.3.4. Working memory capacity. **Counting Recall** is a test of working memory capacity (AWMA; Alloway, 2007), the ability to maintain and update active working memory representations even in the face of interference. In this task, the child is presented with a series of screens. Each screen shows a number of red triangles, blue triangles, red circles, and blue circles. When a screen appears, the child must count out loud all the red circles and repeat the final number out loud. In the three-screen condition, for example, the child does this with a series of three screens and is then presented with a blank screen. Now the child is asked to repeat, in order, the final number of red circles on each of the three previously viewed screens. Trials are blocked by set sizes that range from one to seven screens, with six trials per block. Four of the six trials within a block must be correct in order to proceed to the next block. The score is the total number of correct trials.

2.2.3.5. Fluid IQ. **Matrices** (KBIT-2; Kaufman & Kaufman, 2004) provides an age-appropriate standardized measure of fluid IQ. The test includes both verbal analogies presented pictorially as well as a larger set of abstract spatial/geometric reasoning items typical of tests of fluid intelligence. Testing ends when four consecutive items are answered incorrectly. The score is the total number of correct responses (i.e., total raw score).

2.2.3.6. Knowledge enrichment. **Factual Knowledge.** The Academic Knowledge test (Woodcock-Johnson III; Woodcock, McGrew, & Mather, 2001) provides a general measure of factual knowledge. Children are asked questions that draw on their knowledge of science, social studies, and the humanities. Each subsection is terminated when children have answered all questions or when they have incorrectly answered three questions in a row. Questions range widely in content and difficulty (e.g., What animal quacks? What is it called when garbage and wastes are dumped into our lakes and rivers?). Learning the answers to these questions requires no conceptual construction because children already have the necessary concepts to represent these facts (animal kinds and their characteristic sounds for the former question; garbage and bodies of water for forming the concept *pollution* needed for the latter). As we are interested in the child's accumulated knowledge of generic facts whose learning draws only on mechanisms underlying knowledge enrichment, the dependent measure is total number of correct responses (i.e., total raw score).

Receptive Vocabulary (RV). The Verbal Knowledge test (KBIT-2; Kaufman & Kaufman, 2004) is, primarily, a test of receptive vocabulary. On most trials, children are asked to point to the picture (out of a set of six pictures) that illustrates the meaning of an orally presented word (e.g., *whispering, furry, precipitation*). Occasionally, a description is orally presented, rather than a single word (e.g., What tells you how much something weighs?), and thus probes generic factual knowledge. There are 60 trials in all, but administration is halted after four consecutive incorrect answers. The child's score is the total number of pictures correctly pointed to (i.e., total raw score).

We take RV to be a measure of knowledge enrichment because much lexical learning—including the words on this test—occurs via fast mapping, as mentioned above (Carey & Bartlett, 1978; Markson & Bloom, 1997). Accumulated vocabulary, like accumulated fast-mapped generic knowledge, is a function of the richness of the input as well as the capacity of the learning mechanisms that support fast mapping. The two knowledge enrichment tasks should therefore be highly correlated with each other, even controlling for age and EF.

3. Results

The results are analyzed in three steps. First, we discuss the Vitalism battery: scores on the Animism, Death, and Body Parts interviews, and the relations among them. Then we discuss the predictor variables: measures of knowledge enrichment (Factual Knowledge, Receptive Vocabulary), EF (Hearts & Flowers, Flanker, Verbal Fluency, Counting Recall), and fluid IQ (Matrices), and the relations among them. Finally, we analyze the relations between the predictor and outcome variables.

3.1. The Vitalism battery

As in our previous studies (Bascandziev et al., 2018; Zaitchik et al., 2013), children showed considerable variance on each task, with some children demonstrating almost no vitalist biology and others explicitly articulating the bare bones of a vitalist understanding of life, death, and bodily function that is tapped in these interviews (Table 1).

Although the three interviews probe different aspects of the vitalist theory, bivariate correlation analyses show that all pairs of tasks were significantly correlated (Animism/Death: $r(81) = .44$, $p < .001$; Animism/Body Parts: $r(81) = .28$, $p = .01$; Death/Body Parts: $r(81) = .48$, $p < .001$). Moreover, partial correlations controlling for age yielded similar results (Animism/Death: $r(80) = .42$, $p < .001$; Animism/Body Parts: $r(80) = .25$, $p = .03$; Death/Body Parts: $r(80) = .45$, $p < .001$). Given the significant intercorrelations, the three variables were standardized (M = 0, SD = 1) and then aggregated into a single composite variable by computing an average. This variable, Composite Vitalism, was then standardized (M = 0, SD = 1) as well.

Table 1
Descriptive statistics for the Vitalism battery (n = 83).

Measure	M	SD	Min	Max	Possible Range
Animism interview	7.83	3.94	−1	18	−4 to 19
Death interview	1.65	2.35	−3	6	−9 to 8
Body Parts interview	10.66	3.96	3	20	0 to 26

Table 2
Descriptive statistics for the predictor variables (n = 83).

Measure	M	SD	Min	Max	Possible Range
EF					
H&F	0.79	0.14	0.33	1.00	0 to 1
FL	0.67	0.17	0.25	0.93	0 to 1
VF Animals	13.60	4.60	3	28	
VF Foods	12.06	4.07	4	26	
Counting Recall	13.25	4.58	6	26	0 to 42
Fluid IQ					
Matrices	22.34	6.24	13	38	0 to 46
Knowledge Enrichment					
Factual Knowledge	39.59	4.08	32	51	0 to 78
RV	23.75	4.68	16	37	0 to 60

3.2. The predictor variables

All of our predictor variables—measures of EF, fluid IQ, and Knowledge Enrichment—showed sufficient variance for correlation analysis (Table 2). All EF scores in our study are raw scores that have not been adjusted to account for the age of the participant, and some of the variance in EF scores is due to age differences. Therefore, in order to have comparable predictor variables, we used the raw scores on Matrices, Receptive Vocabulary, and Factual Knowledge as well. For those measures for which national norms are available, we note that the standardized scores were well above the national norm—Matrices (KBIT-2): 0.48 SD above the norm; RV (KBIT-2 Verbal Knowledge): 0.98 SD above the norm. On Factual Knowledge (Woodcock-Johnson III Academic Knowledge), the estimated age associated with our sample's mean score was 92 months, though their actual mean age was 77 months.

3.2.1. The EF battery

Table 3 displays the correlations among the EF measures. H&F and FL scores each reflect exogenous shifting/inhibition, and as expected, scores of these two measures were correlated. They were thus aggregated to form a composite H&F/FL score. Scores on the two VF tests, which reflect endogenous shifting, were correlated, so they were aggregated to form a composite VF score. As above, the aggregates are averages of z-scores, which were then standardized. Counting Recall, our single measure of working memory, was kept as a separate variable on theoretical grounds. Nevertheless, the fact that it was correlated with all other EF measures (except VF Animals) confirms that it is a robust measure, with enough variance to reveal relations with other variables.

3.2.2. Knowledge enrichment

Participants varied substantially in their raw scores on our two tests hypothesized to reflect knowledge enrichment (RV and Factual Knowledge; Table 2). As expected, the two measures were highly correlated ($r(81) = .70, p < .001$) and remained so when age was partialled out ($r(80) = .66, p < .001$). This finding confirms our previous use of receptive vocabulary raw scores as a measure of accumulated factual knowledge (Bascandziev et al., 2018; Johnson & Carey, 1998; Zaitchik et al., 2013) and justifies the aggregation of the two measures into a composite Knowledge Enrichment measure.

3.3. Relations among the predictor variables

Simple bivariate correlations among the predictor variables, as well as those correlations after partialling out age, show exactly the same pattern (Table 4). Counting Recall was correlated with H&F/FL and with Matrices. The first relationship makes sense, given that the mixed trials of H&F and FL demand active maintenance of multiple rules in working memory. The second relationship, reflecting an association between working memory and fluid IQ, is well-known in the adult psychometric literature (Duncan, Schramm, Thompson, & Dumontheil, 2012; Friedman et al., 2006; Harrison, Shipstead, & Engle, 2015; Wiley, Jarosz, Cushen, & Colflesh, 2011), and has also been found in children (Alloway, Gathercole, Willis, & Adams, 2004; Demetriou et al., 2014; Gray et al., 2017; see Fry & Hale, 2000, for review). Knowledge Enrichment was correlated with both VF and Matrices. The former relationship makes sense given that VF tasks draw on vocabulary, as do both RV and Factual Knowledge. The latter relationship between Knowledge Enrichment and Matrices is sensible for the same reason; the matrices measure used here included some verbal analogies, which similarly draw on lexical knowledge.

There are two important lessons from these results in the present context. First, there is sufficient variance in our predictor measures to reveal correlations among them. Second, each predictor variable demonstrates a unique pattern of relations to the others. The partial independence of these measures leaves open the possibility that only some of them will be related to vitalist biology. We now turn to addressing the central focus of this study, first asking which domain-general constructs are correlated with the development of vitalist biology, and then using multiple-regression analysis to determine which constructs predict unique variance in progress in this domain.

3.4. Relationships between the predictor variables and the Vitalism battery

3.4.1. Correlation analysis

The relationships between the predictor variables and the vitalism measures are clear and consistent (Table 5). First, replicating previous results (Bascandziev et al., 2018; Zaitchik et al., 2013), H&F/FL, VF, and Knowledge Enrichment were correlated with Composite Vitalism. With only one exception (the relation between H&F/FL and Body Parts), these variables were also correlated with each component of the Vitalism battery.

The new result is that not *all* measures of domain-general cognitive functions predict variance on the Vitalism battery. Specifically, Counting Recall was not correlated with Composite Vitalism nor with any component of the Vitalism battery. Thus, EF fractionates—some specific EFs predict variance in the progress children have made in constructing a framework vitalist biology, and others do not. Matrices was also not correlated with Composite Vitalism nor with any component of the Vitalism battery. Therefore, the relations between the EF measures and vitalism are unlikely to reflect shared variance between EF and fluid IQ, at least not as reflected in performance on Matrices.

Table 3
Correlations among the EF measures.

	H&F	FL	VF Animals	VF Foods	Counting Recall
H&F	*	**.30 (.007)**	.12 (.27)	.10 (.38)	**.23 (.03)**
FL	**.31 (.005)**	*	.10 (.38)	.18 (.11)	**.28 (.01)**
VF Animals	.13 (.26)	.04 (.70)	*	**.52 (< .001)**	.05 (.64)
VF Foods	.10 (.38)	.14 (.20)	**.50 (< .001)**	*	**.25 (.02)**
Counting Recall	**.24 (.03)**	**.24 (.03)**	.001 (.99)	**.22 (.05)**	*

Note. p-values in parentheses. Partial correlations controlling for age below the diagonal. Boldface indicates $p < .05$.

Table 4
Correlations among the predictor variables.

	H&F/FL	Verbal Fluency	Counting Recall	Knowledge Enrichment	Matrices
H&F/FL	*	.18 (.11)	**.32 (.003)**	.14 (.22)	**.26 (.02)**
Verbal Fluency	.15 (.19)	*	.17 (.12)	**.33 (.002)**	.13 (.24)
Counting Recall	**.30 (.007)**	.13 (.25)	*	.21 (.06)	**.35 (.001)**
Knowledge Enrichment	.09 (.44)	**.28 (.01)**	.14 (.19)	*	**.35 (.001)**
Matrices	**.23 (.03)**	.09 (.40)	**.33 (.003)**	**.31 (.004)**	*

Note. p-values in parentheses. Partial correlations controlling for age below the diagonal. Boldface indicates $p < .05$.

Table 5
Bivariate correlations between the vitalism measures and predictor variables.

	Animism	Death	Body Parts	Composite Vitalism
H&F/FL	**.24 (.03)**	**.24 (.02)**	.14 (.22)	**.27 (.01)**
Verbal Fluency	**.36 (.001)**	**.34 (.002)**	**.38 (< .001)**	**.46 (< .001)**
Counting Recall	.15 (.18)	.12 (.27)	.01 (.94)	.12 (.28)
Knowledge Enrichment	**.33 (.002)**	**.42 (< .001)**	**.49 (< .001)**	**.53 (< .001)**
Matrices	–.08 (.50)	.16 (.15)	.21 (.06)	.13 (.25)

Note. p-values in parentheses. Boldface indicates $p < .05$.

3.4.2. Regression analysis

Though the correlations presented above are suggestive with regard to the role of domain-general mechanisms in the construction of vitalist biology, they leave open the possibility that unique variance in this construction may be attributable to accumulated fast-mapped factual knowledge alone. We therefore ran a hierarchical regression analysis (Table 6) to test whether any of our predictor variables account for unique variance in Composite Vitalism, after controlling for age and all other predictor variables.

Inspection of Model A in Table 6, which includes only age and Knowledge Enrichment as predictor variables, shows that Knowledge Enrichment significantly predicted Composite Vitalism scores even after controlling for age. On average, children who scored 1.0 SD higher on Knowledge Enrichment scored 0.49 SD higher on Composite Vitalism. The R^2 statistic in Model A indicates that 30% of the variance in Vitalism was associated with Knowledge Enrichment and age ($F(2,80) = 16.82$, $p < .001$). It stands to reason that measures of accumulated generic factual knowledge/knowledge enrichment capacity predict progress on vitalism— children must learn at least some facts concerning the bodily machine in order to create the new knowledge structure in which concepts of life, death, and the bodily machine are interrelated in a new way.

The question is not, however, whether measures of knowledge enrichment predict the acquisition of vitalist biology; the question is whether the construction process *additionally* draws on executive mechanisms, even controlling for measures of knowledge enrichment. As evident in Model B, H&F/FL and VF each accounted for independent variance in Composite Vitalism scores (though the effect of H&F/FL was marginal), controlling for Knowledge Enrichment and age. Children who scored 1.0 SD higher on H&F/FL were, on average, 0.16 SD higher on Composite Vitalism, and children who scored 1.0 SD higher on VF were, on average, 0.30 SD higher on Composite Vitalism. The ΔR^2 statistic indicated that H&F/FL and VF together explained an additional 11% of the variance in performance on Composite Vitalism, over and above age and Knowledge Enrichment. This increase in variance was significant ($F(2,78) = 7.25$, $p = .001$)[3] .

Finally, the last question was whether Counting Recall and Matrices would independently predict variance in Vitalism, controlling for all other variables. This was explored in Model C, which showed that neither Counting Recall nor Matrices was a significant predictor, and that taken together they did not significantly increase the explanatory power of the model ($F(2,76) = 1.01$, $p = .37$). The beta values and significance levels for the other predictors remained relatively unchanged between Models B and C (save for H&F/FL, which went from trend-level to significant), further suggesting that each contributes unique variance to Composite Vitalism scores.

In sum, the results of the regression analysis support the view that in the domain of vitalist biology the processes involved in knowledge enrichment are distinct from those involved in theory construction and conceptual change. This is demonstrated most clearly by the fact that H&F/FL and VF each independently explained variance in Composite Vitalism, even after partialling out Knowledge Enrichment and all other predictor variables. Our interpretation of these results is that shifting/inhibition play an important role in the construction of a vitalist theory, and that the construction process draws on different domain-general mechanisms than does mere knowledge enrichment. These results complement prior findings of a double-dissociation between mechanisms underlying knowledge enrichment and conceptual change (Bascandziev et al., 2018), demonstrating that the association between shifting/inhibition and the construction of vitalist biology is robust, even when controlling for knowledge enrichment capacity and accumulated factual knowledge, working memory, and fluid IQ. Most importantly, neither working memory nor fluid IQ was significantly associated with variance in vitalism. We will return to implications of

Table 6
Hierarchical regression analysis predicting Composite Vitalism.

Predictor	β	*t*	R^2	ΔR^2
Model A			.30	**.30*** **
Age	0.12	1.17		
Knowledge Enrichment	**0.49**	**4.90*** **		
Model B			.41	**.11** **
Age	0.06	0.67		
Knowledge Enrichment	**0.39**	**4.04*** **		
H&F/FL	0.16	1.74~		
VF	**0.30**	**3.14** **		
Model C			.42	.01
Age	0.08	0.80		
Knowledge Enrichment	**0.43**	**4.25*** **		
H&F/FL	**0.20**	**2.07* **		
VF	**0.30**	**3.15** **		
Counting Recall	–0.06	–0.66		
Matrices	–0.10	–1.06		

Note. ~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$; Boldface indicates $p < .05$.

[3] H&F/FL and VF were also both significant predictors when individually entered into the model controlling for age and Knowledge Enrichment (H&F/FL: $\beta = 0.19$, $t(79) = 2.05$, $p = .04$; VF: $\beta = 0.32$, $t(79) = 3.35$, $p = .001$).

these findings in the Discussion.

3.5. The coherence of the Vitalism battery, revisited

In addition to the results of the regression analysis, our data provide a second source of support for the claim that the processes involved in knowledge enrichment are distinct from those involved in theory construction and conceptual change in this domain. Here our outcome measure is not Composite Vitalism, but rather the correlations among the biology tasks. These correlations reflect the internal coherence of the three tasks, no matter how far along the child has come in the conceptual construction. To appreciate the meaning of these correlations, it is important to recall that each task in the Vitalism battery probes a very different aspect of the vitalist theory. Therefore, beyond justifying the creation of the Composite Vitalism score, these measures of relatedness have theoretical significance; they support the conclusion that the concepts drawn upon in the interviews are co-constructed as related elements of a single coherent theory.

There is, however, another possible explanation for these significant correlations in which theory construction involving conceptual change plays no role. As each of the biology interviews surely taps factual knowledge, performance on each interview may reflect how much fast-mapped factual knowledge the child has accrued. To address this possibility, a final set of correlational analyses among the Vitalism interviews partialled out age and Knowledge Enrichment. The correlations between Animism and Death ($r(79) = .35, p = .002$), which reflect the interrelatedness of the concepts *life* and *death* (i.e., that death is understood as the end of life), and those between Death and Body Parts ($r(79) = .34, p = .002$), which reflect the interrelatedness of the concept *death* and the breakdown of the bodily machine (i.e., that death is understood as the result of the cessation of all bodily function), remained significant. Only the correlation between Animism and Body Parts was no longer significant ($r(79) = .14, p = .22$). This no doubt reflects the fact that many of the facts tapped on the Body Parts interview are acquired by knowledge enrichment alone, and the functioning of the bodily machine is not as important to the distinction between *animate* and *inanimate* as it is to the distinction between *alive* and *dead*.

In sum, the intercorrelations among the vitalism interviews cannot be fully explained by age and Knowledge Enrichment. Even after controlling for these variables, the interviews are still intercorrelated. These intercorrelations reflect the co-construction of the constituent elements of the vitalist theory—children who better understand the concepts *animate* and *inanimate* also better understand the concepts *alive* and *dead*, and those who better understand the concepts *alive* and *dead* also better understand the bodily processes that sustain life and whose cessation leads to death. This pattern of results provides strong evidence that the understanding of vitalism, which partly drives the intercorrelations, is separable from fast-mapped accumulated factual knowledge. This in turn suggests that this conceptual construction requires learning mechanisms that are at least partially distinct from those that support knowledge enrichment alone.

4. Discussion

This study elucidates the domain-general mechanisms underlying the development of vitalist biology with far greater specificity than previous investigations. Besides replicating the finding that exogenous shifting/inhibition predicts progress in constructing a vitalist biology (Bascandziev et al., 2018; Zaitchik et al., 2013), this study is the first to establish that 1) endogenous shifting determines independent variance in that progress, and 2) working memory and fluid IQ are completely unrelated to the early stages of acquiring vitalist biology. Additionally, by utilizing measures of both receptive vocabulary and general factual knowledge, we establish far more robustly the limits of knowledge enrichment mechanisms in accounting for development in this domain. Thus, the association of shifting/inhibition to vitalist understanding cannot be explained away by other constructs that could plausibly be implicated in episodes of conceptual change, namely knowledge enrichment mechanisms, working memory, and fluid IQ. This pattern of results makes it clear that neither *general* intelligence, nor *general* executive functioning, nor *general* capacity for acquiring fast-mappable factual knowledge can—alone or in any combination—fully explain variance in the development of this domain. Rather, specific executive resources must be brought to bear in this conceptual construction, and the identity of these specific resources constrains our account of the learning processes needed to master and deploy the framework theory of vitalist biology.

4.1. Factual knowledge: Necessary but not sufficient

Though the Factual Knowledge and RV tasks do not specifically probe knowledge of biological facts that would be relevant to the construction of the vitalist theory, the Knowledge Enrichment composite nevertheless strongly predicts Vitalism scores. As noted above, this is most likely because the knowledge enrichment tasks reflect variance in both the richness of the child's environmental input and in the child's knowledge enrichment capacity itself (see Bascandziev et al., 2018). These sources of variance would also be expected to contribute to variance in accumulated fast-mapped *biological* facts, facts that are specifically relevant to the construction of the vitalist theory. Such facts provide rich content and structural support to the process of theory building. After all, theories provide explanations for observations and facts, and explanation seeking cannot begin without an explanandum.

There are many biological facts that can be stated in terms of concepts available both to children who have mastered vitalist biology and those who have not. These facts are acquired through knowledge enrichment mechanisms by both groups of children. For example, four-year-olds can easily learn that one needs to eat to grow, for they have relevant concepts of eating (ingesting food) and growth (getting bigger). Of course, they do not yet know the vitalist explanation of this simple fact. Bascandziev et al. (2018) spell out three quite different roles that such facts might play in that construction. First, *anomalous* facts—facts that cannot be understood and integrated with one's current theory—provide important motivation for conceptual change, often triggering its underlying processes. Second, some facts will be learned through the combinatorial machinery of language alone. If children are told that plants are alive, or that Grampa is dead, *before* they have the vitalist concepts of life and death, then these facts are stored in placeholder structures that are built during the long process of conceptual change. Finally, facts stated in vocabulary common across both conceptual systems, Theory 1 and Theory 2, play a crucial role in the modeling processes through which the Theory 2 meanings of the placeholder terms are constructed. Thus, it is both important and interpretable that, with other predictor variables controlled, our composite measure of general fast-mapped factual knowledge was the strongest predictor of Composite Vitalism scores and demonstrated the strongest overall correlations with the individual interviews in the Vitalism battery as well.

Nevertheless, conceptual change demands more than fact learning alone. Awareness of an anomaly may be motivating but does not in and of itself tell the learner how to resolve that anomaly. As the history of science repeatedly shows, scientists can be aware of anomalous facts, and even be aware that some conceptual change is necessary, and still decades or even centuries of scientific development are required before the anomaly is resolved (see Carey, 2009, pp. 371–376 for a worked example). Similarly, progress in building and deploying the vitalist theory also demands more than the accumulation of factual knowledge articulated in terms of concepts available in the enriched agency theory. Clear evidence for this derives from the study of adults with Williams syndrome described above (Johnson & Carey, 1998). The present study suggests that the impairments in EF that are caused by

Williams syndrome, specifically in inhibition and shifting, preclude virtually any progress on the conceptual changes that are necessary for construction of a vitalist biology.

The patterns of intercorrelations among the tasks provide further evidence that the learning mechanisms underlying knowledge enrichment differ from those underlying conceptual change. First, our measures of Receptive Vocabulary and Factual Knowledge, which constitute our Knowledge Enrichment construct, were highly correlated, as predicted. This suggests that they rest on the same associative/statistical learning mechanisms, mechanisms that are presumably highly constrained by knowledge of the syntactic and semantic structure of natural language. Second, Knowledge Enrichment was not correlated with H&F/FL. Third, the relation between Composite Vitalism and H&F/FL, and the relation between Composite Vitalism and VF, each remained significant even after partialling out Knowledge Enrichment. Clearly, if the learning mechanisms underlying conceptual change were the same as those underlying fast-mapped knowledge enrichment, the measures predicting each would not dissociate in this way.

However, this dissociation was only partial in the case of Verbal Fluency, as VF and Knowledge Enrichment were correlated. VF is a measure of executive function, but insofar as the measure reflects the size and organization of fast-mapped labels for animals and foods, it is also a measure of knowledge enrichment. It might thus predict Composite Vitalism for the same reason that other measures of knowledge enrichment do. If this were the only reason, however, then its relationship with Composite Vitalism should not have survived partialling out Knowledge Enrichment. That is does suggests that it is the executive components of VF—those responsible for endogenous shifting and selection among competing representations—that account for the unique variance it contributes to predicting vitalist understanding.

4.2. Why shifting and inhibition predict vitalism

There are at least two possible roles for shifting and inhibition in predicting progress on measures of conceptual development: they could be involved in the learning processes that lead to the *construction* of theoretical knowledge, or they could be required for online *expression* of that knowledge. These two possibilities are not mutually exclusive; indeed, there is reason to believe that shifting and inhibition mechanisms are drawn on for *both* the construction of vitalist theory and its online expression.

The role of shifting and inhibition in adults' expression of vitalist biology has been clearly documented on the Animism interview. Several studies show that a concept *alive* that is rooted in a biologically enriched theory of agency interferes with the categorization of living and non-living entities according to the vitalist concept of life. Under speeded response conditions, adolescents, college-aged adults, and even biology professors are less accurate and slower at judging that plants are alive and the sun is not alive than that dogs are alive and a rock is not alive (Babai, Sekal, & Stavy, 2010; Goldberg & Thompson-Schill, 2009). Importantly, this interference is seen only on entities for which the vitalist theory and the enriched agency theory provide different answers (see also Shtulman & Valcarcel, 2012; Vosniadou et al., 2018). Furthermore, a large percentage of healthy elderly adults, given all the time they need in order to respond, make the same sort of animist errors on the Animism interview as do young children, explicitly justifying their animist responses with respect to activity and movement (Tardiff et al., 2017; Zaitchik & Solomon, 2008). Notably, declines in inhibition and shifting among healthy elderly adults are associated with increased animist responding (Tardiff et al., 2017).

While shifting and inhibition resources certainly play a role in the expression of vitalist understanding, particularly on the Animism interview, they also play a role in the construction of the vitalist theory. This was demonstrated in Bascandziev et al.'s (2018) study, which trained children on the functions of several body organs, training that never related body functions to the concepts *life* and *death*. That is, the training provided instruction that could only directly improve performance on the Body Parts interview. Despite this, gains from pretest to posttest were found across all three interviews, demonstrating far transfer to understanding of these never-mentioned vitalist concepts. Crucially, it was shifting/inhibition, not receptive vocabulary, that predicted training-induced gains, clearly implicating these EFs in the construction process.

Similarly, in the present study children were at the age where they are just beginning to construct a vitalist biology, so it is likely that their performance on the Vitalism battery reflects their early progress in building this new theory rather than their ability to merely express a theory they have already built. Importantly, our measures of shifting and inhibition, and particularly VF, were related not only to performance on Animism but to the other two interviews as well. Since neither of these other interviews demands the inhibition of prepotent agency responses for accurate expression, such inhibition cannot be the only role played by shifting and inhibition on these tasks in the present data.

Thus, we conclude that inhibition and shifting play a privileged role in the early stages of the construction of vitalist biology. We see two ways that these EFs may be important to episodes of theory construction and conceptual change such as this one. First, they are important to processes of comprehension monitoring and contradiction resolution that are central to conceptual construction. Second, they are important to the bootstrapping mechanisms that underlie the conceptual constructions needed for some contradiction resolution.

I. *Comprehension monitoring and explanation seeking* processes, which are important drivers of conceptual change (Lombrozo, 2012), are likely to draw heavily on inhibition and shifting. Before the vitalist theory has been fully acquired and entrenched, including during the very episodes of learning when a vitalist theory of biology is under construction, there is constant interference from prepotent agency representations. The same is likely true for the construction of intuitive physics, theory of mind, rational numbers, or any episode of theory building in which earlier prepotent concepts are present (Shtulman & Valcarcel, 2012; Vosniadou et al., 2018). Recognition of such conflicts is an important motivator of conceptual change. Any parent of normally-developing three- and four-year-old children is familiar with "why" or "how come" questions, in which children seek explanations. Often their explanation-seeking is motivated by a conflict in their beliefs or a failed prediction. These are very common in the domain of intuitive biology, for children care about death, where babies come from, what it means that they are going to grow up to be adults, and the like. For example, consider the following exchange between an almost-four-year-old and her mother, from Carey (1985):

> Child. That's funny, statues are not alive but you can still see them.
> Mother. What's funny about that?
> Child. Well, Grampa's dead and that's sad because we can never see him again.
> Mother. Oh, I see. Well some things, like tables are chairs and statues are never alive so they can't die, but other things, like animals and plants and people, first they are alive and then they die, and that's sad because when they die they don't exist anymore, and we can never see them again.
> Child (excitedly). Isn't that funny? Tables and chairs are not alive and we can still SEE them.

The child's question derived from a contradiction between being *not alive* (which applies to statues) and being *absent* (which applies to people who are not alive). This exchange also illustrates that noticing a contradiction does not tell the child (or the scientist, see Carey, 2009; Wiser & Carey, 1983) how to resolve it. The three-year-old's contradiction was articulated in terms of her undifferentiated concept *alive*. Importantly, she was in no position to understand her mother's attempt to point out to her that she had failed to differentiate the concept *dead* from the concept *inanimate.*

Comprehension monitoring and interference resolution are key goals of the EF system (Hasher & Zacks, 1988). Attempts to resolve a contradiction involve searching one's knowledge for relevant information and inhibiting one possibility as one entertains another. But, as just noted, noticing a contradiction does not entail that the learner is in a position to resolve it; often its resolution awaits conceptual constructions not yet made (as in the case of the child above). It is likely that inhibition and shifting also play important roles in the bootstrapping processes underlying such construction.

II. *One bootstrapping process,* dubbed "Quinian bootstrapping," is implicated in many cases of conceptual change, both in the history of science and in ontogenesis (Carey, 2009). The first step in Quinian bootstrapping involves the creation of a placeholder structure, a structure formulated through knowledge enrichment alone. The representations in the placeholder structure are not fully interpretable to a child who lacks the concepts of Theory 2. At this point, the child can only use the placeholder structure to model phenomena stated in terms of concepts that are common to Theory 1 and Theory 2. The modeling processes include thought experimentation, limiting case analyses, and most importantly, constructing analogical mappings. Consider the demands of constructing such an analogy: one must select the conceptually relevant structures among competing alternatives and inhibit the irrelevant ones; one must shift between the two structures being aligned; one must select the mappable features and inhibit the non-mappable features of both. These processes—*selection among competing representations*, *switching between data structures*, and *inhibiting goal-irrelevant information*—are exactly the ones tapped by H&F/FL and VF.

That said, there are limitations to the present work that license future study. First, because we used a combined exogenous shifting/inhibition measure, we cannot isolate whether it was shifting, inhibition, or both that drive the results. We did not attempt to separate them for two reasons. First, we were primarily interested in replicating prior results and differentiating shifting/inhibition from alternative constructs. Second, there is controversy over whether the EFs are fully differentiated in the preschool and early elementary school years. In particular, a number of studies find support for a single unitary EF construct at age 3 (e.g., Wiebe et al., 2011; Willoughby, Blair, Wirth, Greenberg, & The Family Life Project Investigators, 2010). By age 5–7, there is evidence of some differentiation, but also differences across studies in the nature of the factors. Crucially, in studies that measure tasks tapping all three constructs found in adults (inhibition, shifting, WM), multiple studies find no evidence for a separable inhibition factor (Lee et al., 2013; van der Ven et al., 2013; cf. Monette, Bigras, & Lafrenière, 2015). The inability to isolate inhibition in particular is in line with recent work in adults arguing that rather than standing on its own, inhibition is subsumed by a common factor underlying all three of these EFs, perhaps reflecting the shared need for goal maintenance among all EF tasks (Friedman & Miyake, 2017). For this reason, both Zaitchik et al. (2013) and the present study of conceptual change in younger children used a combined shifting/inhibition measure. For older children, however, it has been shown that these two EFs are separable and have different roles to play in on-line processing (Vosniadou et al., 2018). Future work utilizing carefully designed batteries of multiple inhibition and shifting tasks will be necessary to determine whether they can be separated in younger children and whether they have different roles to play in conceptual construction.

Some caution is also warranted in interpreting the role of VF. VF is a complex task, and its precise underlying executive components and relationships to other measures of EF is still under investigation. Notably, recent work suggests it shares variance with all three component EFs tested here—shifting, inhibition, and working memory—but that it also taps unique variance not attributable to these EFs, vocabulary, or nonverbal intelligence (Gustavson et al., 2019; see also Hedden & Yoon, 2006; Shao, Janse, Visser, & Meyer, 2014; Unsworth, Spillers, & Brewer, 2011). Care must be taken in applying the results of these studies to our population, however, as they were not conducted in young children. This complexity highlights the importance of the multiple regression approach taken in the present study, which demonstrated the unique variance shared between VF and Vitalism. Because VF uniquely measures endogenous flexibility in young children (Barker et al., 2014; Snyder & Munakata, 2010), that seems to be the most likely source of its contribution to construction of vitalist biology, with all other variables controlled. Still, additional studies are needed to precisely characterize the component processes of VF at this age and their contributions to conceptual construction.

4.3. Why working memory and fluid IQ do not predict vitalism

In contrast to measures of shifting and inhibition, working memory was not associated with progress in constructing a vitalist biology. Counting Recall provides a robust measure of working memory capacity, and there was sufficient variance in performance to reveal relations between working memory and the vitalism measures. Of course, caution must be taken in interpreting null results, but one explanation for a lack of relation between working memory and vitalist biology may be that the three interviews do not strongly draw on working memory for on-line expression of the vitalist theory. Indeed, the on-line demands are likely minimal. Responses are given immediately after each question; moreover, the information drawn upon is unlikely to consist of elaborate or numerous data structures that might draw heavily on working memory. The majority of questions ask for properties or exemplars of various items. In familiar cases this merely involves accessing these in long-term memory (e.g., Is a cat alive? What is your heart for?). In unfamiliar cases, inferences must be drawn based on the properties of the item in question (e.g., Is the sun alive? What would happen if you didn't have a stomach?)—but these inferences need not require holding information in mind at anywhere near the capacity limits of working memory.

This result diverges from findings among the elderly, where working memory does indeed predict performance on the Animism interview (Tardiff et al., 2017). For this population, presumably, there are strong demands on the ability to hold in mind the biological context of the task. But elderly adults have an intact vitalist theory, and the children under study do not. In young children, who do not have a rich and elaborated theory to hold in mind, this specific demand on expression is likely to be weak or absent. With respect to the construction of the vitalist theory in young children, the present results thus suggest that working memory does not play an important role in the early steps. Of course, the same might not be true for other conceptual constructions, or even for later stages in the construction of vitalist biology.

We similarly found no relationship between progress in the development of the vitalist theory and fluid IQ. This is in line with evidence that the particular executive function most associated with fluid IQ is working memory; indeed, these two constructs are considered by some to be highly overlapping or even identical (Duncan et al., 2012; Friedman et al., 2006; Harrison et al., 2015; Wiley et al., 2011). Inhibition and shifting, in contrast, appear to be separable from working memory and fluid IQ (Friedman et al., 2006). Our findings support these claims. Children's scores on Counting Recall and Matrices were significantly correlated with each other and highly similar in their relations to other variables as well.

However, this result is still surprising. Our Matrices measure of fluid IQ is a measure of analogical reasoning, and analogical reasoning, as mentioned above, is central to any episode of conceptual change (Carey, 2009; Gentner, 2002; Nersessian, 1992). With respect to vitalist biology in particular, Hatano and Inagaki (1994) argue that the analogy between humans' consumption of food and water and plants' consumption of water is crucial to the coalescence of animals and plants into the ontological category *living things*. Moreover, analogies (and other modeling processes involved in episodes of conceptual change) at least sometimes rely heavily on working memory. For example, numerous behavioral, developmental, and neuroimaging studies support a role for working memory and underlying frontal executive processes in analogical reasoning (Cho, Holyoak, & Cannon, 2007, 2010; Crone et al., 2009; Green, Fugelsang, Kraemer, Shamosh, & Dunbar, 2006; Halford, 1993; Richland et al., 2006; Waltz, Lau, Grewal, & Holyoak, 2000; Watson & Chatterjee, 2012).

In view of the important role of analogical reasoning and other modeling processes in conceptual construction, why is it that Counting Recall and Matrices fail to predict children's Composite Vitalism scores? We see two possible reasons for this. First, as noted above, the earliest stages of the construction of a theory of vitalism may not place high demands on working memory. Second, Matrices probes variance in highly constrained analogical reasoning problems, and this variance is swamped by a different and much stronger source of variance in analogical reasoning—the ability to set up the relevant structure-mapping problem to begin with. Adult frontal lobe patients, for example, are equal to controls in simple relational reasoning problems but severely impaired with more complex relations that presumably require more working memory to solve (Waltz et al., 1999). It seems likely that both Matrices and the construction of vitalist biology, at least at this earliest stage, require simple relational reasoning. The really difficult problem, at least in the early stages of constructing a vitalist biology, appears to lie in selecting relevant information and inhibiting irrelevant information.

4.4. Conclusion

Our results illuminate the specificity of the relationship between domain-general mechanisms and the earliest stages in the construction of vitalist biology. Had it been the case that all three of our EF measures *and* fluid IQ predicted progress in this construction, we would be left with the unsatisfying conclusion that more processing capacity leads to better performance. This was not the case. Nor was it the case that the association between vitalist biology and shifting and inhibition was better explained by variation in either of two plausible third common factors: working memory and fluid IQ. This finding of specificity invites the following questions: Do these relationships between domain-general mechanisms and vitalism hold at later ages, as both continue to develop, or do new patterns of relations emerge? How general is the finding that shifting and inhibition are involved in conceptual construction—is this true for all conceptual constructions? Finally, with respect to the learning processes that underlie conceptual construction, what is it exactly that draws on shifting and inhibition? Attempts to explore these questions are just beginning.

Acknowledgments

This work was funded with support from the National Science Foundation#1247396 (Co-PIS: Susan Carey and Deborah Zaitchik) and with support from the Spencer Foundation#201200109 (Co-PIS: Susan Carey and Deborah Zaitchik). We are grateful to the children and their families who participated in this study.

References

Ackerman, P. L. (2000). Domain-specific knowledge as the" dark matter" of adult intelligence: Gf/Gc, personality and interest correlates. *The Journals of Gerontology: Series B, 55*, 69–84.
Alloway, T. P. (2007). *Automated Working Memory Assessment.* London, UK: Pearson Assessment.
Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology, 87*, 85–106.
Babai, R., Sekal, R., & Stavy, R. (2010). Persistence of the intuitive conception of living things in adolescence. *Journal of Science Education and Technology, 19*, 20–26.
Barker, J. E., Semenov, A. D., Michaelson, L., Provan, L. S., Snyder, H. R., & Munakata, Y. (2014). Less-structured time in children's daily lives predicts self-directed executive functioning. *Frontiers in Psychology, 5*, 593.
Bascandziev, I., Powell, L. J., Harris, P. L., & Carey, S. (2016). A role for executive functions in explanatory understanding of the physical world. *Cognitive Development, 39*, 71–85.
Bascandziev, I., Tardiff, N., Zaitchik, D., & Carey, S. (2018). The role of domain-general cognitive resources in children's construction of a vitalist theory of biology. *Cognitive Psychology, 104*, 1–28.
Benson, J. E., Sabbagh, M. A., Carlson, S. M., & Zelazo, P. D. (2013). Individual differences in executive functioning predict preschoolers' improvement from theory-of-mind training. *Developmental Psychology, 49*, 1615–1627.
Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences, 21*, 327–336.
Blair, C., & Diamond, A. (2008). Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure. *Development and Psychopathology, 20*, 899–911.
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.
Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology, 33*, 205–228.
Carey, S. (1978). The child as word learner. In J. Bresnan, G. Miller, & M. Halle (Eds.). *Linguistic theory and psychological reality* (pp. 264–293). Cambridge, MA: MIT Press.
Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA: MIT Press.
Carey, S. (2009). *The origin of concepts.* New York: Oxford University Press.
Carey, S. (2015). Why theories of concepts should not ignore the problem of acquisition. In E. Margolis, & S. Laurence (Eds.). *Concepts: New directions* (pp. 415–454). Cambridge, MA: MIT Press.
Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Proceedings of the Stanford Child Language Conference,* 17–29 15.
Carey, S., Zaitchik, D., & Bascandziev, I. (2015). Theories of development: In dialog with Jean Piaget. *Developmental Review, 38*, 36–54.
Chieffo, D., Lettori, D., Contaldo, I., Perrino, F., Graziano, A., Palermo, C., ... Guzzetta, F. (2011). Surgery of children with frontal lobe lesional epilepsy: Neuropsychological study. *Brain and Development, 33*, 310–315.
Cho, S., Holyoak, K. J., & Cannon, T. D. (2007). Analogical reasoning in working memory: Resources shared among relational integration, interference resolution, and maintenance. *Memory & Cognition, 35*, 1445–1455.
Cho, S., Moody, T. D., Fernandino, L., Mumford, J. A., Poldrack, R. A., Cannon, T. D., ... Holyoak, K. J. (2010). Common and dissociable prefrontal loci associated with component mechanisms of analogical reasoning. *Cerebral Cortex, 20*, 524–533.
Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics, 50*, 66–71.
Contento, I. (1981). Children's thinking about food and eating: A Piagetian-based study. *Journal of Nutrition Education, 13*, 86–90.
Crone, E. A., Wendelken, C., Van Leijenhorst, L., Honomichl, R. D., Christoff, K., & Bunge, S. A. (2009). Neurocognitive development of relational reasoning. *Developmental Science, 12*, 55–66.
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. *Neuropshychologia, 44*, 2037–2078.
Demetriou, A., Spanoudis, G., Shayer, M., van der Ven, S., Brydges, C. R., Kroesbergen, E., ... Swanson, H. L. (2014). Relations between speed, working memory, and intelligence from preschool to adulthood: Structural equation modeling of 14 studies. *Intelligence, 46*, 107–121.
Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science, 318*, 1387–1388.
Duncan, J., Schramm, M., Thompson, R., & Dumontheil, I. (2012). Task rules, working memory, and fluid intelligence. *Psychonomic Bulletin & Review, 19*, 864–870.
Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex, 86*, 186–204.
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.
Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology, 54*, 1–34.
Gathercole, S. E., Tiffany, C., Briscoe, J., Thorn, A., & 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.
Geary, D. C., vanMarle, K., Chu, F. W., Hoard, M. K., & Nugent, L. (2019). Predicting age of becoming a cardinal principle knower. *Journal of Educational Psychology, 111*, 256–267.
Gentner, D. (2002). Analogy in scientific discovery: The case of Jonannes Kepler. In L. Magnani, & N. J. Nersessian (Eds.). *Model-based reasoning: Science, technology, values* (pp. 21–39). Boston, MA: Springer.
Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences, 7*, 287–292.
Goldberg, R. F., & Thompson-Schill, S. L. (2009). Developmental 'roots' in mature biological knowledge. *Psychological Science, 20*, 480–487.
Gray, S., Green, S., Alt, M., Hogan, T., Kuo, T., Brinkley, S., ... Cowan, N. (2017). The structure of working memory in young children and its relation to intelligence. *Journal of Memory and Language, 92*, 183–201.
Green, A. E., Fugelsang, J. A., Kraemer, D. J., Shamosh, N. A., & Dunbar, K. N. (2006). Frontopolar cortex mediates abstract integration in analogy. *Brain Research, 1096*, 125–137.

Gustavson, D. E., Panizzon, M. S., Franz, C. E., Reynolds, C. A., Corley, R. P., Hewitt, J. K., ... Friedman, N. P. (2019). Integrating verbal fluency with executive functions: Evidence from twin studies in adolescence and middle age. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0000589 Advance online publication.

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

Harrison, T. L., Shipstead, Z., & Engle, R. W. (2015). Why is working memory capacity related to matrix reasoning tasks? *Memory & Cognition, 43*, 389–396.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Vol. Ed.), *The psychology of learning and motivation: Vol. 22*, (pp. 193–225). New York: Academic Press.

Hatano, G., & Inagaki, K. (1994). Young children's naïve theory of biology. *Cognition, 50*, 171–188.

Hedden, T., & Yoon, C. (2006). Individual differences in executive processing predict susceptibility to interference in verbal working memory. *Neuropsychology, 20*, 511–528.

Henry, J. D., & Crawford, J. R. (2004). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology, 18*, 284–295.

Hernandez, M. T., Sauerwein, H. C., Jambaqué, I., De Guise, E., Lussier, F., Lortie, A., ... Lassonde, M. (2002). Deficits in executive functions and motor coordination in children with frontal lobe epilepsy. *Neuropsychologia, 40*, 384–400.

Inagaki, K., & Hatano, G. (1993). Young children's understanding of the mind-body distinction. *Child Development, 64*, 1534–1549.

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 folkbiology: Evidence from people with Williams syndrome. *Cognitive Psychology, 37*, 156–200.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test* (2nd ed.). Circle Pines, MN: American Guidance Services.

Kavé, G., Kigel, S., & Kochva, R. (2008). Switching and clustering in verbal fluency tasks throughout childhood. *Journal of Clinical and Experimental Neuropsychology, 30*, 349–359.

Korkman, M., Kemp, S. L., & Kirk, U. (2001). Effects of age on neurocognitive measures of children ages 5 to 12: A cross-sectional study on 800 children from the United States. *Developmental Neuropsychology, 20*, 331–354.

Laski, E. V., & Dulaney, A. (2015). When prior knowledge interferes, inhibitory control matters for learning: The case of numerical magnitude representations. *Journal of Educational Psychology, 107*, 1035–1050.

Lee, K., Bull, R., & Ho, R. M. (2013). Developmental changes in executive functioning. *Child Development, 84*, 1933–1953.

Lombrozo, T. (2012). Explanation and abductive inference. In K. J. Holyoak, & R. G. Morrison (Eds.). *Oxford handbook of thinking and reasoning* (pp. 260–276). Oxford, UK: Oxford University Press.

Luo, Y., & Baillargeon, R. (2005). Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. *Psychological Science, 16*, 601–608.

Markson, L., & Bloom, P. (1997). Evidence against a dedicated system for word learning in children. *Nature, 385*, 813–815.

McCloskey, M. (1983). Naive theories of motion. In D. Gentner, & A. Stevens (Eds.). *Mental models* (pp. 229–324). Hillsdale, NJ: Erlbaum.

McCloskey, M., Caramazza, A., & Green, B. (1980). Curvilinear motion in the absence of external forces: Naïve beliefs about the motion of objects. *Science, 210*, 1139–1141.

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.

Monette, S., Bigras, M., & Lafrenière, M. A. (2015). Structure of executive functions in typically developing kindergarteners. *Journal of Experimental Child Psychology, 140*, 120–139.

Nersessian, N. J. (1992). How do scientists think? Capturing the dynamics of conceptual change in science. In R. N. Giere (Ed.). *Cognitive models of science* (pp. 3–45). Minneapolis, MN: University of Minnesota Press.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*, 255–258.

Osório, A., Cruz, R., Sampaio, A., Garayzábal, E., Martínez-Regueiro, R., Gonçalves, Ó. F., ... Fernández-Prieto, M. (2012). How executive functions are related to intelligence in Williams syndrome. *Research in Developmental Disabilities, 33*, 1169–1175.

Piaget, J. (1929). *The child's conception of the world.* London, UK: Routledge & Kegan Paul.

Rhodes, S. M., Riby, D. M., Fraser, E., & Campbell, L. E. (2011). The extent of working memory deficits associated with Williams syndrome: Exploration of verbal and spatial domains and executively controlled processes. *Brain and Cognition, 77*, 208–214.

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology, 94*, 249–273.

Riva, D., Avanzini, G., Franceschetti, S., Nichelli, F., Saletti, V., Vago, C., ... Bulgheroni, S. (2005). Unilateral frontal lobe epilepsy affects executive functions in children. *Neurological Sciences, 26*, 263–270.

Riva, D., Nichelli, F., & Devoti, M. (2000). Developmental aspects of verbal fluency and confrontation naming in children. *Brain and Language, 71*, 267–284.

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., ... Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia, 42*, 1029–1040.

Sauzéon, H., Lestage, P., Raboutet, C., N'Kaoua, B., & Claverie, B. (2004). Verbal fluency output in children aged 7–16 as a function of the production criterion: Qualitative analysis of clustering, switching processes, and semantic network exploitation. *Brain and Language, 89*, 192–202.

Saxe, R., Tenenbaum, J., & Carey, S. (2005). Secret agents: 10 and 12-month-olds infer an unseen cause of the motion of an inanimate object. *Psychological Science, 16*, 995–1001.

Schipolowski, S., Wilhelm, O., & Schroeders, U. (2014). On the nature of crystallized intelligence: The relationship between verbal ability and factual knowledge. *Intelligence, 46*, 156–168.

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology, 5*, 772.

Shtulman, A. (2006). Qualitative differences between naive and scientific theories of evolution. *Cognitive Psychology, 52*, 170–194.

Shtulman, A., & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition, 124*, 209–215.

Slaughter, V., Jaakkola, K., & Carey, S. (1999). Constructing a coherent theory: Children's biological understanding of life and death. In M. Siegel, & C. 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.

Snyder, H. R., & Munakata, Y. (2013). So many options, so little control: Abstract representations can reduce selection demands to increase children's self-directed flexibility. *Journal of Experimental Child Psychology, 116*, 659–673.

Tardiff, N., Bascandziev, I., Sandor, K., Carey, S., & Zaitchik, D. (2017). Some consequences of normal aging for generating conceptual explanations: A case study of vitalist biology. *Cognitive Psychology, 95*, 145–163.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*, 138–146.

Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P., & Stuss, D. (1998). Clustering and switching on verbal fluency: The effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia, 36*, 499–504.

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2011). Variation in verbal fluency: A latent variable analysis of clustering, switching, and overall performance. *The Quarterly Journal of Experimental Psychology, 64*, 447–466.

van der Ven, S. H., Kroesbergen, E. H., Boom, J., & Leseman, P. P. (2013). The structure of executive functions in children: A closer examination of inhibition, shifting, and updating. *The British Journal of Developmental Psychology, 31*, 70–87.

Vosniadou, S. (2019). The development of students' understanding of science. *Frontiers in Education, 4*, 32.

Vosniadou, S., Pnevmatikos, D., Makris, N., Eikospentaki, K., Lepenioti, D., Chountala, A., & Kyrianakis, G. (2015). Executive functions and conceptual change in science and mathematics learning. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.). *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 2529–2534).

Vosniadou, S., Pnevmatikos, D., Makris, N., Lepenioti, D., Eikospentaki, K., Chountala, A., ... Kyrianakis, G. (2018). The recruitment of shifting and inhibition in on-line science and mathematics tasks. *Cognitive Science, 42*, 1860–1886.

Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Mishkin, F. S., de Menezes Santos, M., ... Miller, B. L. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological Science, 10*, 119–125.

Waltz, J. A., Lau, A., Grewal, S. K., & Holyoak, K. J. (2000). The role of working memory in analogical mapping. *Memory & Cognition, 28*, 1205–1212.

Watson, C. E., & Chatterjee, A. (2012). A bilateral frontoparietal network underlies visuospatial analogical reasoning. *NeuroImage, 59*, 2831–2838.

Wellman, H. M., & Gelman, S. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology, 43*, 337–375.

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology, 108*, 436–452.

Wiley, J., Jarosz, A. F., Cushen, P. J., & Colflesh, G. H. (2011). New rule use drives the relation between working memory capacity and Raven's Advanced Progressive Matrices. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 256–263.

Willoughby, M. T., Blair, C. B., Wirth, R. J., Greenberg, M., & The Family Life Project Investigators (2010). The measurement of executive function at age 3 years: Psychometric properties and criterion validity of a new battery of tasks. *Psychological Assessment, 22*, 306–317.

Wiser, M., & Carey, S. (1983). When heat and temperature were one. In D. Gentner, & A. Stevens (Eds.). *Mental models* (pp. 267–297). Hillsdale, NJ: Erlbaum.

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

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*, 1–34.

Woodward, A. L., & Somerville, J. A. (2000). Twelve-month-old infants interpret action in context. *Psychological Science, 11*, 73–77.

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

Zaitchik, D., Iqbal, Y., & Carey, S. (2013). The effect of executive function on biological reasoning in young children: An individual differences study. *Child Development, 85*, 160–175.