Journal of Experimental Child Psychology 200 (2020) 104932

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

The world within: Children are sensitive to internal complexity cues

Richard E. Ahl [a,b,*], Dorsa Amir [b], Frank C. Keil [a]

[a] Department of Psychology, Yale University, New Haven, CT 06511, USA
[b] Department of Psychology, Boston College, Chestnut Hill, MA 02467, USA

ARTICLE INFO

Article history:
Received 24 April 2019
Revised 23 May 2020
Accepted 16 June 2020

Keywords:
Insides
Artifacts
Causal reasoning
Mechanisms
Scientific education

ABSTRACT

When reasoning about the mechanisms of complex entities, it is important to consider their internal parts. Previous research has shown that young children view "insides" as critical to how objects function. However, whether children hold specific expectations regarding complex objects' insides remains an open question. Here, children (n = 378) and adults (n = 124) made internal and causal complexity judgments regarding real-world objects. In Study 1, 5- and 6-year-olds, but not 4-year-olds, succeeded at internal complexity judgments and matched complex artifacts with complex insides. All age groups succeeded at causal complexity judgments and identified complex artifacts as causally complex. Study 2 tested whether the internal complexity cues of number/area, diversity, and connections of internal parts conveyed complexity to children. The 5-year-olds were sensitive only to number/area of internal parts as a complexity cue, but the older children and adults were sensitive to all three cues plus number of parts when controlling for area (Study 3). Despite limited exposure to insides, even young school-age children hold detailed and abstract expectations concerning internal complexity.

* Corresponding author at: Department of Psychology, Boston College, Chestnut Hill, MA 02467, USA.
E-mail address: richard.ahl@bc.edu (R.E. Ahl).

https://doi.org/10.1016/j.jecp.2020.104932

Introduction

The industrialized world is a complicated place. Unlike the simple tools of prehistoric humans, such as hand axes, the workings of many contemporary devices, such as smartphones, are beyond the comprehension of even well-educated adults (Rozenblit & Keil, 2002). The functions of such devices and the mechanisms by which they are carried out make them different from simpler artifacts, profoundly affecting how we use, fix, and learn from them (Kominsky, Zamm, & Keil, 2018). One source of complex artifacts' complexity, and a point of contrast with simpler artifacts, is that complex artifacts' functions emerge largely through their hidden internal properties. The present studies explore the extent to which children hold expectations regarding insides and how such expectations relate to functional complexity. Mentally representing the unseen is crucial to reasoning about complex entities and processes, an important skill when navigating an increasingly technological society (Fourez, 1997). Considering insides helps us to reason about why a toy's batteries need to be changed and why it is harder to fix a computer than to fix a flashlight.

Even infants expect certain entities to have insides (Setoh, Wu, Baillargeon, & Gelman, 2013), and children privilege information about insides relative to outsides (Gelman & Wellman, 1991). However, beyond recognizing the presence and importance of insides, do children hold detailed expectations about them? Although we can infer from previous research that children believe objects like smartphones have insides that make them work, do children expect their insides to be different from the insides of simpler objects? The understanding that insides matter does not mean children have specific expectations regarding internal parts. Such expectations are useful in many situations, such as attempting to fix an object and reasoning about an object's functional properties or value. Given that the workings of complex things are sophisticated and opaque, it might seem that young children must know little about them. Young children lack the content knowledge required to fully understand the causal mechanisms of animals and machines (Au & Romo, 1999; Carey, 1985; Keil, 2003). Because examining insides can be harmful, insides are often challenging to observe and rarely amenable to direct causal intervention (Grotzer & Tutwiler, 2014). Beyond practical considerations, many researchers from Piaget onward have argued that children unduly focus on what they can see and struggle to represent that which is not visible (see Springer, 2001). Thus, there are reasons to both support and challenge the proposal that children have expectations regarding what insides are like and how insides relate to functional complexity.

Our goal is to replace the "black box" of children's expectations regarding complex objects' insides with a set of schematics. Our empirical studies explore which internal cues convey causal complexity in the domain of artifacts and addressed broader questions regarding the richness of children's expectations for insides. We begin by summarizing what is known about infants' and children's understanding of insides and artifact complexity.

The functional role of insides

The understanding that insides are important to the functions and identities of certain entities begins during infancy (Taborda-Osorio & Cheries, 2017). When 8-month-olds were familiarized with animal-like, self-propelled, and contingently responding objects that were later revealed to be hollow (i.e., lacking insides), they looked longer than when the objects were revealed to be filled (i.e., suggesting insides). This indicates that infants expect animates to have insides (Setoh et al., 2013). Older infants view insides as having special causal significance (Newman, Herrmann, Wynn, & Keil, 2008). Fourteen-month-olds expected animated cats with the internal feature of stomachs of the same color, but not the external feature of hats of the same color, to display similar patterns of motion. These findings accord with studies in which 14-month-olds could categorize novel animate objects and named objects based on shared internal features even when external features were more perceptually salient (Welder & Graham, 2001). Therefore, infancy research indicates that the privileging of insides emerges long before children receive formal education on this topic and suggests its foundational role in human cognition, shaping causal attributions, categorization, and an early interest in underlying properties.

The only clear evidence that infants have expectations about insides comes from studies using animal-like entities. However, preschoolers also acknowledge the importance of insides for artifacts, for instance, endorsing insides as causing a telephone to ring (Gelman & Kremer, 1991). Many complex artifacts have similarities to natural kinds in that they possess internal mechanisms that are made of specific materials, dictate observable functions, and are configured in certain ways (see Gelman, 1988) and may have stronger parallels to animals than to simple artifacts (e.g., tables) regarding judgments of insides (Arbesman, 2016).

During the preschool years, children show an explicit awareness that insides are important to an object's identity, category membership, functional capacity, and essence (Gelman & Wellman, 1991; Keil, 1989). Such awareness helps children to overlook surface features in favor of deeper causal properties. When asked whether the identity and functionality of animals, inanimate natural kinds (e.g., eggs), and artifacts would change or remain the same if their insides versus outsides were removed, preschoolers thought that removing insides would result in greater changes for all three object categories (Gelman & Wellman, 1991). Preschoolers also understand that insides matter for mechanical artifacts. Four-year-olds, but not 3-year-olds, inferred that an object's hidden internal part was responsible for activating another machine and favored internal parts over external features when selecting new objects that would activate the machine (Sobel, Yoachim, Gopnik, Meltzoff, & Blumenthal, 2007; see also Walker, Lombrozo, Legare, & Gopnik, 2014).

Appearance and substance of insides

Perhaps the most basic expectation regarding the appearance of insides is that it is different from what is outside. Even preschoolers who struggle to name object parts often list different parts inside than outside (Gelman, 1990). Preschoolers also know that two objects' external similarities might not indicate internal similarities. For example, preschoolers may choose a pig and a piggy bank as looking more similar than a pig and a cow, whereas they may choose a pig and a cow as having similar insides (Gelman & Wellman, 1991). Such an understanding helps children to ignore external appearances when considering internal parts.

Preschoolers are sensitive to an entity's ontological kind and expect objects to have insides in keeping with their kind. Simons and Keil (1995) found that preschoolers expected animals and artifacts to have different kinds of things inside even if they were unclear on the details, as demonstrated by choosing depictions of natural kinds (either organs or rocks) as belonging inside animals and depictions of human-made things (either gears or blocks) as belonging inside machines. Eight-year-olds performed well across the board and rarely made within-domain errors, consistent with knowledge beyond the natural versus non-natural distinction. More recent research found better performance with preschoolers in a similar task, likely due to greater clarity of the images used and the absence of domain-related distractors (Gottfried & Gelman, 2005). Importantly, unfamiliar entities (e.g., tapirs) were used in the latter study; thus, children were unlikely to have based matches on encounters with particular objects but instead likely based matches on general expectations about categories.

Insides and complexity

Converging evidence across studies with different ages and methods indicates that preschoolers hold abstract expectations regarding insides; they expect animals and complex artifacts to have insides, know insides matter, and recognize that insides are different from outsides. Given that children view insides as crucial and yet rarely see them, it seems plausible for children to seek information about insides and represent them when they are not visible. Our current knowledge of children's expectations regarding insides, however, is limited, and this issue has particular relevance to children's understanding of complex artifacts. Causal complexity is often hidden inside, making the study of intuitions about inside parts and their configurations crucial to understanding complexity itself and how it is represented by human minds (Arbesman, 2016).

Two recent studies show that children associate complex insides with complex functions. Consider a case where children encountered two novel light-producing machines with identical external appearances (Erb, Buchanan, & Sobel, 2013). One machine flashed a single solid color, whereas the other machine cycled through several different colors. Four-year-olds preferentially matched the variable machine with a visual depiction of complex insides conveyed with multiple, diverse connected parts. In a set of studies involving number and diversity of functions as complexity cues, children learned about several pairs of novel machines with identical external appearances (Ahl & Keil, 2017). In Study 1, one machine per pair performed two diverse functions and the other machine performed a single function. In Studies 2–4, one machine performed two diverse functions and the other machine performed two nondiverse functions. Even 4- and 5-year-olds matched two-function machines with complex insides in Study 1, but only older children and adults matched diverse-function machines with complex insides. The complex insides, relative to simple insides, had more numerous, diverse, and connected parts.

These studies on functional and internal complexity indicate that even young children have a rudimentary ability to determine how artifact functions imply internal complexity in novel machines. They appear to indicate that children view number, diversity, and connections of parts as internal complexity cues. However, on closer examination, they do not address which *specific* internal features convey complexity. To demonstrate internal complexity, both studies simultaneously varied the number, diversity, and connections of inside parts. As a result, which cues are necessary, or sufficient, to convey complexity to children remains unknown. For instance, the simple expectation that complex objects have "more" inside, with no further details represented, could have led children to match complex artifacts with complex insides. Moreover, do children hold expectations regarding the complexity of *real-world* objects? These studies used novel objects and presented standardized information about their functions. Such an approach has the benefits of tight control over stimulus features but differs from people's experiences with everyday objects, which are considerably more varied and rarely accompanied with clear summaries of functionality. Moreover, learning about everyday objects may take place over months or years rather than in one brief session. Thus, children's reasoning about novel machines might not map onto their reasoning about real-world objects.

Causal complexity is a highly abstract property, distinct from mere visual complexity, that is present in various entities yet has common features despite its diverse forms (Simon, 1962). Thus, the "architecture of complexity" (Simon, 1962) across an enormous range of entities may be mentally represented in shared ways. Developing an accurate sense of causal complexity is an important developmental achievement because complexity intuitions can guide future learning. Consider the very different kinds of questions one might ask about a couch versus a computer. Complicated objects warrant further exploration, and learning to use or fix complicated objects may require an expert's advice (Kominsky et al., 2018). Beliefs about complexity help children to efficiently allocate attention and seek guidance from others. They are also closely linked to judgments about insides.

The current studies

Across three studies, our focus was on children's and adults' expectations regarding the internal complexity of objects—henceforth referred to as *internal complexity judgments*—as measured in the Insides Block. Our key analyses examined how accuracy on the Insides Block varied across age groups and whether a given age group performed above chance. Here, participants saw pairs of real-world objects of different levels of complexity and two pictures representing objects' insides. Participants were instructed to match each object with the *insides* picture that is most like its insides. In the Learn Block, participants made separate *causal complexity judgments* regarding how hard it is to learn about how those objects work. (We use the terms *internal complexity judgments* and *causal complexity judgments* for brevity and because such judgments are differentially emphasized in each block; we do not mean to imply that participants necessarily judge internal complexity without considering causal complexity or vice versa.) This is one of few studies to examine children's sensitivity to real-world artifacts' internal or causal complexity, and it is the only study we know of to collect both kinds of measures.

The Learn Block was similar in Studies 1, 2, and 3, but the Insides Block differed across studies. In Study 1, we presented the cues of number/area, diversity, and connections of parts simultaneously by having participants identify whether *insides* pictures containing more versus less of these complexity cues are most like various objects' insides. In Study 2, we presented the cues of number/area, diversity, and connections of parts separately to test whether each individual cue conveys complexity. In Study 3, we delved deeper into the number cue by testing whether number of parts, when controlling for area, conveys complexity.

We chose to study number/area, diversity, and connections of parts because aforementioned studies suggest that young children may view them as internal complexity cues. Such cues are also emphasized by cognitive scientists' definitions of complexity, which mention the number, diversity, and interactions of elements within a causal system (Gelman, 1988; Kominsky et al., 2018; Simon, 1962). However, it is unclear whether children consider such cues when reasoning about real-world objects whose functions are not summarized by an experimenter. To effectively use such cues, children must have extracted relevant information about the objects' functions and/or mechanisms prior to testing. The Learn Block primarily served as a basis for evaluating Insides Block performance. Without independent judgments of causal complexity, low Insides scores could be attributed to a failure to view the objects *themselves* as complex; expectations regarding internal complexity specifically would remain unclear. Nonetheless, because few developmental studies have involved causal complexity judgments about artifacts, children's Learn Block proficiency was also a topic of interest.

Study 1 was necessary before separately testing the cues of number/area, diversity, and connections of inside parts in Studies 2 and 3. If an age group performed poorly on the Insides Block or Learn Block of Study 1, inclusion in subsequent studies would be unnecessary. Because the Learn Block was similar in all studies and the Insides Block was likely easiest in Study 1, where all complexity cues were combined, failure in the subtler tasks of Studies 2 and 3 would be inferable from failure in Study 1. We also used Study 1 to select our set of object pairs for subsequent studies; high-scoring pairs were most useful for Study 2′s goal of determining *which specific* internal complexity cues children use.

We set our Study 1 age range of 4-, 5-, and 6-year-olds based on patterns of successes and failures documented by the previous research literature. Erb et al. (2013) found that 4-year-olds, but not 3-year-olds, matched complex insides with variable machines, and Ahl & Keil (2017) found success with 4- and 5-year-olds in another insides-matching task. Kominsky et al. (2018) found that children aged 5–7 years made fairly consistent ratings of real-world objects' causal complexity but did not ask children to make forced-choice complexity judgments between objects of similar sizes. In addition, Trouche, Chuey, Lockhart, and Keil (2017) found that kindergartners, but not older children, often failed to judge a motorcycle as more complex than a bicycle. Thus, we set 4-year-olds as the youngest age group for our task. Above-chance scores on both blocks would warrant an age group's inclusion in Study 2.

Study 1

Method

Participants

Our sample of children included 20 4-year-olds (13 boys; M_{age} = 52.25 months, SD = 3.46, range = 46–58), 20 5-year-olds (11 boys; M_{age} = 66.05 months, SD = 2.58, range = 63–71), and 20 6-year-olds (13 boys; M_{age} = 77.55 months, SD = 2.82, range = 72–82) tested in children's or natural history museums (n = 22), private schools (n = 32), or our lab (n = 6). Participants were White (n = 50), Asian (n = 7), Black (n = 1), or biracial (n = 1); race was unreported for 1 participant. An additional 8 participants were excluded due to tester error (n = 1), limited English (n = 1), equipment failure (n = 1), comprehension check failures (n = 1), severe inattention (n = 2), or perseverative responding (n = 2). We believe that our sample came from predominantly middle- or upper-income families for all studies. In keeping with institutional review board protocols, parental consent and child assent were obtained before testing. Sample sizes were set a priori based on previous research and sample size calculators for comparisons with at-chance scores.

Adults were recruited and tested online using Amazon's Mechanical Turk, were compensated modestly, and were at least 18 years of age and living in the United States. Our adult sample for this study included 20 participants (9 women; M_{age} = 33.40 years, SD = 10.66). Of these adult participants, 11 reported holding a bachelor's degree or higher. An additional 5 participants were excluded for comprehension check failures (n = 4) or very fast completion times (n = 1).

Materials and procedure for children

Participants completed either the Insides Block or Learn Block first, with block order counterbalanced across participants. For simplicity's sake, we discuss the version with the Insides Block presented first. Scripts are included in Supplement 1 of the online supplementary material. The study was administered with an Apple iPad using the Offline Qualtrics app. In addition, we live-coded responses and made video recordings when possible. We used laminated copies of two pictures to represent machine insides; we use the labels *simple insides* and *complex insides* here but did not use such labels with participants. *Simple insides* depicted two machine-like parts on a rectangular grid. *Complex insides* used the same features but also added three more visually distinct parts plus wire connecting the parts. Both images were created by photographing parts from Snap Circuits, which are commercially sold toys. Descriptions are provided in Appendix A. Images are available from the first author upon request.

Participants made judgments regarding eight items, listed in Appendix B (images available upon request). Each item consisted of a pair of real-world objects that were relatively similar in size, with labels of similar word length and syllables, but differed in their causal and internal complexity (e.g., a smartphone vs. a flashlight). To minimize visual confounds, we depicted the objects using simplified line drawings similar in size and visual complexity, although one remaining limitation is that this approach does not account for visual complexity differences between the actual real-world objects themselves; the more causally complex objects were generally more visually complex. The objects we chose were inspired by Kominsky et al. (2018); many new objects were included to create suitable pairs based on the previous criteria. We piloted the objects to test whether young children were familiar with the objects and their functions. All objects used were deemed familiar by at least 79% of 4- to 6-year-olds tested, with similar familiarity for the simple and complex objects (see Supplement 2).

Our pairs included objects of varying complexity and technological sophistication that we anticipated would produce strong Learn scores (but not necessarily strong Insides scores) for all ages. Some pairs involved contrasts between technological and nontechnological devices; such contrasts may be obvious to adults but are not necessarily obvious to children, and they do represent meaningful distinctions between objects' complexity. Children may have strong Learn scores but weak Insides scores if they are insensitive to the internal complexity cues used as stimuli, lack expectations regarding the insides of real-world objects, or fail to consider causal complexity when considering insides.

Insides Block. The experimenter explained that objects have "inside parts [that] help objects work [and] make things happen with the other parts," emphasizing the insides' causal role. First, he showed simple insides and said, "Some objects have inside parts that look kind of like this, [with] one, two kinds of parts inside." Next, he showed complex insides and said, "Some objects have inside parts that look kind of like this, [with] one, two, three, four, five different kinds of connected parts inside." The insides pictures were oriented vertically in front of participants; whichever image was closer to children was randomized for each participant. As a comprehension check and to introduce the action of placing pictures on the iPad, the experimenter asked participants to first indicate which picture had two kinds of parts versus five different kinds of connected parts and then place smaller versions of each image on top of a new set of identical images, oriented vertically and displayed on the iPad.

The Insides Block had eight forced-choice items. For each item, the two objects were labeled verbally by the experimenter and depicted visually in a horizontal orientation on the iPad, with item order and left–right orientation randomized for each participant. The experimenter asked participants to indicate which insides picture looked most like each object's insides by placing the appropriate picture below each object. Most participants made standard forced-choice matches, but those who did not were prompted to do so. For each item, participants received a score of 0 (incorrect) or 1 (correct; matching complex insides with complex objects, shown in Appendix B).

Learn Block. Participants heard explanations of two laminated pictures orientated vertically. The *easy* picture depicted a schematic profile of a mostly empty human head, and the *hard* picture depicted the head filled with dark coloring to represent "need[ing] to learn a whole lot to learn about how [something] work[s]." (Piloting found that young children easily use and understand the images.) Next, as a comprehension check, participants were asked to identify the hard picture. Afterward, participants proceeded to the Learn Block and were asked to place the hard picture below the object "hardest to learn about [in terms of] how it works" as a measure of causal complexity. Participants saw the eight items in a new randomized order and made forced-choice matches.

Materials and procedure for adults

Testing of adult participants was conducted online using Qualtrics. Instructions were conveyed via on-screen text. The script was streamlined and certain features were modified for online administration, but methods were otherwise similar to those used with children for all studies reported here. Additional questions were asked to screen out inattentive participants.

Results

General analysis approach

To analyze the data, we built mixed-effects logistic regression models using the lme4 package (Bates et al., 2015) in R Version 4.0.0 (R Core Team, 2020). The dependent variable across all studies was accuracy of responses (binary; 1 if correct). Our main predictors across studies included age group (factor; grouping varies by study) and block type (factor; Insides or Learn). We further controlled for item (factor; number of items varies by study) and block order (binary; Insides Block presented first coded as 1) in all models. Additional predictor variables are described in the relevant sections. Due to repeated measures within participants, participant ID numbers were fit as random intercepts in all our models. For predictions of estimates and 95% confidence intervals (CIs), we used the effects package (Fox, 2003). Because of our key interest in children's data and expectation that adults would achieve near-ceiling scores, we included only the children's data for the models reported in text. However, in the interest of allowing for visual comparisons with children's performance, we added adult data to the models used for data visualizations. Aside from the inclusion of adults, the models used for data visualizations are identical to those reported in text. (Overall results do not differ with the inclusion of adults' data, although predicted estimates and CIs vary slightly.)

Study 1 results

To assess children's performance in this study, we followed the approach described earlier to build mixed-effects logistic regression models predicting accuracy by age group and block type when controlling for block order and item (factor; eight levels). When comparing the accuracy of responses with chance (50%) across age groups and blocks, we found that only 4-year-olds failed to perform better than chance on the Insides Block (predicted 95% CI [0.41, 0.59] encompasses 0.50). Four-year-olds on the Learn Block (predicted 95% CI [0.53, 0.70]), and all other age groups on both blocks, performed better than chance (all lower bounds of 95% CIs above 0.50). Fig. 1 visualizes predicted accuracy by block and age group, including adults, who are virtually at ceiling. Item-specific results are shown in Supplements 3 and 4.

We further found that accuracy did not significantly vary across block type (odds ratio [*OR*] = 1.27, 95% CI [0.96, 1.69], z = 1.66, p = .10), nor was there an interaction between block type and age group [likelihood ratio test (LRT), $\chi^2(2)$ = 1.55, p = .46]. We did find age-related increases in accuracy, such that when compared with 4-year-olds, 5-year-olds (*OR* = 1.65, 95% CI [1.09, 2.50], z = 2.35, p = .02) and 6-year-olds (*OR* = 3.03, 95% CI [1.95, 4.70], z = 4.95, p < .001) were significantly more likely to respond accurately.

Discussion

All age groups scored above chance in the Learn Block. Although the ability to make causal complexity judgments was relatively fragile in 4-year-olds, children of this age were still likely to agree with adults about which objects were complex. Our findings echo those of Kominsky et al. (2018), but here we document success in younger children. Knowing exactly how things work is not required

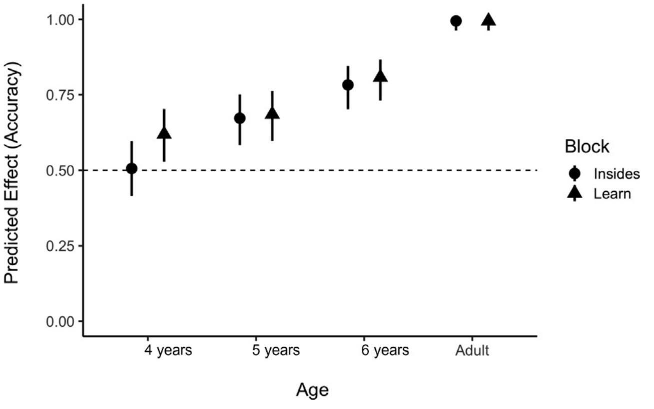

Fig. 1. Predicted effects of block type and age on participants' Study 1 accuracy. Error bars show 95% confidence intervals.

for an abstract sense of which things are causally complex. Our data do not indicate the sources of children's intuitions about causal complexity. However, we speculate that children may learn through interacting with real objects (Kelemen, Seston, & Saint Georges, 2012), adults' responses to mechanistic questions (Frazier, Gelman, & Wellman, 2009), tracking and recall of features such as number and diversity of functions (Ahl & Keil, 2017), and observing whether functions arise through obvious versus nonobvious mechanisms (Wu, Muentener, & Schulz, 2013). Opportunities for learning and abilities to make sense of such learning increase with age, contributing to age-related changes in children's performance.

Our wording for causal complexity questions emphasized the difficulty of learning "how objects work," that is, their causal mechanisms (see Kominsky et al., 2018). One concern is that this prompt did not tell participants what specific factors to consider. Participants may have interpreted this wording differently, for example, how hard it is to use the object. (We addressed this issue via modified wording in our Study 2 script.) Despite this, children and adults generally chose similar objects as complex, and Learn scores provide a comparison with Insides scores: Do age groups that correctly identify complex objects also match such objects with more complex internal parts?

Four-year-olds' scores were above chance in the Learn Block only, whereas older age groups scored well in both blocks. Why did 4-year-olds fail to achieve above-chance Insides scores? One possibility, which we believe can be dismissed, is that they failed to appreciate the causal role of insides. Prior research indicates that preschoolers view insides as important, a point also emphasized in our script. Another possibility is that they did not view our insides pictures as conveying complexity contrasts, but an additional study (Supplement 5) indicates that 4-year-olds usually identified complex insides as such when explicitly asked to do so. We believe that this age group's Study 1 performance is best attributed to difficulties with simultaneously and spontaneously considering real-world objects' internal and causal complexity rather than total insensitivity to internal complexity cues; perhaps 4-year-olds require explicit cuing in order to reflect on visual and causal complexity contrasts. However, we cannot rule out the possibility that 4-year-olds uniquely struggled with the Insides task's abstraction and lack of realism (e.g., matching electronic insides pictures with nonelectronic objects), even as older children succeeded, raising the possibility that 4-year-olds would have performed better had different insides images been used.

Regardless of why 4-year-olds had low Insides scores, which *specific* internal features do older children associate with complex objects? In Study 2, we separately tested the cues of number/area, diversity, and connections of parts. We believed that testing cues separately would result in a more difficult task with weaker effect sizes. Thus, we set our sample size at 26 participants per age group and condition, eliminated 4-year-olds (recall their chance-level Study 1 Insides scores, obviating the need to test specific cues), and included two older age groups in addition to adults and 5-year-olds.

We also shortened our task to five items. We did so to reduce study length (some Study 1 participants became bored) and eliminate lower-scoring items while maintaining item diversity (i.e., a range of objects and complexity levels). Study 2 compared success rates for *individual* internal complexity cues, an approach predicated on successful performance with combined cues. Items with lower scores in Study 1 would have been less informative for our inquiry into which individual cues convey complexity. Our item selection process could be expected to inflate performance in Study 2. However, items with high scores in Study 1 should yield high scores in the various Insides conditions of Study 2 only if the given complexity cue is one to which children are sensitive.

Study 2

Method

Participants

Our sample of children included 78 5-year-olds (39 boys; M_{age} = 65.81 months, SD = 3.72, range = 59–72), 78 6- and 7-year-olds (40 boys; M_{age} = 84.69 months, SD = 6.46, range = 72–95), and 78 8- and 9-year-olds (36 boys; M_{age} = 107.79 months, SD = 7.29, range = 97–121) tested in museums (n = 129), private schools (n = 94), or our lab (n = 11). Participants were White (n = 170), Asian (n = 21), Black (n = 13), Latino/a (n = 9), American Indian (n = 1), Middle Eastern (n = 1), or biracial (n = 16); race was unreported for 3 participants. An additional 7 participants were excluded due to comprehension check failures (n = 4), tester error (n = 1), developmental disability (n = 1), or limited English (n = 1). Detailed demographic reporting is provided in Supplement 6. Our adult sample included 78 participants (36 women; M_{age} = 33.51 years, SD = 10.35). Of these adult participants, 31 reported holding a bachelor's degree or higher. An additional 12 participants were excluded for comprehension check failures (n = 10) or very fast completion times (n = 2).

Materials and procedure

Our method was similar to that of Study 1, with a few notable exceptions. Here, the cues of number/area, diversity, and connections of parts were conveyed in separate Insides conditions, with resulting changes to *simple insides* and *complex insides*. The number/area condition presented contrasts of number and area of parts when controlling for diversity and connections, the diversity condition presented contrasts of diversity of parts when controlling for number/area and connections, and the connections condition presented contrasts of connections between parts when controlling for number/area and diversity. Comprehension questions and related scaffolding were done with condition-specific prompts. In the Learn Block, the new script referenced the workings of "inside parts" ("It's really hard to learn about how the inside parts of one of these objects makes it work . . .") to make the Learn questions relate to internal mechanisms and address ambiguities of the corresponding prompt in Study 1. For each item, participants received a score of 0 (incorrect) or 1 (matching the insides with more parts, more diverse parts, and more connected parts with complex objects). Condition and block order were counterbalanced across participants.

At the end of the study, participants identified which of the two insides pictures they viewed previously were "complicated . . . harder to fix . . . build . . . and learn about how it works." Success on this question indicates participants' broad sensitivity to the given cue but does not imply that participants consider this cue in relation to real-world objects. All age groups identified complex insides as such in each condition, as Supplement 7 reports in detail.

Results

Overall results

We followed the approach described earlier to build mixed-effects logistic regression models predicting accuracy by age group, condition (factor; three levels), and block type when controlling for block order and item (factor; five levels). When comparing the accuracy of responses with chance across age groups, blocks, and conditions, we found that 5-year-olds failed to perform better than chance in the diversity and connections conditions of Insides (predicted 95% CIs [0.44, 0.67] and [0.47, 0.71], respectively, encompass 0.50) but performed better than chance in the number/area condition of Insides (predicted 95% CI [0.62, 0.82]). The older age groups performed better than chance for all conditions in Insides, and all age groups performed better than chance in Learn (all lower bounds of 95% CIs above 0.50). Fig. 2 visualizes predicted accuracy in Study 2 by block, condition, and age group, including adults, who were virtually at ceiling. Table 1 displays the raw accuracy scores for each item.

Accuracy was higher in Learn than in Insides (OR = 2.48, 95% CI [1.97, 3.14], z = 7.65, p < .001). We found no evidence of a three-way interaction among age group, condition, and block type [LRT, $\chi^2(4)$ = 6.90, p = .14], nor was there a two-way interaction between age group and condition (p = .52) or block type and condition (p = .54). In spite of the lack of significant interactions, given the importance of Insides scores to our research questions, we subsetted our data to consider how children performed within each block.

In the Insides Block, compared with 5-year-olds, both 6- and 7-year-olds (OR = 2.70, 95% CI [1.70, 4.28], z = 4.20, p < .001) and 8- and 9-year-olds (OR = 4.40, 95% CI [2.69, 7.18], z = 5.93, p < .001) were significantly more likely to respond accurately. Compared with 6- and 7-year-olds, 8- and 9-year-olds were marginally more likely to respond accurately (OR = 1.63, 95% CI [0.988, 2.69], z = 1.91, p = .06). We also found a main effect of condition, such that compared with performance in the number/area condition, participants were less likely to answer accurately in both the diversity condition (OR = 0.41, 95% CI [0.25, 0.67], z = − 3.58, p < .001) and the connections condition (OR = 0.52, 95% CI [0.32, 0.85], z = − 2.61, p = .009).

In the Learn Block, compared with 5-year-olds, both 6- and 7-year-olds (OR = 3.14, 95% CI [1.86, 5.32], z = 4.27, p < .001) and 8- and 9-year-olds (OR = 5.83, 95% CI [3.22, 10.50], z = 5.83, p < .001) were significantly more likely to respond accurately. Compared with 6- and 7-year-olds, 8- and 9-year-olds were marginally more likely to respond accurately (OR = 1.85, 95% CI [0.99, 3.47], z = 1.93, p = .05). We did not find a significant main effect of condition (all ps > .05).

Discussion

In the Learn Block, all age groups scored above chance. Older children performed better than 5-year-olds, and participants performed similarly across conditions. In the Insides Block, however, there were condition differences. Participants in the number/area condition performed significantly better than those in the diversity and connections conditions. Still, older age groups scored above chance in all three conditions, suggesting that individuals of these ages associate each internal complexity

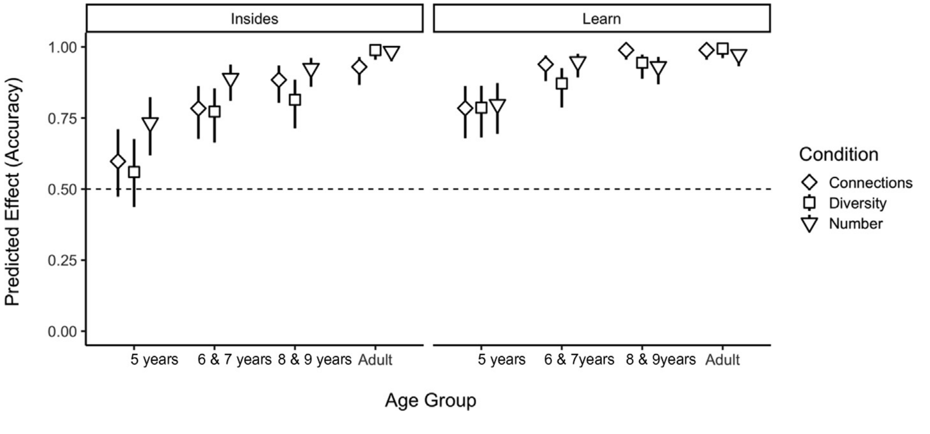

Fig. 2. Predicted effects of block type, condition, and age on participants' Study 2 accuracy. Error bars show 95% confidence intervals.

Table 1
Percentages of children receiving correct scores in Study 2.

Age and item	Condition and blocks		
	Number/area Insides/Learn	Diversity Insides/Learn	Connections Insides/Learn
5-year-olds			
Flashlight vs. smartphone	77/85	39/77	42/69
Table lamp vs. microwave	69/65	65/69	54/69
Jungle gym vs. helicopter	58/81	62/89	65/89
Office chair vs. vacuum	73/69	58/81	69/85
Grandfather clock vs. X-ray machine	73/81	54/65	62/65
6- and 7-year-olds			
Flashlight vs. smartphone	85/89	73/89	81/92
Table lamp vs. microwave	89/100	69/81	65/89
Jungle gym vs. helicopter	92/100	65/89	85/100
Office chair vs. vacuum	77/96	77/85	81/100
Grandfather clock vs. X-ray machine	89/81	85/77	62/77
8- and 9-year-olds			
Flashlight vs. smartphone	89/92	73/85	81/100
Table lamp vs. microwave	92/85	69/96	69/100
Jungle gym vs. helicopter	96/100	85/96	89/92
Office chair vs. vacuum	85/92	81/96	96/100
Grandfather clock vs. X-ray machine	89/85	81/89	92/100

Note. Percentages are rounded to the nearest whole number. n = 26 per age group and condition.

cue with complex real-world objects. School-aged children and adults seem to think that internal complexity has certain features and agree on what they are; complexity means more, rather than less, number/area, diversity, and connections of parts inside.

Five-year-olds, however, performed above chance in the number/area condition only; this cue is both sufficient and necessary for this age. For 5-year-olds, complex objects may simply have a greater number or area of inside parts relative to simpler objects. Their expectations regarding internal complexity lack the detail possessed by older children. The heightened visual salience of the number/area cue in this experimental context could have aided children's success in this condition, but seems unlikely to fully account for it given that differences between simple insides and complex insides were scaffolded in all conditions.

One possibility is that 5-year-olds' success mirrors young children's willingness to make inferences based on information about number, sample size, and related quantitative information (e.g., Lawson & Fisher, 2011), whereas their appreciation of diversity's causal significance is context dependent and develops more gradually throughout the elementary school years (Ahl & Keil, 2017; Rhodes & Liebenson, 2015). A focus on quantitative and numerical information is often supported by children's teachers (McMullen, Chan, Mazzocco, & Hannula-Sormunen, 2019), raising the possibility that children would spontaneously attend to number when encountering real insides and also in experimental contexts, facilitating success in the number/area condition; the cues of diversity and connections seem unlikely to be the targets of extensive explicit teaching from teachers, particularly for young children.

Another possibility is that 5-year-olds' success in the number/area condition, and this condition's overall higher scores, is attributable to external features of our stimuli. One version of this account holds that visual features of our stimuli or their corresponding real-world objects encouraged a perceptual matching strategy specific to the number/area condition. This strategy would consist of matching more inside parts with objects that have more external parts or buttons (such objects were generally the complex objects). In light of this, we conducted Study 4 with new object pairs and methodological changes (Supplements 8–11). We replicated Study 2's findings; children aged 5 years and older expected complex objects to have a greater number/area of inside parts. However, the greater number of external visual features on some of our complex objects may have facilitated children's success in this Study 2 condition and may have contributed to high scores in Study 4 as well.

Although we attempted to control our visual stimuli on this dimension via simplified line drawings, differences in the objects themselves, which children may represent when performing the task, could have influenced children's scores. Our findings must be qualified by the limitations of our object pairs, particularly regarding the number/area cue.

Another version of the external features account holds that our complex objects differed from our simpler objects specifically in their number of external features *that affect internal functioning,* and this difference contributed to the number/area condition's higher scores. In other words, the fact that a microwave has many buttons *offering opportunities to interact with its internal parts* presents problems for its contrast with its paired simple object (a table lamp). This account could also explain why the number/area condition outperformed the others, and it remains unaddressed by our Study 4 data. Because complex artifacts often have multiple external components that influence internal functions, this property is not restricted to the specific complex objects we chose but rather is true of complex artifacts more generally, although it nonetheless complicates interpretations of our findings. Therefore, further investigation of this issue is best achieved with the use of novel artifacts (e.g., Erb et al., 2013), which allow for a level of control that is difficult to achieve with real-world objects. We suspect that when external opportunities for interaction are not truly indicative of underlying causal complexity (e.g., a child's toy with many buttons to press but little genuine mechanistic complexity), children aged 5 and older would not mistakenly attribute high levels of causal and internal complexity to such objects.

Another issue concerns whether number versus area of parts was responsible for participants' success in the number/area condition; complex insides had more parts *and* a larger area of parts relative to simple insides. Therefore, it is unclear whether children believe that complex objects have a greater number, or simply a greater area, of inside parts than simple objects; the latter is a cruder heuristic. Thus, to test the unique contribution of number as a complexity cue, we presented a contrast of number while holding area constant. In Study 3, *complex insides* had several small parts, whereas *simple insides* had a single large part similar in area to the combined area of complex insides' small parts. We tested 6- and 7-year-olds, 8- and 9-year-olds, and adults. (We did not test 5-year-olds because we viewed the subtler cue of number alone as challenging, and their Study 2 number/area performance was weaker than that of older children.) If children expect more complex objects to have a greater number of internal parts, they should match complex insides, which has more parts, with complex objects.

Study 3

Method

Participants

Our sample of children included 26 6- and 7-year-olds (13 boys; M_{age} = 83.89 months, SD = 7.46, range = 72–95) and 26 8- and 9-year-olds (12 boys; M_{age} = 105.04 months, SD = 6.86, range = 96–19) tested in private schools (n = 49) and museums (n = 3). Participants were White (n = 38), Black (n = 6), Asian (n = 2), Latino (n = 1), or biracial (n = 5). An additional 3 participants were excluded due to comprehension check failures (n = 1), contamination from another participant (n = 1), or severe inattention (n = 1). Our adult sample included 26 participants (14 women; M_{age} = 38.65 years, SD = 11.45). Of these adult participants, 13 reported holding a bachelor's degree or higher. An additional 7 participants were excluded for comprehension check failures (n = 3) or very fast completion times (n = 4).

Materials and procedure

Our method was similar to that of Study 2, with the exception of different insides images and related script changes. *Simple insides* had a battery and a single yellow part that was digitally enlarged to occupy a similar area as the four smaller yellow parts in complex insides. *Complex insides* had a battery and four yellow parts. Thus, both images had a similar area of parts, although complex insides had more parts. For each item, participants received a score of 0 or 1 (matching complex insides with complex objects).

Results

We followed the approach described earlier to build mixed-effects logistic regression models predicting accuracy by age group and block type when controlling for block order and item (factor; five levels). When comparing the accuracy of responses with chance across age groups and blocks, we found that all groups performed better than chance (all lower bounds of 95% CIs above 0.50). Fig. 3 visualizes predicted accuracy in Study 3, including adults who were virtually at ceiling. Participants performed significantly better in Learn than in Insides (OR = 4.54, 95% CI [2.75, 7.50], z = 5.92, p < .001). We did not find a significant difference in accuracy between age groups (OR = 1.51, 95% CI [0.86, 2.64], z = 1.45, p = .15).

Because Study 3 and Study 2 were conducted at different times, with only partial overlap in testing locations, a between-study comparison must be interpreted cautiously. Nonetheless, with this analysis, we found a significant two-way interaction between study and block [LRT, $\chi^2(1)$ = 7.30, p = .007] for 6- to 9-year-olds (the age range common to both studies). We found that whereas Learn accuracy was similar in Study 2′s number/area condition as in Study 3 (OR = 1.28, 95% CI [0.66, 2.49], z = 0.73, p = .47), Insides accuracy was higher in Study 2′s number/area condition than in Study 3 (OR = 3.75, 95% CI [2.18, 6.45], z = 4.78, p < .001) (see Supplement 12 for a visualization).

Discussion

All age groups achieved above-chance Insides scores, indicating that number of parts, above and beyond area, is viewed as an internal complexity cue. Yet, children's Study 3 Insides scores were approximately 1 point lower than in Study 2′s number/area condition; some of this condition's strength may be attributable to complex insides' greater area of parts relative to simple insides. Children aged 6 years and older seem to view complex objects' insides as having both "more stuff" and "a greater number of parts" compared with simpler objects. One alternative explanation for children's lower Study 3 Insides scores is that the large size of simple insides' yellow part caused children to view this part as special and reduced their focus on the numerical contrast with complex insides. We cannot determine how much of the cross-study Insides score difference is attributable to this feature. We did not test 5-year-olds, another limitation of this study; however, based on older children's results, we

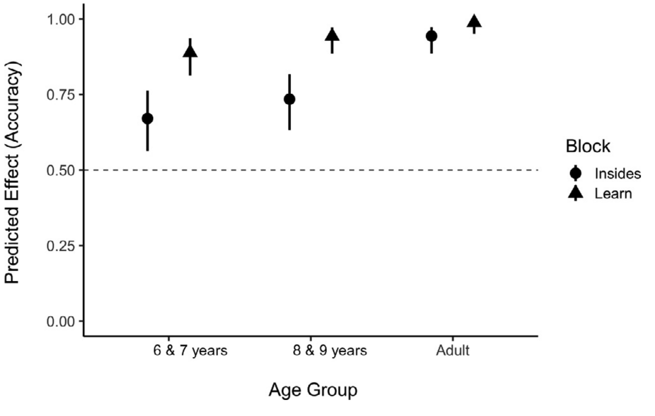

Fig. 3. Predicted effects of block type and age on participants' Study 3 accuracy. Error bars show 95% confidence intervals.

believe that 5-year-olds would have attained chance-level scores had they been tested. Study 3 indicates that, during the early school years, children are sensitive to the cue of number of parts in addition to the cue of area, revealing a finer distinction than was shown in Study 2.

General discussion

Our studies are the first to document the internal complexity features children expect in complex real-world objects and to compare judgments of internal and causal complexity. We demonstrate a clear age-based progression of complexity cue sensitivity. Four-year-olds achieved above-chance scores when comparing objects' causal complexity but achieved only chance-level scores when comparing objects' internal complexity. Five-year-olds succeeded at both judgments; however, children of this age possess a simple "more is more" heuristic regarding internal complexity. Starting at 6 and 7 years of age, children expect complex objects' insides to have a greater area of parts, a greater number of parts, more diverse parts, and more connections between parts relative to simpler objects, thereby surpassing 5-year-olds' heuristic. Such intuitions regarding internal complexity cues align with those of adult participants.

During the early elementary school years, children have specific yet abstract expectations for internal complexity in the domain of artifacts that extend beyond expectations regarding ontological category (cf. Simons & Keil, 1995). Given that children deem insides to be important, it is unsurprising that children would represent something about them. However, merely knowing that insides matter does not mean that children will associate specific internal features with complex objects, let alone features that adults also view as internal complexity cues.

It is unlikely that our participants were previously taught about our complexity cues or our objects' real insides. Instead, our stimuli appeared to elicit intuitive theories that were abstract and generative enough to be applied during our study. Building beliefs about insides is challenging; insides are difficult to observe and act upon. Despite this, children eventually hold differentiated expectations about insides. Several factors, which likely contribute to age-related performance improvements, may be involved in this process. Children may use general principles of mechanical causality regarding how observable entities interact (Au & Romo, 1999; Springer & Keil, 1989) to infer the workings of unseen entities through a process of analogical transfer. For instance, they may view connected elements as implying more causal interactions than disconnected ones and may apply this understanding to insides. In addition, the early-emerging insight that insides matter, combined with children's general preference for causal information (Alvarez & Booth, 2015), may lead children to disproportionately seek and retain information about internal properties, adding more details to the "black box" of insides as they grow older. Children may get glimpses into internal complexity by, for instance, encountering devices' insides during repairs, and they may extract enough from such experiences to develop a framework that can be applied to unfamiliar objects.

Echoing findings from other studies on artifact-based cognition (Ahl & Keil, 2017; Neldner et al., 2019), older children outperformed younger children. In addition to factors mentioned above and in the Discussions of Studies 1 and 2, reasons for children's age-related improvements include domain-general cognitive changes (e.g., improved executive function and analogical reasoning skills, with such improvements supporting the ability to simultaneously consider causal complexity and its implications for insides, an ability that appears to elude 4-year-olds) and domain-specific conceptual developments (e.g., greater knowledge base of mechanistic information, more experiences with insides). Structured explorations and conversations about artifacts and insides surely accumulate with age and can occur in science museums (Benjamin, Haden, & Wilkerson, 2010), some of which have exhibits devoted to the insides of artifacts (e.g., a traveling exhibition called *Toys: The Inside Story*). Domain-specific knowledge and domain-general skill improvements likely work in tandem to allow older children to make sense of such experiences (see Bascandziev, Tardiff, Zaitchik, & Carey, 2018). The aforementioned factors may also account for individual differences between same-aged children. We suspect that children gradually develop abstract skeletal models of different causal patterns and link them to artifacts of varying levels of complexity. As such models become articulated with age, children can deploy them to make complexity judgments with the aid of analogies between novel cases and stored exemplars facilitated by executive function abilities. Information about an artifact's complexity therefore can evoke specific expectations regarding its internal properties. The convergence of developmental changes may allow increasing insights into how interior structures imply different levels of causal complexity.

Our studies focused on causal, rather than purely visual, complexity. However, these different kinds of complexity can overlap. In previous research on pure visual complexity judgments, number and diversity of elements, which conveyed causal complexity in our studies, conveyed visual complexity to children and adults (Chipman & Mendelson, 1975, 1979; Tinio & Leder, 2009). Despite this, visual and causal complexity cues are not equivalent; five-year-olds matched complex insides with complex objects in the number/area condition alone (4-year-olds never did). In addition, more symmetrical images are viewed as less visually complex than less symmetrical images (Chipman & Mendelson, 1979), but we suspect that symmetry, which implies planning, signifies a causally complex system. Exploring how judgments regarding visual and causal complexity may diverge is a topic for further study.

We made design choices to allow for success at causal complexity judgments, such as choosing artifacts known to American children. Exploring which objects do and do not lead to adult-like causal complexity judgments, and why, is a topic for future research. However, our focus was on internal complexity cues *given* recognition of their causal complexity. Success at causal complexity judgments with our specific object pairs might not generalize to other objects. (We speculate that if children identify one object in a pair as more causally complex, they would also be likely to identify that object as more internally complex, a pattern we have seen across the current studies.) Our results do not demonstrate what kinds of information children value when making causal complexity judgments (e.g., how hard objects are to fix; Kominsky et al., 2018). Particularly given the difficulties of studying the inherently complicated topic of complexity, more research is needed on the extent to which children attend to genuine indications of complexity.

Some object pairs contrasted technological and nontechnological objects, whereas others contrasted two technological objects. The small number of object pairs of the latter kind is a limitation of Study 2. More generally, the small number of items in our studies prevented us from exploring how children may respond differently to complexity contrast subtypes. However, although contrasts between technological and nontechnological objects may seem easier to adults, children's success in our studies was not driven solely by such contrasts; the smartphone versus the flashlight and the microwave versus the table lamp object pairs performed well. In addition, any obviousness in causal complexity judgments would not explain condition- and age-specific differences in internal complexity judgments. Another limitation of our study was that our insides were not veridical. We used simplified pictures to achieve uniformity across stimulus items and to reduce demands on working memory and quantitative reasoning. How participants would perform on tasks involving realistic insides is an open question. Despite this, we can conclude that the insides' lack of realism did not pose substantial problems for participants older than 4 years, as judged by their high scores, nor could it straightforwardly account for age- and condition-specific differences in children aged 5 years and older. The technological appearance of the insides was common to both simple and complex insides and thus could not, on its own, lead participants to correct insides matches. Further studies can work to address the limitations surrounding our insides and object pairs, either with the use of more realistic insides or with the use of novel artifacts rather than real-world artifacts.

Might our results generalize beyond predominantly middle-class, American children? Our procedure is predicated on familiarity with our object pairs. Children without exposure to such artifacts would struggle with our tasks, but whether children across cultures would generally link internal complexity to functional complexity is an open question. Children across cultures capably use and manufacture tools (Neldner et al., 2019) and arguably possess special cognitive adaptations (Hernik & Csibra, 2009) that may aid in the understanding of artifacts. However, children from industrialized cultures that provide exposure to diverse artifacts have advantages with certain kinds of artifact-based reasoning relative to children from nonindustrialized cultures (Neldner et al., 2019) and may have more experiences with artifacts' insides. In the domain of biology, by contrast, the experiences of children from nonindustrialized cultures may lead to advantages in insights about biological complexity.

Do children view number/area, diversity, and connections of internal parts as signs of complexity in the natural world? Infants understand the importance of insides to animals (Newman et al., 2008), with little comparable evidence regarding machines. However, knowledge from the biological domain may be transferred to artifacts and vice versa (Kelemen & Carey, 2007). The initial understanding that insides matter could be transferred from animals to complex artifacts at a young age. In turn, perhaps children's experiences with mechanistic interventions in artifacts could be used to form expectations regarding animals' insides. Whether children hold beliefs about animals' internal complexity is an open question, particularly given the domain-specific causal patterns in biology and their construal by human cognition (Carey, 1985).

The industrialized world is full of devices whose mechanisms are opaque, intricate, and hidden. Given this complexity, one might expect children to have no expectations regarding the inner workings of such objects. However, rather than merely positing a complex essence that brings about an object's functionality, children have specific expectations for how complexity is instantiated inside. Such expectations show a developmental trajectory of increasing differentiation and ultimately converge with those of adults. Years before children take classes in physics or engineering, they nonetheless possess abstract and generative expectations about the functional configuration of complex objects' insides. These expectations may set the path for informal causal explorations into how objects function as well as for formal scientific and technological learning in secondary school and beyond.

Acknowledgments

This research was supported by a National Science Foundation grant (DRL 1561143) awarded to Frank C. Keil. Portions of the data were presented as a poster at the 2017 Biennial Meeting of the Society for Research on Child Development (Study 1) and as a talk at the 2017 Conference of the Cognitive Development Society (Study 2). We thank Kayla Good for stimulus preparation, data collection, participant recruitment, and the design of the study reported in Supplement 5; Rachel Han and Paige Cunningham for participant recruitment; and Alexander Wisowaty for data collection and participant recruitment. We thank our participating children, their families, and the administrators, teachers, and staff at our research sites, including Abington Friends School, Pine Point School, St. Thomas's Day School (New Haven, CT), Kinder Works of Doylestown, PA, Yale Peabody Museum of Natural History, and Living Laboratory® at Stepping Stones Museum for Children.

Appendix A. Descriptions of insides pictures.

Study	Simple insides	Complex insides
Study 1	One battery, one yellow part	One battery, one yellow part, plus three unique parts with connecting wires
Study 2: Number/area	One battery, one yellow part	One battery, four yellow parts
Study 2: Diversity	One battery, four yellow parts	One battery, one yellow part, plus three unique parts
Study 2: Connections	One battery, four yellow parts	One battery, four yellow parts with connecting wires
Study 3	One battery, one large yellow part	One battery, four yellow parts

Appendix B. Items for studies 1, 2, and 3

	Study 1	Study 2 and Study 3
1. Flashlight vs. smartphone	x	x
2. Toaster vs. computer	x	
3. Table lamp vs. microwave	x	x
4. Jungle gym vs. helicopter	x	x
5. Office chair vs. vacuum cleaner	x	x
6. Beanbag chair vs. dishwasher	x	
7. Grandfather clock vs. X-ray machine	x	x
8. Clothing iron vs. video camera	x	

Note. Within each pair, the simple objects are listed first and the complex objects are listed second here, but presentation order was randomized for each participant.

Appendix C. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2020.104932.

References

Ahl, R. E., & Keil, F. C. (2017). Diverse effects, complex causes: Children use information about machines' functional diversity to infer internal complexity. *Child Development, 88*, 828–845.

Alvarez, A. L., & Booth, A. E. (2015). Preschoolers prefer to learn causal information. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.00060.

Arbesman, S. (2016). *Overcomplicated: Technology at the limits of comprehension*. New York: Penguin Books.

Au, T. K., & Romo, L. F. (1999). Mechanical causality in children's "folkbiology". In D. L. Medin & S. Atran (Eds.), *Folkbiology* (pp. 355–401). Cambridge, MA: MIT Press.

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.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., ... Bolker, M. B. (2015). Package lme4. *Convergence, 12*, 2.

Benjamin, N., Haden, C. A., & Wilkerson, E. (2010). Enhancing building, conversation, and learning through caregiver–child interactions in a children's museum. *Developmental Psychology, 46*, 502–515.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Chipman, S. F., & Mendelson, M. J. (1975). The development of sensitivity to visual structure. *Journal of Experimental Child Psychology, 20*, 411–429.

Chipman, S. F., & Mendelson, M. J. (1979). Influence of six types of visual structure on complexity judgments in children and adults. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 365–378.

Erb, C. D., Buchanan, D. W., & Sobel, D. M. (2013). Children's developing understanding of the relation between variable causal efficacy and mechanistic complexity. *Cognition, 129*, 494–500.

Fourez, G. (1997). Scientific and technological literacy as a social practice. *Social Studies of Science, 27*, 903–936.

Fox, J. (2003). Effect displays in R for generalised linear models. *Journal of Statistical Software, 8*, 1–27.

Frazier, B. N., Gelman, S. A., & Wellman, H. M. (2009). Preschoolers' search for explanatory information within adult–child conversation. *Child Development, 80*, 1592–1611.

Gelman, R. (1990). First principles organize attention to and learning about relevant data: Number and the animate–inanimate distinction as examples. *Cognitive Science, 14*, 79–106.

Gelman, S. A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology, 20*, 65–95.

Gelman, S. A., & Kremer, K. E. (1991). Understanding natural cause: Children's explanations of how objects and their properties originate. *Child Development, 62*, 396–414.

Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition, 38*, 213–244.

Gottfried, G. M., & Gelman, S. A. (2005). Developing domain-specific causal-explanatory frameworks: The role of insides and immanence. *Cognitive Development, 20*, 137–158.

Grotzer, T. A., & Tutwiler, S. M. (2014). Simplifying causal complexity: How interactions between modes of causal induction and information availability lead to heuristic-driven reasoning. *Mind, Brain, and Education, 8*, 97–114.

Hernik, M., & Csibra, G. (2009). Functional understanding facilitates learning about tools in human children. *Current Opinion in Neurobiology, 19*, 34–38.

Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.

Keil, F. C. (2003). Folkscience: Coarse interpretations of a complex reality. *Trends in Cognitive Sciences, 7*, 368–373.

Kelemen, D., & Carey, S. (2007). The essence of artifacts: Developing the design stance. In S. Laurence & E. Margolis (Eds.), *Creations of the mind: Theories of artifacts and their representation* (pp. 212–230). Oxford, UK: Oxford University Press.

Kelemen, D., Seston, R., & Saint Georges, L. (2012). The designing mind: Children's reasoning about intended function and artifact structure. *Journal of Cognition and Development, 13*, 439–453.
Kominsky, J. F., Zamm, A. P., & Keil, F. C. (2018). Knowing when help is needed: A developing sense of causal complexity. *Cognitive Science, 42*, 491–523.
Lawson, C. A., & Fisher, A. V. (2011). It's in the sample: The effects of sample size and sample diversity on the breadth of inductive generalization. *Journal of Experimental Child Psychology, 110*, 499–519.
McMullen, J., Chan, J. Y. C., Mazzocco, M. M., & Hannula-Sormunen, M. M. (2019). Spontaneous mathematical focusing tendencies in mathematical development and education. In A. Norton & M. W. Alibali (Eds.), *Constructing number* (pp. 69–86). Cham, Switzerland: Springer.
Neldner, K., Redshaw, J., Murphy, S., Tomaselli, K., Davis, J., Dixson, B., & Nielsen, M. (2019). Creation across culture: Children's tool innovation is influenced by cultural and developmental factors. *Developmental Psychology, 55*, 877–899.
Newman, G. E., Herrmann, P., Wynn, K., & Keil, F. C. (2008). Biases towards internal features in infants' reasoning about objects. *Cognition, 107*, 420–432.
R Core Team (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
Rhodes, M., & Liebenson, P. (2015). Continuity and change in the development of category-based induction: The test case of diversity-based reasoning. *Cognitive Psychology, 82*, 74–95.
Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science, 26*, 521–562.
Setoh, P., Wu, D., Baillargeon, R., & Gelman, R. (2013). Young infants have biological expectations about animals. *Proceedings of the National Academy of Sciences of the United States of America, 110*, 15937–15942.
Simon, H. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society, 106*, 467–482.
Simons, D. J., & Keil, F. C. (1995). An abstract to concrete shift in the development of biological thought: The insides story. *Cognition, 56*, 129–163.
Sobel, D. M., Yoachim, C. M., Gopnik, A., Meltzoff, A. N., & Blumenthal, E. J. (2007). The blicket within: Preschoolers' inferences about insides and causes. *Journal of Cognition and Development, 8*, 159–182.
Springer, K. (2001). Perceptual boundedness and perceptual support in conceptual development. *Psychological Review, 108*, 691–708.
Springer, K., & Keil, F. C. (1989). On the development of biologically specific beliefs: The case of inheritance. *Child Development, 60*, 637–648.
Taborda-Osorio, H., & Cheries, E. W. (2017). Developmental origins of biological explanations: The case of infants' internal property bias. *Psychonomic Bulletin & Review, 24*, 1527–1537.
Tinio, P. P., & Leder, H. (2009). Just how stable are stable aesthetic features? Symmetry, complexity, and the jaws of massive familiarization. *Acta Psychologica, 130*, 241–250.
Trouche, E., Chuey, A., Lockhart, K. L., & Keil, F. C. (2017). Why teach how things work? Tracking the evolution of children's intuitions about complexity. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 3368–3373). Austin, TX: Cognitive Science Society.
Walker, C. M., Lombrozo, T., Legare, C. H., & Gopnik, A. (2014). Explaining prompts children to privilege inductively rich properties. *Cognition, 133*, 343–357.
Welder, A. N., & Graham, S. A. (2001). The influence of shape similarity and shared labels on infants' inductive inferences about nonobvious object properties. *Child Development, 72*, 1653–1673.
Wu, Y., Muentener, P. J., & Schulz, L. (2013). The invisible hand: Toddlers infer hidden agents when events occur probabilistically. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 3807–3810). Austin, TX: Cognitive Science Society.