Cognitive Psychology 125 (2021) 101368

Contents lists available at ScienceDirect

Cognitive Psychology

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

How do the object-file and physical-reasoning systems interact? Evidence from priming effects with object arrays or novel labels

Yi Lin[a,*], Jie Li[b], Yael Gertner[c], Weiting Ng[d], Cynthia L. Fisher[a], Renée Baillargeon[a]

[a] Department of Psychology, University of Illinois at Urbana-Champaign, IL 61820, USA
[b] Department of Graduate Medical Education, Stanford University, CA 94305, USA
[c] Department of Computer Science, University of Illinois at Urbana-Champaign, IL 61801, USA
[d] School of Humanities and Behavioural Sciences, Singapore University of Social Sciences, Singapore

ARTICLE INFO

Keywords:
Infant cognition
Event representation
Physical reasoning
Object-file system
Physical-reasoning system
Priming

ABSTRACT

How do infants reason about simple physical events such as containment, tube, and support events? According to the two-system model, two cognitive systems, the object-file (OF) and physical-reasoning (PR) systems, work together to guide infants' responses to these events. When an event begins, the OF system sends categorical information about the objects and their arrangements to the PR system. This system then categorizes the event, assigns event roles to the objects, and taps the OF system for information about features previously identified as causally relevant for the event category selected. All of the categorical and featural information included in the event's representation is interpreted by the PR system's domain knowledge, which includes core principles such as persistence and gravity. The present research tested a novel prediction of the model: If the OF system could be primed to also send, at the beginning of an event, information about an as-yet-unidentified feature, the PR system would then interpret this information using its core principles, allowing infants to detect core violations involving the feature earlier than they normally would. We examined this prediction using two types of priming manipulations directed at the OF system, object arrays and novel labels. In six experiments, infants aged 7–13 months ($N = 304$) were tested using different event categories and as-yet-unidentified features (color in containment events, height in tube events, and proportional distribution in support events) as well as different tasks (violation-of-expectation and action tasks). In each case, infants who were effectively primed reasoned successfully about the as-yet-unidentified feature, sometimes as early as six months before they would typically do so. These converging results provide strong support for the two-system model and for the claim that uncovering how the OF and PR systems represent and exchange information is essential for understanding how infants respond to physical events.

1. Introduction

How do infants represent the physical world and reason about it? How do these capacities guide their expectations about events' outcomes and their prospective actions on objects? And how do these capacities improve with experience, resulting in more accurate

* Corresponding author.
E-mail address: yilin5@illinois.edu (Y. Lin).

https://doi.org/10.1016/j.cogpsych.2020.101368
Received 18 June 2020; Received in revised form 18 December 2020; Accepted 19 December 2020

expectations and better adapted actions? Developmental researchers have long been interested in shedding light on these questions (Baillargeon, 1995; Carey, 2011; Leslie, 1995; Piaget, 1954; Smith et al., 2019; Spelke et al., 1992; Stavans et al., 2019; Téglás et al., 2011; Wang, 2019; Xu, 2019). Much of this research has focused on the simple event categories our distant human ancestors routinely observed and produced, such as occlusion, containment, cover, tube, burying, support, and collision events (for a recent review, see Lin et al., in press). To investigate infants' reasoning about these events, researchers have used a wide range of empirical approaches including search methods (Van de Walle et al., 2000; Zosh & Feigenson, 2015), preferential-reaching methods (Hauf et al., 2012; Hespos & Baillargeon, 2006), and violation-of-expectation (VOE) methods featuring change violations (i.e., objects spontaneously undergo changes that are not physically possible; Kibbe & Leslie, 2019; Wilcox, 1999), interaction violations (i.e., objects interact in ways that are not physically possible given their respective properties; Baillargeon, 1987; Mou & Luo, 2017), and individuation violations (i.e., fewer objects are revealed at the end of an event than were presented during the event; Stavans & Baillargeon, 2018; Xu & Carey, 1996). Over the past few decades, investigations using these methods have produced several sets of puzzling and even seemingly contradictory findings. In each set, infants are shown to succeed at some tasks but to fail at other very similar tasks. Below, we describe three such sets of findings.

First, young infants often succeed at tasks that require reasoning about broad, *categorical* information about an object (e.g., its existence, location, and ontological properties), but fail at tasks that require reasoning about more fine-grained, *featural* information about the object (e.g., its size, pattern, and color). For example, consider results from experiments on occlusion events. When an object became hidden behind a screen, infants aged 2.5–6 months detected a violation if the object magically disappeared (Kibbe & Leslie, 2011; Luo et al., 2009), if it reappeared from behind a different screen without having appeared in the gap between them (Aguiar & Baillargeon, 1999; Wilcox et al., 1996), and if it surreptitiously changed ontological category behind the screen (e.g., from a human-like doll's head to a ball; Kibbe & Leslie, 2019). In contrast, infants under 3.5 months did not detect a violation if the object failed to appear in a high opening in the screen (Baillargeon & DeVos, 1991; Luo & Baillargeon, 2005); infants under 7.5 months did not detect a violation if the object changed pattern behind the screen (Wilcox et al., 2011; Wilcox, 1999); and infants under 11.5 months did not detect a violation if the object changed color behind the screen (Káldy & Leslie, 2003; Wilcox & Chapa, 2004). Similar patterns of successes and failures were also obtained with containment (Hespos & Baillargeon, 2001a, 2001b), cover (Wang & Baillargeon, 2006; Wang et al., 2005), tube (Baillargeon, 1995; Wang & Goldman, 2016), burying (Newcombe et al., 1999, 2005), support (Baillargeon et al., 1992; Needham & Baillargeon, 1993), and collision (Baillargeon, 1995; Kotovsky & Baillargeon, 1998) events.

Second, when infants begin to succeed at tasks that require reasoning about particular featural information, they often do so with events from *one category*, but still fail with events from *another category*. For example, 5–6.5-month-olds detected a violation if an object changed shape when behind an occluder (Káldy & Leslie, 2005; Wilcox, 1999) or when inside a container (Wang & Onishi, 2017), but

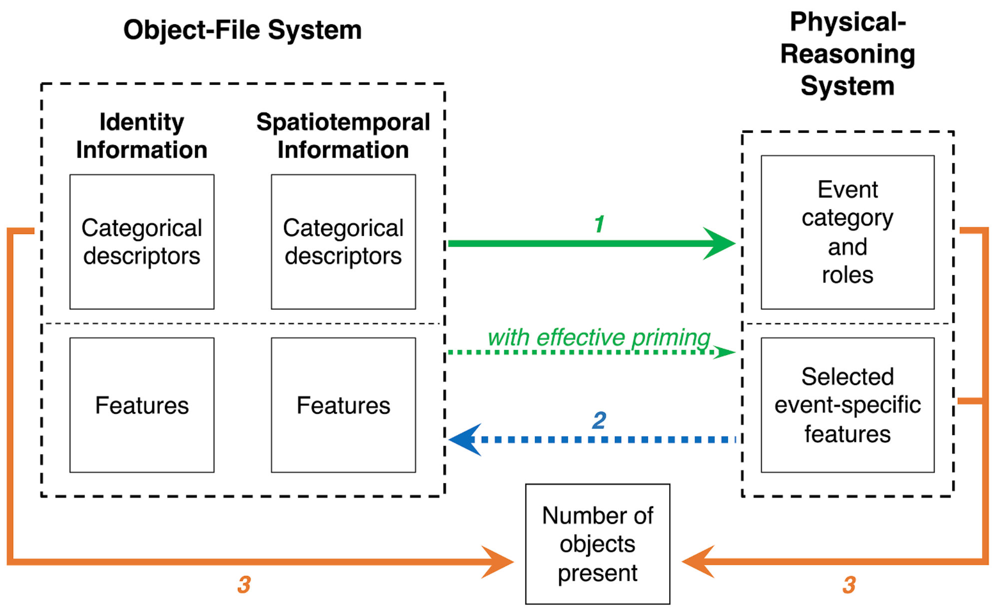

Fig. 1. Schematic depiction of the two-system model. Each numbered arrow illustrates one of the steps the OF and PR systems go through when representing an event. *Arrow 1 (green-solid)*: At the beginning of the event, the OF system spontaneously passes on to the PR system the identity and spatiotemporal categorical descriptors for the objects in the event; the PR system then uses these descriptors to determine the event's category and to assign appropriate event roles to the objects. *Arrow 2 (blue-dashed)*: The PR system taps the OF system for information about the features it has previously identified as causally relevant for the event category; it then interprets this selected featural information using its domain knowledge. *Arrow 3 (orange-solid)*: When the event comes to an end, the two systems must agree on how many objects were present for the OF system to continue tracking the objects past the event's endpoint. Finally, the *green dashed line* below Arrow 1 denotes that with effective priming, the OF system can be induced to also pass on information about as-yet-unidentified features to the PR system. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

not if it changed shape when buried in sand (Newcombe et al., 1999). Similarly, infants as young as 3.5 months could detect violations involving height in occlusion events, as we just saw, but such violations were not detected until much later in other event categories: at about 7.5 months in containment events, 12 months in covering events, and 14 months in tube events (Hespos & Baillargeon, 2001a; Wang & Goldman, 2016; Wang et al., 2005). In preferential-reaching tasks, 6-month-olds correctly searched for a tall frog behind a tall as opposed to a short occluder—but they searched randomly when faced with a tall and a short container (Hespos & Baillargeon, 2006). Similarly, after being "taught" to attend to the feature height in covering events, 9-month-olds searched for a tall toy under a tall as opposed to a short cover—but they searched randomly when faced with a tall and a short tube (Wang & Kohne, 2007).

Finally, when infants begin to succeed at tasks that require reasoning about particular featural information in an event category, they often succeed at tasks that require detecting *change or interaction* violations, but still fail at tasks that require detecting *individuation* violations. Xu and Carey (1996; Xu et al., 2004) were the first to report this baffling failure, and their results were subsequently confirmed in a wide range of individuation tasks (Bonatti et al., 2002; Futó et al., 2010; Leslie et al., 1998; Stavans & Baillargeon, 2018; Surian & Caldi, 2010; Wilcox & Chapa, 2002). In one task, for example, two objects emerged in alternation on either side of a large screen; after several repetitions of this occlusion event, the screen was removed to reveal only one of the objects. When the two objects differed only in their featural properties (e.g., a small red ball with yellow stripes and a large white ball with green dots), infants aged 12 months and younger failed to detect this violation, as though they did not clearly expect to see two objects when the screen was removed (Lin & Baillargeon, 2018; Lin et al., 2019; Stavans et al., 2019; Xu et al., 2004). Thus, although by 12 months most infants can detect change and interaction violations involving size, pattern, and color information in occlusion events, they still fail to detect individuation violations involving these features. Similar failures have also been found with containment events (Stavans et al., 2019; Van de Walle et al., 2000).

It has been a protracted and challenging task for infancy researchers to make sense of these different patterns of successes and failures and to integrate them into a single, coherent account of early physical reasoning (Baillargeon et al., 2012; Carey, 2011; Needham & Baillargeon, 2000; Stavans et al., 2019; Wang & Baillargeon, 2008b; Wilcox et al., 2003; Xu & Carey, 2000; Xu, 2007). According to the recent *two-system* model (Lin et al., in press; Stavans et al., 2019), such an account requires understanding the nature, development, and interactions of two key systems in the cognitive architecture that underlies infants' responses to physical events: the object-file (**OF**) and physical-reasoning (**PR**) systems (Fig. 1). The overarching idea of the model is that the puzzling findings listed above can all be better understood by considering how the two systems exchange and compare information about objects in physical events.

In the following sections, we first describe the OF and PR systems and their interactions. Next, we introduce the present research, which tested a striking prediction of the two-system model: Priming experiences that cause the OF system, at the beginning of an event, to send information about a feature the PR system has not yet learned to include in its event representations should, serendipitously, enable infants to detect core violations involving this feature earlier than they normally would.

1.1. Two systems

The two-system model (Lin et al., in press; Stavans et al., 2019) builds on work in the adult and infant literatures (Casasola, 2008; Gordon & Irwin, 1996; Huttenlocher & Lourenco, 2007; Huttenlocher et al., 1991; Kahneman et al., 1992; Leslie et al., 1998; Pylyshyn, 1989, 2007; Rips et al., 2006; Wang & Baillargeon, 2008b). It assumes that two cognitive systems contribute to early physical reasoning: the OF system (Gordon & Irwin, 1996; Kahneman et al., 1992) and the PR system (Baillargeon et al., 2011; Wang & Baillargeon, 2008b). The two systems serve different functions and have at least partly distinct neural substrates (Fischer et al., 2016; Grill-Spector et al., 2001).

OF system. Imagine a simple static scene involving two objects: A small red ball with yellow stripes rests on a surface to the left of an open box. As infants attend to the scene, the OF system builds a temporary file for each object, drawing on incoming perception as well as on stored knowledge. Each file is created by two closely related processes within the OF system, the identity ("what") and spatiotemporal ("where") processes; each process lists broad, *categorical* information about the object as well as fine-grained, *featural* information. As the names suggest, the identity process lists categorical and featural information about the object's identity (e.g., for the ball, categorical: inanimate, non-human, closed; featural: small, round, striped, red-and-yellow), whereas the spatiotemporal process lists categorical and featural information about the object's location (e.g., for the ball, categorical: occupies the left of the two filled locations on the surface; featural: located a short distance from and slightly in front of the box).[1]

With development, significant changes occur in the information represented by each process (Huttenlocher & Lourenco, 2007;

[1] Over the past 20 years, our research team has produced a succession of cognitive models aimed at explaining some or all of the puzzling physical-reasoning findings listed at the start of the Introduction (Baillargeon et al., 2009, 2011, 2012; Levine & Baillargeon, 2016; Lin et al., in press; Needham & Baillargeon, 2000; Stavans et al., 2019; Wang & Baillargeon, 2008b; Wilcox & Baillargeon, 1998). In earlier, three-system models, the identity and spatiotemporal processes were assumed to constitute two separate systems that each interacted with the PR system (Baillargeon et al., 2011, 2012; Levine & Baillargeon, 2016; Wang & Baillargeon, 2008b). Our current, two-system model (Lin et al., in press; Stavans et al., 2019) acknowledges the differences between these two processes but groups them into a single system to highlight (a) the similar ways in which the processes exchange categorical and featural information with the PR system (see Fig. 1) and (b) the similar ways in which these exchanges can be modified via contextual manipulations, as shown in the present research and elsewhere (Lin et al., 2019). Of course, our primary purpose is to understand infants' physical reasoning, and we recognize that researchers with other purposes might well slice up infants' cognitive architecture differently, for example by returning to a three-system model or by grouping the OF and PR systems into a single, object-related system.

Pauen, 2002; Wilcox & Schweinle, 2003; Xu & Carey, 1996). To illustrate, a key development concerns the categorical descriptors represented by the identity process. Prior to about 12 months, these descriptors typically include only abstract, ontological descriptors, such as whether an object is animate or inanimate, human-like or non-human, and open or closed (Bonatti et al., 2002; Kibbe & Leslie, 2019; Mou & Luo, 2017; Setoh et al., 2013; Surian & Caldi, 2010; Wang et al., 2005). By the first birthday, however, identity descriptors begin to also include objects' basic-level categories (e.g., ball, shoe, spoon), presumably as a result of language acquisition (Bonatti et al., 2002; Cacchione et al., 2013; Schaub et al., 2013; Van de Walle et al., 2000; Wilcox & Baillargeon, 1998; Xu & Carey, 1996). Returning to our example, this means that by 12 months, the identity categorical descriptors for the ball would also include the basic-level descriptor 'ball'.

PR system. If the ball and box become involved in a causal interaction, the OF system updates its files and sends the identity and spatiotemporal categorical descriptors for each object to the PR system, as though activating it. This system is a causal-reasoning system; its function is to reason about causal interactions between objects so as to guide expectations and actions. As the event unfolds, the PR system builds a specialized representation of the event, in three broad steps.

In the first step (see arrow 1 in Fig. 1), the PR system uses the categorical descriptors supplied by the OF system to identify the event category involved (assuming it has already formed that category). For example, the event would be categorized as an occlusion event if the ball was lowered behind the box, as a containment event if it was placed inside the box, and as a collision event if it collided with the box. Once the PR system has categorized the event, it assigns event-specific roles to the objects (e.g., occludee/occluder, containee/container, hittee/hitter; Leslie & Keeble, 1987; Lin & Baillargeon, 2018; Yin & Csibra, 2015).

In the second step (see arrow 2 in Fig. 1), the PR system accesses the list of features it has previously identified (through an explanation-based learning process, described below) as causally relevant for predicting outcomes in the event category selected. Next, it taps the OF system for information about these—and only these—selected features. The retrieved information (e.g., about the relative sizes of the occludee and occluder and about the shape and pattern of the occludee) is then added to the event's representation. All of the information included in this representation is interpreted by the PR system's domain knowledge, which includes core knowledge and acquired rules. Two core principles are particularly relevant to the research reported in this article. One is the *persistence* principle (all other things being equal, objects persist, as they are, in time and space), with its corollaries of solidity, continuity, cohesion, boundedness, and unchangeableness (an object cannot occupy the same space as another object, nor can it spontaneously disappear, break apart, fuse with another object, or change into a different object; Baillargeon, 2008; Baillargeon et al., 2009; Spelke et al., 1992; Spelke et al., 1995). The other is the *gravity* principle (objects fall when inadequately supported; Baillargeon & DeJong, 2017; Baillargeon et al., 1992; Needham & Baillargeon, 1993; Wang et al., 2016). As the PR system interprets the information in its event representation, it guides infants' expectations about how the event will unfold (outcomes that deviate from these expectations are flagged as violations), and it directs infants' exploratory and prospective actions in response to the event (Hauf et al., 2012; Hespos & Baillargeon, 2006; Stahl & Feigenson, 2015; Van de Walle et al., 2000).

Finally, the third step (see arrow 3 from each sysem in Fig. 1) occurs as the event comes to an end and the OF and PR systems compare notes about how many objects are involved in the event (it is as though the PR system, before de-activating, makes sure it is on the same page as the OF system, which will continue to track the objects past the endpoint of the event). Because the OF system uses primarily categorical information to keep track of objects within events, whereas the PR system uses both categorical and featural information, disagreements can occur between the two systems. Returning to our example, if the small red ball with yellow stripes was lowered into the box, and a large white ball with green dots was then lifted from the box, the OF system would assume that it was the same ball because it had the same categorical descriptors, whereas the PR system would assume that these were two different balls because the persistence principle specifies that an object cannot spontaneously change its featural properties. When, at the end of an event, the OF system assumes that a single object is present in a hiding location, but the PR system assumes that two objects are present in that same location, a *catastrophic individuation failure* ensues. The OF system realizes that its object file is corrupted: It does not cleanly refer to a single object in the world but instead contains a tangled mix of information that pertains to two separate objects. The OF system then discards its corrupted file, leading infants to have *no expectation at all* about how many objects will be revealed (Stavans et al., 2019).

Explanation-based learning. How does the PR system learn what features are causally relevant for predicting outcomes in each event category? There is growing evidence that *explanation-based learning* (**EBL**) is one of the key mechanisms in this process (Baillargeon & DeJong, 2017; Wang, 2019; Wang & Baillargeon, 2008a; Wang & Kohne, 2007; for other mechanisms in infants' predictions of physical outcomes, see e.g., Smith et al., 2019; Téglás & Bonatti, 2016; Wang, 2019; Xu, 2019). EBL is typically triggered when infants observe two or more events from the same category that result in contrastive outcomes they cannot explain (e.g., when object A is lowered into container X, it protrudes above the rim; when it is lowered into container Y, however, it does not protrude above the rim). Infants then search for a feature whose values map onto these outcomes (e.g., the relative heights of A and X, and A and Y, predict whether A will protrude or not). Next, infants bring to bear their domain knowledge to generate a plausible explanation for how the selected feature might have contributed to the observed outcomes (e.g., the persistence principle dictates that for A to persist as it is, with its full height, it must protrude above the rim when lowered into X because it is taller than X). This explanation is then generalized, resulting in a candidate rule that incorporates only the feature specified in the explanation (e.g., the relative heights of a container and containee determine whether the containee will protrude above the rim of the container). This candidate rule is then evaluated against further empirical evidence; if it proves accurate in predicting outcomes for a few additional exemplars, it is adopted, and the feature identified by the rule becomes part of the PR system's list of features known to be causally relevant for the event category. From that point on, the PR system routinely taps the OF system for information about the feature when representing events from the category.

As the preceding description suggests, a main assumption of the two-system model is that when it comes to the simple event categories we have been discussing, infants learn primarily *what* features to include in their event representations, not *how* to interpret these features. The PR system's core knowledge, which emerged over the course of evolution to facilitate reasoning and learning about these event categories, allows infants to interpret causally relevant features as soon as these are included in their event representations. Thus, in the preceding example, EBL enabled infants to learn (very efficiently) to include height information when representing containment events; but infants did not need to learn how to interpret that information. As soon as the information that A was taller than X but shorter than Y was included in the events' representations, the PR system could bring to bear its persistence principle to correctly interpret that height information and guide expectations and actions. We return to this point when introducing the prediction tested in the present research.

1.2. *Explanations for puzzling findings*

The two-system model helps explain the three sets of puzzling findings described earlier. First, consider the findings about core violations involving categorical vs. featural information. Because (a) the OF system spontaneously sends categorical but not featural information to the PR system,[2] and (b) the PR system selectively taps the OF system for only those features it has identified as causally relevant for the event category involved, this has two major consequences. One is that even very young infants, whose event representations may be very sparse and include only the categorical information supplied by the OF system, will nonetheless succeed at detecting core violations involving that information. This explains, for example, why infants as young as 2.5–3.5 months detect a violation when an object is hidden behind a screen that then rotates through the space occupied by the object (Baillargeon, 1987), when an object is hidden in a container that is then slid forward and to the side to reveal the object standing in the container's initial position (Hespos & Baillargeon, 2001b), and when a cover is lowered over an object, slid to the side, and finally lifted to reveal no object (Wang et al., 2005). In each case, the categorical information represented (e.g., cover over object) is sufficient, when interpreted by the PR system's core knowledge, to flag the event as a violation. The other consequence is that even older infants will fail to detect core violations that involve featural information they have not yet identified as relevant for the event category involved. This explains, for example, why infants under 11.5 months fail to detect a violation when an object changes color behind a screen (Wilcox, 1999), when an asymmetrical object remains stable with its larger end unsupported (Baillargeon & DeJong, 2017), and when a tall object becomes fully hidden in a short tube (Wang et al., 2005). In each case, infants have not yet identified the relevant feature (e.g., height in tube events), and so no information about the feature is requested, even though that information is available in the OF system.[3]

Second, the two-system model explains why infants who succeed in detecting feature-based core violations in one event category may fail to do so in a different category. Because (a) features are identified separately for each event category and months can separate the identification of the same feature in different categories, and (b) the PR system taps the OF system for only those features it has identified as relevant for the event category involved, striking décalages (or lags) can arise in infants' detection of similar feature-based violations in different event categories (Hespos & Baillargeon, 2001a, 2006; Wang & Goldman, 2016; Wang & Kohne, 2007; Wang et al., 2005).

Finally, the two-system model explains why infants who succeed at detecting change or interaction violations involving particular features in an event category may still fail to detect individuation violations involving the same features and event category. Because the OF system uses primarily categorical information to individuate and track the objects in a physical event, whereas the PR system uses both categorical and featural information, the two systems can disagree about how many objects are involved in the event, leading infants to have no expectation at all about the number of objects present. This explains why infants aged 12 months and younger typically fail at individuation tasks involving two objects that (from infants' perspective) are identical except for already-identified

[2] The two-system model assumes that when two or more objects begin to interact, the OF system spontaneously passes on the categorical but not the featural information at its disposal to the PR system. Why would that be the case? One possible answer, in light of our discussion of the EBL mechanism, has to do with *learnability*. Infants begin with sparse event representations and add new features one by one as they identify them as causally relevant; having all of the featural information in the OF system's files added to an event representation at once would make it difficult or even impossible for infants to identify new features (Baillargeon & DeJong, 2017). Another, complementary answer is suggested by the fact that the information the OF system spontaneously passes on at the start of an event (arrow 1 in Fig. 1) is the same information it uses to track objects from event to event (arrow 3 from the OF systems in Fig. 1). Using primarily identity and spatiotemporal categorical descriptors to track objects may be effective under a wide range of conditions and may be particularly helpful for an infant system with limited information-processing resources.

[3] Evidence that information about a feature may be available in the OF system and yet not be requested by the PR system comes from experiments that compared infants' ability to detect changes to objects in *object-recognition* tasks that exclusively tapped the OF system and in *physical-reasoning* tasks that also tapped the PR system (Wang & Goldman, 2016; Wang & Mitroff, 2009). In one experiment (Wang & Goldman, 2016), for example, 12-month-olds first saw a tall tube being lowered over a short block. Next, the tube was lifted to reveal either the same block as before (no-change event) or a much taller block (change event). Consistent with prior findings that height is not identified as a tube feature until about 14 months (Wang et al., 2005), infants failed to detect the change to the block's height. However, they detected this change if they were briefly *turned away* from the apparatus while the tube was lowered over the block and lifted back again. In line with the two-system model, when infants *witnessed no causal interaction* between the tube and the block, the OF system guided their responses in what was for them essentially an object-recognition task; it detected the change to the block's height and updated the featural information in its file, leading infants to show enhanced attention to the novel, taller block (i.e., infants showed a novelty response). However, when infants *witnessed the causal interaction* between the tube and the block, the PR system took over and guided their responses in this physical-reasoning task; because it had not yet identified height as a tube feature, it did not tap the OF system for this information, causing infants to fail to detect the persistence violation they were shown (for further evidence that the PR system takes control during causal interactions, see Stavans et al., 2019; Wilcox & Chapa, 2002).

features (e.g., a striped cup and a dotted cup). After the two objects are brought out in alternation from behind a screen, infants detect no violation if the screen is removed to reveal only one of the objects (Xu et al., 2004), or if it is removed to reveal no object at all (Stavans et al., 2019). Similarly, after the two objects are brought out in alternation from inside a box, infants detect no violation if the box is shown to contain only one of the objects (Van de Walle et al., 2000), or if it remains silent when shaken, as though empty (Stavans et al., 2019).

1.3. New prediction

Our research was designed to test a new prediction suggested by two assumptions of the two-system model. One of these assumptions is that although the OF system spontaneously passes on only categorical information (i.e., identity and spatiotemporal descriptors) to the PR system, it does pass on featural information whenever the PR system requests it. Given that the OF system is capable of passing on both categorical and featural information, this gave rise to an interesting possibility: What if the OF system could be primed to pass on featural as well as categorical information to the PR system at the beginning of an event? More specifically, *what if the OF system could be primed to pass on information about an as-yet-unidentified feature relevant to the event*? This information would then become part of the PR system's event representation.

Of course, receiving information about an as-yet-unidentified feature would be of little use to the PR system if it did not know how to interpret it. This brings us to the second assumption, which we discussed when describing the EBL mechanism: When it comes to the simple event categories we have been discussing, infants learn primarily what featural information to include in their event representations, not how to interpret that information; their core knowledge enables them to correctly reason about the information as soon as it becomes available. This meant that if the OF system could be primed to pass on information about an as-yet-unidentified feature, *the PR system would be able to interpret this information using its domain knowledge, making possible successful expectations and actions involving the information.*

Across experiments, we used two different types of manipulations to prime the OF system to pass on information about an as-yet-unidentified feature to the PR system. One type of manipulation (used with infants aged 7–12 months) involved exposure to a *multi-value object array*. Infants saw a row of three or four objects that were identical except that each represented a different value of a single varying feature (e.g., dolls that each differed only in color, or blocks that each differed only in height). We speculated that if exposure to these objects was sufficient to highlight the varying feature in the OF system, then this might cause the system, when subsequently faced with a test event involving one of the objects, to send information about the feature along with the usual categorical information. Once included in the PR system's event representation and interpreted by its domain knowledge, this featural information should allow infants to respond correctly to the event. To add to the generalizability of our results, we tested this possibility using three different features and event categories: color in containment events (Experiment 1), height in tube events (Experiments 3 and 5), and proportional distribution in support events (Experiment 6). In addition, Experiments 1 and 3 used VOE tasks with change violations, Experiment 5 used a preferential-reaching task, and Experiment 6 used a VOE task with an interaction violation.

The second type of manipulation (used with infants aged 12–13 months) involved providing a *novel label for an object*: Infants heard an experimenter utter a novel label for an object using a nominal naming phrase (e.g., "A gormet!"). Our manipulation built on two prior findings in the language-acquisition literature. One is that by the end of the first year, a novel label applied to an object serves as an invitation to form an object category (Balaban & Waxman, 1997; Dewar & Xu, 2007; Waxman & Markow, 1995; Xu et al., 2005). The other is that when infants first hear a novel label applied to an object, they richly encode the object's properties so that this information is available, when the label is subsequently applied to a new object, to find the commonalities between them and help determine the boundaries of the object category referred to by the label (Havy & Waxman, 2016; LaTourette & Waxman, 2020; Waxman & Braun, 2005). Our manipulation took advantage of this initial rich encoding. We speculated that when the experimenter labeled the object "A gormet!", English-speaking infants would be uncertain about the boundaries of this object category (e.g., were all gormets this shape, size, and color, or did they vary in these properties?). Consequently, when building a file for the object, the OF system might (a) list the label as one of the object's identity categorical descriptors and (b) bind some of the object's featural properties (e.g., shape, size, and color) to the label, to specify its possible meaning. This would cause the OF system, when subsequently faced with a test event involving the object, to pass on these featural properties along with the object's label. Once included in the PR system's event representation and interpreted by its domain knowledge, this featural information should allow infants to detect violations involving the information. We tested this possibility in two VOE tasks: One focused on color information in containment events and presented infants with a change violation (Experiment 2), and the other focused on height information in tube events and presented infants with an interaction violation (Experiment 4).

Although we used two types of priming manipulations across experiments, we assumed that, ultimately, both types would cause the OF system to send *a mix of categorical information* (arrow 1 in Fig. 1) *and featural information* (dashed arrow immediately below arrow 1 in Fig. 1) to the PR system when a test event started. In the object-array manipulations, this featural information would be about the as-yet-unidentified feature highlighted in the arrays and necessary for success at the task. In the novel-label manipulations, it would be about the featural properties bound to the label, which hopefully would include the as-yet-unidentified feature necessary for success at the task.

Many prior experiments have found that subtle contextual manipulations can lead infants aged 4–13 months to succeed at physical-reasoning tasks they would otherwise fail. These include *categorical-encoding* manipulations designed to induce infants' OF system to assign distinct categorical descriptors to two or more objects (Futó et al., 2010; Lin et al., 2019; Stavans & Baillargeon, 2018; Wilcox & Chapa, 2004; Wilcox et al., 2011; Xu, 2002); *teaching* experiments designed to teach infants' PR system a new physical rule that identifies a causally relevant feature for predicting outcomes in an event category (Baillargeon & DeJong, 2017; Wang & Baillargeon, 2008a; Wang & Kohne, 2007); and *carryover* experiments in which infants see the same objects in two successive events from different categories, causing the PR system to carry over the featural information included in its representation of the first event to its representation of the second event (Baillargeon et al., 2009; Wang & Baillargeon, 2005; Wang & Onishi, 2017; Wang, 2011). In light of these prior experiments (described in more detail in the General Discussion), the potential contributions of the present research were twofold. First, it sought to introduce two novel priming manipulations: In six experiments, we tested whether infants aged 7–13 months ($N = 304$) might serendipitously succeed at a physical-reasoning task involving an as-yet-unidentified feature if their OF system was primed, via exposure to an object array or a novel label, to pass on information about the feature to the PR system (for a summary of all experiments, see Table 1). Second, positive findings across experiments would provide strong evidence for the two-system model and for its description of how the OF and PR systems represent and exchange information about objects in physical events.

2. Experiment 1

Experiment 1 examined whether 12-month-old infants would detect that an object surreptitiously changed color when hidden in a container if their OF system was first primed using an object array to pass on information about the object's color to the PR system.

We first tested a *baseline* group of 12-month-olds, to compare their ability to detect color change violations in occlusion and containment events. Prior research indicated that by their first birthday, most infants have identified color as an occlusion feature (Wilcox, 1999); however, no information was available about the age at which infants typically identify color as a containment feature. Infants saw test events in which a brightly colored Boohbah toy (a non-human doll with a smooth head and a furry body; Fig. 2) was lowered either behind an occluder (*occluder* condition) or inside a container (*container* condition). When lifted again, the toy was either the same as before (*expected* event) or a different color (*unexpected* event). To anticipate our results, we found that infants in the occluder condition looked significantly longer at the unexpected than at the expected event, confirming that they had identified color as an occlusion feature and could detect the change violation in the unexpected event. In contrast, infants in the container condition looked equally at the events, suggesting that they had not yet identified color as a containment feature and thus failed to detect the violation in the unexpected event.

Building on this last result, we next tested a *priming* group of 12-month-olds, to examine whether they would succeed in the container condition if color information was first highlighted in the OF system via a priming trial. Infants saw a row of Boohbah toys that differed only in color and thus represented distinct values of this single varying feature. We speculated that presenting four different colors might be sufficient to render color information salient, but presenting only two colors might not be; in various areas of research on infant cognition, three or more exemplars are often required to draw infants' attention to particular information (Gerken & Bollt, 2008; Henderson & Woodward, 2012; Needham et al., 2005; Quinn & Bhatt, 2005; Xu & Denison, 2009; Xu & Tenenbaum, 2007). In line with these speculations, infants in an *effective-priming* condition received a priming trial showing toys in four colors, two of which were used again in the test trials. In contrast, infants in an *ineffective-priming* condition saw toys in only two colors, which were both used again in the test trials.

We made two predictions. First, if exposure to Boohbah toys in four different colors in the priming trial was sufficient to highlight the toys' colors in the OF system, then this might cause the OF system, in the test trials, to send the PR system information about the toys' colors along with their categorical descriptors. This color information would then be included in the PR system's event representations, along with the width, height, and shape information requested by the system (recall that by 12 months these features have all been identified as causally relevant for containment events; Hespos & Baillargeon, 2001a; Wang & Onishi, 2017; Wang et al., 2004). Consequently, infants would be able to reason that (a) only one Boohbah toy could fit into the container and (b) the toy could not magically change color when hidden. Infants should therefore detect the change violation they were shown and look significantly longer at the unexpected than at the expected event. Second, if exposure to Boohbah toys in only two different colors was insufficient to highlight the toys' colors in the OF system, then no color information would be sent to the PR system. Infants should therefore fail to detect the violation in the unexpected event and look equally at the two events. This negative result would help rule out alternative interpretations of positive results in the effective-priming condition, such as that exposure to Boohbah toys in the priming trial facilitated infants' processing of the test trials (e.g., by reducing overall processing demands).

2.1. Design

Infants in the *baseline* group were randomly assigned to an occluder or a container condition. Each infant faced a puppet-stage apparatus and saw an unexpected and an expected test event on alternate trials for two pairs of trials. Half of the infants in each condition saw the unexpected event first, and half saw the expected event first. At the beginning of each trial in the *occluder* condition, a Boohbah toy (e.g., purple) sat on the apparatus floor to the right of an occluder, an upright screen that was too small and too close to the back wall to be able to hide more than one Boohbah toy at a time. A primary experimenter's gloved hand rested on the ledge of a window (filled with a fringe curtain) in the right wall of the apparatus. To start, the hand grasped the toy, lifted it above the occluder, lowered it behind the occluder, and then released it and moved back above the occluder; after a short pause, the hand retrieved the toy and returned it to its original position on the apparatus floor. Each event cycle lasted 15 s, and cycles were repeated until the trial ended (see Procedure section for criteria). In the *expected* event, the toy was the same color as before when retrieved (e.g., purple). In the *unexpected* event, the toy was a new color (e.g., orange); as event cycles continued, the toy changed back and forth between its old and new color each time it went out of view (e.g., first purple, then orange, then purple, then orange, and so on). To effect these changes, a hidden secondary experimenter reached through an opening in the back wall behind the occluder and surreptitiously replaced one toy with the other (in the expected event, she simply removed and replaced the same toy, to equate procedures across

Table 1
Description and effect size estimate of each condition in Experiments 1–6. Each row represents a different condition, and columns (left to right) describe its experiment number, age tested, sample size, event category and feature studied, type of task used (VOE or action task), type of violation shown in the test events (for VOE tasks only), priming manipulation and condition (for priming conditions only), and effect size (Cohens' *d* for VOE tasks and Cohen's *g* for action tasks). Experiments 1, 2, and 5 used a within-subject design, and Experiments 3, 4, and 6 used a between-subjects design.

A summary of the six experiments reported in this paper.

Experiment	Age	*N*	Event Category	Feature	Task	Type of Violation	Priming Manipulation	Priming Condition	Effect Size (Cohen's *d* or *g*)
Expt. 1	12 months	16	occlusion	color	VOE	change	–	–	1.09
Expt. 1	12 months	16	containment	color	VOE	change	–	–	0.05
Expt. 1	12 months	16	containment	color	VOE	change	object array	effective	0.85
Expt. 1	12 months	16	containment	color	VOE	change	object array	ineffective	−0.10
Expt. 2	12 months	16	containment	color	VOE	change	verbal label	effective	1.10
Expt. 2	12 months	16	containment	color	VOE	change	verbal label	ineffective	0.16
Expt. 3	8 months	16	containment	height	VOE	change	–	–	2.40
Expt. 3	8 months	16	tube	height	VOE	change	–	–	−0.17
Expt. 3	8 months	16	tube	height	VOE	change	object array	effective	2.13
Expt. 3	8 months	16	tube	height	VOE	change	object array	ineffective	−0.44
Expt. 4	13 months	16	tube	height	VOE	interaction	verbal label	effective	1.22
Expt. 4	13 months	16	tube	height	VOE	interaction	verbal label	ineffective	−0.51
Expt. 5	10 months	16	containment	height	action	–	–	–	*0.25
Expt. 5	10 months	16	tube	height	action	–	–	–	*0.06
Expt. 5	10 months	16	tube	height	action	–	object array	effective	*0.25
Expt. 5	10 months	16	tube	height	action	–	object array	ineffective	*-0.06
Expt. 5	10 months	16	tube	height	action	–	object array	irrelevant	*-0.19
Expt. 6	7 months	16	support	proportional distribution	VOE	interaction	object array	effective	1.37
Expt. 6	7 months	16	support	proportional distribution	VOE	interaction	object array	ineffective	−0.47

events). Infants were assigned to one of four color sets: purple/orange, orange/purple, yellow/pink, or pink/yellow. Infants in the *container* condition watched similar events except that the occluder was replaced with a container; the front of the container was identical to the occluder and was similarly positioned in the apparatus, so that events in the two conditions were perceptually similar. To make possible the surreptitious removal and replacement of the toy in each event cycle, the container had an opening (not visible to infants) in the lower portion of its back wall.

Infants in the *priming* group were randomly assigned to an effective-priming or an ineffective-priming condition and saw the same test events as in the container condition of the baseline group. Prior to these events, infants received a priming trial that differed between conditions. In the *effective-priming* condition, infants saw a row of four different-color Boohbah toys; from left to right, these were purple, yellow, pink, and orange. The row was centered in front of the container, which hid the opening in the back wall; the gloved hand was not present during the trial. In the *ineffective-priming* condition, infants saw Boohbah toys in the two colors appropriate for their set. Half of the infants saw one exemplar of each toy (e.g., purple and orange), and half saw two exemplars of each toy (e.g., purple, purple, orange, and orange). This manipulation served to rule out the possibility that it was seeing four toys, as opposed to four colors, that made possible infants' success in the effective-priming condition.

2.2. Method

2.2.1. Participants

Participants were 64 healthy term infants (32 male, $M = 12$ months, 15 days, range = 11 months, 20 days to 13 months, 8 days).

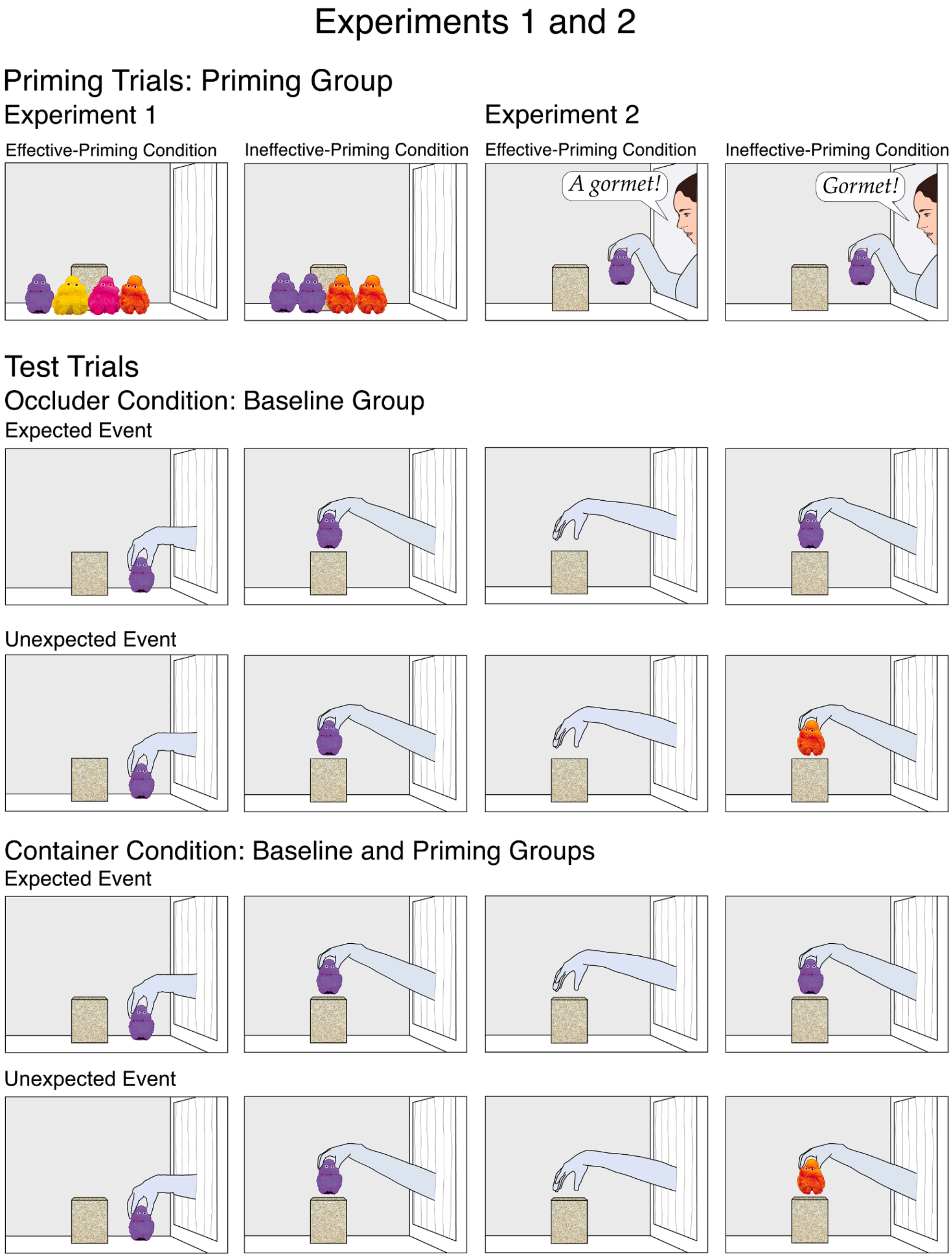

Fig. 2. Schematic depiction of the events in the priming and test trials of each condition in Experiments 1 and 2. (For information about the toys' colors in the trials, the reader is referred to the text or the web version of this article.)

Half of the infants were assigned to the baseline group (which was conducted first), and half to the priming group. Within each group, 16 infants were randomly assigned to each condition (occluder or container in the baseline group, effective- or ineffective-priming in the priming group). Another 19 infants were tested but excluded from the final analyses. Of these, 14 looked for the maximum time allowed (60 s) on three or more of the four test trials (4 in the occluder condition, 5 in the container condition, and 5 in the effective-priming condition).[4] The remaining 5 infants were excluded because they were active (e.g., trying to get down from their parent's lap; 3, 2 in the container condition and 1 in the ineffective-priming condition), or because the mean difference in their looking times at the two test events was over 2.5 standard deviations from the relevant condition means (2, 1 in the container condition and 1 in the effective-priming condition).

In this and the following experiments, infants' names were obtained from a university-maintained database of parents interested in participating in child development research. Each infant's parent gave written informed consent prior to the testing session, and the protocol was approved by the Institutional Review Board at the University of Illinois at Urbana-Champaign.

2.2.2. Apparatus and stimuli

The apparatus consisted of a brightly lit display box (201 cm high × 101 cm wide × 44 cm deep) with a large opening (55 × 95 cm) in its front wall; between trials, a hidden supervisor lowered a curtain to cover this opening. Inside the apparatus, the sidewalls were painted white, and the back wall and floor were covered with pastel adhesive paper. The primary experimenter sat at a window (51 × 38 cm), filled with a white fringe curtain, in the right wall of the apparatus; she wore a pale shirt and introduced her right hand, encased in a long pale glove, into the apparatus through the curtain.

Across conditions, eight Boohbah toys (each about 12 × 10 cm in diameter at the largest points) were used. They were identical except for color, with two toys in each of four colors: purple, orange, yellow, and pink. In the occluder condition, the occluder consisted of a screen (16 × 13 cm) covered with tan granite-patterned adhesive paper; it stood 14 cm from the back wall and 28 cm from the right wall. At the start of each test trial, the Boohbah toy sat about 9 cm to the right of and 1.5 cm behind the occluder. The secondary experimenter used an opening (11.5 × 11.5 cm) in the back wall, hidden behind the occluder, to surreptitiously remove and replace toys during the test trials. In the container, effective-priming, and ineffective-priming conditions, the screen was replaced with a tan granite-patterned container (16 × 13 × 13 cm) whose front wall looked identical to the occluder and occupied the same position on the apparatus floor. The container had an opening (12 × 12.5 cm) in its back wall, 4 cm below its rim; this opening was aligned with that in the back wall, to allow the secondary experimenter to remove and replace toys.

During each testing session, a metronome beat softly to help the two experimenters adhere to the events' second-by-second scripts. Two cameras captured images of the infant and events; the two images were combined, projected onto a monitor located behind the apparatus, and checked by the supervisor to confirm that the events followed the prescribed scripts.

2.2.3. Procedure

Each infant sat on a parent's lap centered in front of the apparatus. Parents were instructed to keep their infants seated (standing infants might be able to see behind the occluder or inside the container), to remain silent, and to close their eyes during the test trials. Prior to the testing session, the primary experimenter showed the infant her gloved hand as well as the occluder (tapping on both sides; occluder condition) or a copy of the container with no opening in its back wall (tapping on all sides including the insides; all other conditions); this manipulation ensured that infants knew they were facing a screen or a container in the session. During the trials, two observers hidden behind cloth-covered panels on either side of the apparatus monitored the infant's looking behavior; they were not told and could not determine the infant's test order and condition (occluder or container in the baseline group, effective- or ineffective-priming in the priming group).

In both groups, each test trial began with a paused pretrial that ended when the infant had accumulated 2 s of looking, to allow the infant to orient to the apparatus before the trial proper began. Each trial ended when the infant either (a) looked away from the event for 2 consecutive seconds after looking at it for at least 15 cumulative seconds or (b) looked for 60 cumulative seconds. The 15-s minimum value corresponded to one event cycle and helped ensure that infants had sufficient information to distinguish the unexpected and expected events. For the priming group, the priming trial that preceded the test trials ended when the infant either (a) looked away from the array for 2 consecutive seconds after looking at it for at least 5 cumulative seconds or (b) looked for 30 cumulative seconds. The 5-s minimum value helped ensure that infants had the opportunity to inspect and compare the Boohbah toys. One infant, in the ineffective-priming condition, completed only one test pair because he was overly active and climbed into the apparatus. Interobserver agreement during each test trial was calculated by dividing the number of 100-ms intervals in which the two observers agreed by the total number of intervals in the trial. Agreement was measured for 60/64 infants (only one observer was present for four infants) and averaged 97% per trial per infant.

Preliminary analyses of the test data in the ineffective-priming condition revealed no significant interaction of event with the number of exemplars shown in the priming trial (e.g., purple and orange vs. purple, purple, orange, and orange), $p = .789$; the data

[4] Infants who look for the maximum amount of time allowed in all or most test trials in a task (i.e., infants who "reach ceiling") are often excluded on the assumption that they needed additional familiarization to successfully process the test events they were shown (for other reports with excluded ceiling babies, see e.g., Baillargeon & DeJong, 2017; Jin et al., 2018; Margoni et al., 2018; Scott & Baillargeon, 2009). Of course, these ceiling infants might have performed better had they been provided with familiarization trials, for a better introduction to the task. However, familiarization trials also increase the risk of inadvertently inducing subtle novelty or familiarity preferences that can then affect test responses (Wang et al., 2004), so we typically err in the direction of using as few familiarization trials as possible.

were therefore collapsed across this factor in subsequent analyses. In addition, preliminary analyses of the test data in each group revealed no significant interaction of condition and event with infants' sex, color set, or test order, all *ps* > 0.182; the data were therefore collapsed across these latter three factors. All of the data from Experiments 1–6 are available via the Open Science Framework and can be accessed at https://osf.io/aewyn/.

2.3. Results

2.3.1. Baseline group

Looking times in the test trials (Fig. 3) were averaged across the two test pairs and analyzed using a 2 × 2 ANOVA with condition (occluder, container) as a between-subjects factor and event (unexpected, expected) as a within-subject factor. The analysis yielded a significant main effect of event, $F(1, 30) = 7.74$, $p = .009$, as well as a significant Condition × Event interaction, $F(1, 30) = 6.57$, $p = .016$, $\eta_p^2 = 0.180$. Planned comparisons indicated that infants in the occluder condition looked significantly longer at the unexpected ($M = 48.37$, $SD = 10.30$) than at the expected ($M = 36.47$, $SD = 11.56$) event, $F(1, 30) = 14.28$, $p < .001$, Cohen's $d = 1.09$, with 13/16 infants showing this pattern; in contrast, infants in the container condition looked about equally at the unexpected ($M = 38.58$, $SD = 9.14$) and expected ($M = 38.09$, $SD = 10.05$) events, $F(1, 30) = 0.02$, $p = .888$, $d = 0.05$, with 11/16 infants looking longer at the unexpected event. Nonparametric Wilcoxon signed-rank tests confirmed the results of the occluder ($Z = 2.60$, $p = .009$) and container ($Z = 0.38$, $p = .706$) conditions.

2.3.2. Priming group

Looking times in the priming trial were compared using a one-way ANOVA with condition (effective-priming, ineffective-priming) as a between-subjects factor. This factor was not significant, $F(1, 30) = 0.08$, $p = .782$, suggesting that infants in the effective-priming ($M = 19.04$, $SD = 9.11$) and ineffective-priming ($M = 18.15$, $SD = 8.89$) conditions looked about equally at the static arrays of Boohbah toys they were shown.

Looking times during the test trials (Fig. 3) were averaged and analyzed using a 2 × 2 ANOVA with condition (effective-priming, ineffective-priming) as a between-subjects factor and event (unexpected, expected) as a within-subject factor. The analysis yielded only a significant Condition × Event interaction, $F(1, 30) = 5.18$, $p = .030$, $\eta_p^2 = 0.147$. Planned comparisons indicated that infants in the effective-priming condition looked significantly longer at the unexpected ($M = 44.41$, $SD = 11.20$) than at the expected ($M = 34.89$, $SD = 11.13$) event, $F(1, 30) = 8.34$, $p = .007$, $d = 0.85$, with 12/16 infants showing this pattern; in contrast, infants in the ineffective-priming condition looked about equally at the unexpected ($M = 43.62$, $SD = 12.44$) and expected ($M = 44.71$, $SD = 8.47$) events, $F(1, 30) = 0.11$, $p = .742$, $d = -0.10$, with 7/16 infants looking longer at the unexpected event. Wilcoxon signed-rank tests confirmed the results of the effective-priming ($Z = 2.32$, $p = .020$) and ineffective-priming ($Z = 0.09$, $p = .925$) conditions.

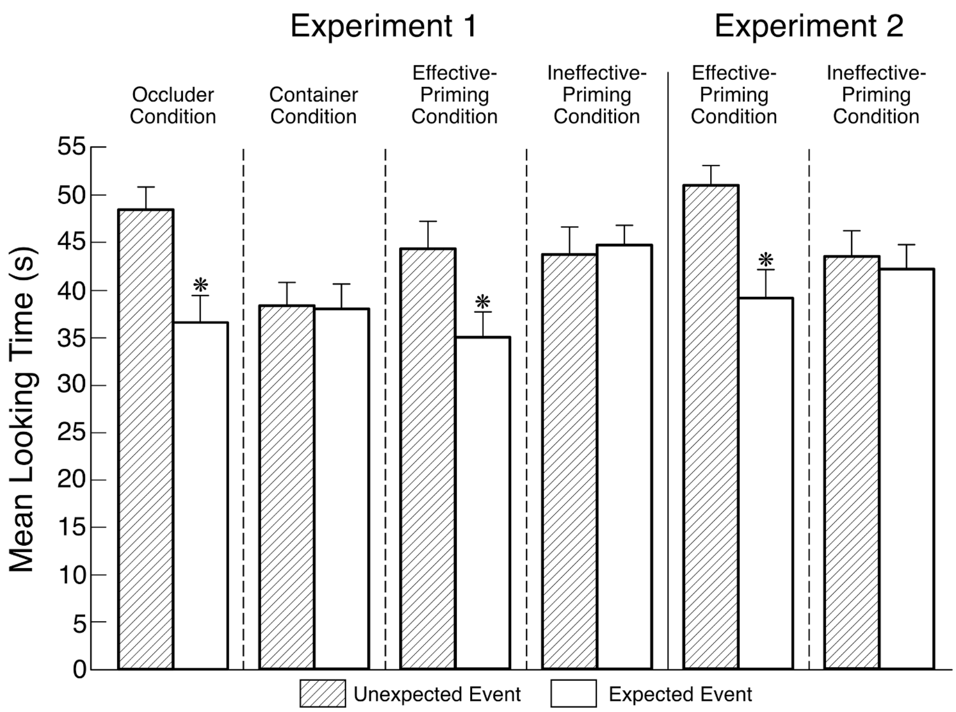

Fig. 3. Mean looking times in the test trials of Experiments 1 and 2, by condition and event. Error bars represent standard errors, and an asterisk denotes a significant difference between events within a condition.

2.4. Discussion

In the baseline group, infants in the occluder condition looked significantly longer at the unexpected than at the expected event, confirming that by their first birthday, infants have identified color as an occlusion feature (Wilcox, 1999). As the PR system represented each test event, it categorized it as an occlusion event, tapped the OF system for information about the toy's color (and other identified occlusion features such as width, height, and shape; Hespos & Baillargeon, 2001a; Wang & Onishi, 2017; Wang et al., 2004), interpreted this information using its domain knowledge, and hence detected the persistence violation in the unexpected event: The toy appeared to magically change color each time it was occluded. In contrast, infants in the container condition tended to look equally at the two test events, suggesting that they had not yet identified color as a containment feature. When representing each test event, the PR system requested information about identified features such as width, height, and shape, but not color, and it therefore failed to detect the persistence violation in the unexpected event. From the PR system's perspective, in each test event a toy of the same width, height, and shape was alternately lowered into and lifted from the container; there was no information in the unexpected event's representation to suggest that anything was amiss.

In the priming group, infants in the effective-priming condition succeeded in detecting the persistence violation from the container condition after receiving a priming trial depicting four different-color Boohbah toys. This positive finding suggests that (a) the multi-value array shown in the priming trial highlighted information about the toys' colors in the OF system; (b) this caused the OF system to send information about the toy's color to the PR system in each test event; (c) this color information, once included in the PR system's event representation and interpreted by its domain knowledge, allowed infants to detect the persistence violation in the unexpected event: An object cannot spontaneously change color when briefly hidden in a container. In contrast, infants in the ineffective-priming condition looked equally at the unexpected and expected events, confirming the results of the container condition in the baseline group and suggesting that exposure to only two different-color toys (even if there were two exemplars of each toy) was not sufficient to highlight information about the toys' colors in the OF system. This negative result also rules out the possibility that infants in the effective-priming condition succeeded simply because prior exposure to the two Boohbah toys used in the test events facilitated the processing of the events.

3. Experiment 2

In Experiment 1, we found that after seeing an array of four different-color Boohbah toys, 12-month-olds detected a violation if one of the toys surreptitiously changed color when briefly hidden in a container. We suggested that exposure to the multi-value array highlighted information about the toys' colors in the OF system, causing it to pass on this information to the PR system. In Experiment 2, we asked whether a *verbal* priming manipulation that did not focus on color information directly might also allow infants to detect the same violation.

As in the priming group of Experiment 1, infants were assigned to an effective- or an ineffective-priming condition. Infants in the *effective-priming* condition received three priming trials in which a Boohbah toy (the one they would see at the start of each test event; e.g., the purple toy) was given a novel label using the nominal naming phrase, "A gormet!" (Fig. 2). As was explained in the Introduction, we speculated that upon hearing "A gormet!", English-speaking infants would be uncertain about the boundaries of this object category (e.g., were all gormets this size, shape, and color, or did they vary in these properties?). Consequently, when building the toy's file, the OF system might bind some of the toy's featural properties to its label (one of its identity categorical descriptors), to specify its possible meaning. If information about color was included among these properties, then this information (a) would be passed on to the PR system at the start of each test event, (b) would become part of the event's representation, and (c) would be interpreted by the PR system's domain knowledge, allowing infants to detect the persistence violation in the unexpected event: An object cannot spontaneously change color when briefly hidden in a container.

Priming trials in the *ineffective-priming* condition were identical except that the experimenter simply said, "Gormet!" Without an article such as "a" or "the" before it, the novel word "gormet" becomes ambiguous: It is not clear whether it is intended as a label for the toy, as an attention-getter (e.g., "Look!", "Watch!", "Wow!"), or as a reference to some aspect of the experimenter's action (e.g., "Tilt!", "Right!", "Fun!"). Prior evidence indicated that by 13–15 months of age, English-speaking infants are already sensitive to this subtle linguistic difference: They are more likely to interpret a novel word as a label when it is preceded by "a" or "the" than when it is not (Booth & Waxman, 2009; Fennell & Waxman, 2010; Waxman & Booth, 2001). Assuming the same might be true of 12-month-olds, we took advantage of this sensitivity to create an ineffective-priming condition that was highly similar to the effective-priming condition but, because it did not unambiguously provide a label for the Boohbah toy, was unlikely to induce infants to detect the violation in the unexpected event.

In short, as in Experiment 1, infants in the effective-priming condition were expected to detect the persistence violation they were shown and look significantly longer at the unexpected than at the expected event, whereas infants in the ineffective-priming condition were expected to fail to detect this violation and look equally at the two events. We reasoned that negative findings in the ineffective-priming condition would help rule out alternative interpretations of positive findings in the effective-priming condition, such as that prior exposure to the Boohbah toy in the priming trials facilitated infants' processing of the test trials, or that hearing speech in the priming trials enhanced infants' encoding of the toy's featural properties.

3.1. Design

Infants received four test trials identical to those in the container condition of Experiment 1 except that two new color sets were used, purple/yellow and yellow/purple. Prior to the test trials, infants received three priming trials that differed between conditions. Each priming trial had an initial phase and a final phase; the initial phase was computer-controlled and presented the experimenter's scripted actions, ending with a paused scene; the final phase was infant-controlled and showed this paused scene. Looking times during the two phases were computed separately. At the start of the (14-s) initial phase in the *effective-priming* condition, a Boohbah toy (e.g., purple) sat on the apparatus floor to the right of the container, and the experimenter knelt at her open window, with her gloved hand on the ledge. The experimenter lifted the toy, tilted it left and right (from infants' perspective), and said, "A gormet!" each time she tilted it to the right. After labeling the toy five times, the experimenter returned the toy to the apparatus floor and paused at her window; infants watched this paused scene until the trial ended. The use of separate initial and final phases in each trial thus ensured that infants heard all five repetitions of the label. Priming trials in the *ineffective-priming* condition were identical except that the experimenter simply said "Gormet!"

3.2. Method

3.2.1. Participants

Participants were 32 healthy term infants (15 male, $M = 12$ months, 18 days, range = 11 months, 29 days to 13 months, 4 days). Half of the infants were randomly assigned to the effective-priming condition and half to the ineffective-priming condition. Another 5 infants were excluded because they were fussy (2, 1 in each condition) or distracted (1, in the ineffective-priming condition), looked for the maximum time allowed on all four test trials (1, in the ineffective-priming condition), or stood during the surreptitious removal of the toy from the container in one of the test trials and hence could have noticed the removal (1, in the ineffective-priming condition).

3.2.2. Apparatus, stimuli, and procedure

The apparatus and stimuli were the same as in the containment condition of Experiment 1 with two exceptions. First, the fringe curtain in the primary experimenter's window was removed for the priming trials so that she could be seen, in her pale shirt, when she labeled the toy; white ceiling-to-floor curtains behind her hid the testing room. Second, the color sets used were purple/yellow and yellow/purple.

The procedure was also similar to that in the containment condition of Experiment 1 except that infants now received three priming trials, each with an initial and a final phase. Infants were highly attentive during these initial phases and looked, on average, for 96% of each phase. The final phase of each priming trial ended when the infant either (a) looked away from the paused scene for 2 consecutive seconds after looking at it for at least 2 cumulative seconds or (b) looked for 20 cumulative seconds. Criteria for ending the test trials were the same as in Experiment 1. Interobserver agreement during the test trials was calculated for 27/32 infants (only one observer was present for five infants) and averaged 96% per trial per infant. Finally, preliminary analyses of the test data revealed no significant interaction of condition and event with infants' sex, color set, or test order, all $ps > 0.880$; the data were therefore collapsed across these latter three factors.

3.3. Results

Looking times during the final phases of the priming trials were averaged and analyzed using a one-way ANOVA with condition (effective-priming, ineffective-priming) as a between-subjects factor. This factor was not significant, $F(1, 30) = 2.49$, $p = .125$, suggesting that infants in the effective-priming ($M = 14.89$, $SD = 3.97$) and ineffective-priming ($M = 12.46$, $SD = 4.72$) conditions looked about equally at the (identical) paused scene at the end of each trial

Looking times during the test trials (Fig. 3) were averaged and analyzed using a 2 × 2 ANOVA with condition (effective-priming, ineffective-priming) as a between-subjects factor and event (unexpected, expected) as a within-subject factor. The analysis yielded a significant main effect of event, $F(1, 30) = 9.68$, $p = .004$, as well as a significant Condition × Event interaction, $F(1, 30) = 5.62$, $p = .024$, $\eta_p^2 = 0.158$. Planned comparisons indicated that infants in the effective-priming condition looked significantly longer at the unexpected event ($M = 50.81$, $SD = 8.69$) than at the expected event ($M = 39.03$, $SD = 12.39$), $F(1, 30) = 15.04$, $p < .001$, $d = 1.10$, with 13/16 infants showing this pattern; in contrast, infants in the ineffective-priming condition looked about equally at the unexpected ($M = 43.63$, $SD = 10.30$) and expected ($M = 42.04$, $SD = 10.27$) events, $F(1, 30) = 0.27$, $p = .607$, $d = 0.16$, with only 9/16 infants looking longer at the unexpected event. Wilcoxon signed-rank tests confirmed the results of the effective-priming ($Z = 2.82$, $p = .005$) and ineffective-priming ($Z = 0.36$, $p = .720$) conditions.

3.4. Discussion

Infants who heard "A gormet!" as the Boohbah toy was tilted back and forth in the priming trials looked significantly longer at the unexpected than at the expected event in the test trials, whereas infants who simply heard "Gormet!" looked equally at the events. These results suggest that when the word gormet was presented in a nominal naming phrase, infants interpreted it as an object label. Because the boundaries of the object category referred to by this label were unclear, the OF system bound to the label some of the featural properties of the toy, including its color. Consequently, when the OF system passed on the toy's label (along with other categorical descriptors) to the PR system at the start of each test event, information about the toy's color was fortuitously passed on as well. Once included in the PR system's event representation, this color information enabled infants to detect the persistence violation in the unexpected event. In contrast, infants who simply heard "Gormet!" did not interpret the word as an object label. As a result, no color information was passed on to the PR system, who therefore failed to detect the violation in the unexpected event.

4. Experiment 3

In Experiment 1, 12-month-olds succeeded in detecting a *color* change violation in a *containment* event following exposure to a multi-value object array designed to highlight color information in the OF system. To provide converging evidence for this finding, in Experiment 3 we used multi-value object arrays to induce 8-month-olds to detect a *height* change violation in a *tube* event (Fig. 4). Infants saw a tall block being lowered into a tall tube; when lifted again, the block was either the same height as before (*expected* event) or much shorter (*unexpected* event). Such a violation is typically not detected until about 14 months of age (Wang et al., 2005).

Beyond using a violation involving a different feature (height) and a different event category (tube events), Experiment 3 differed from Experiments 1 and 2 in one other respect: Infants saw a single event in a single test trial. This single-event design allowed us to garner evidence that once primed to attend to a feature they have not yet identified as causally relevant to an event category, infants

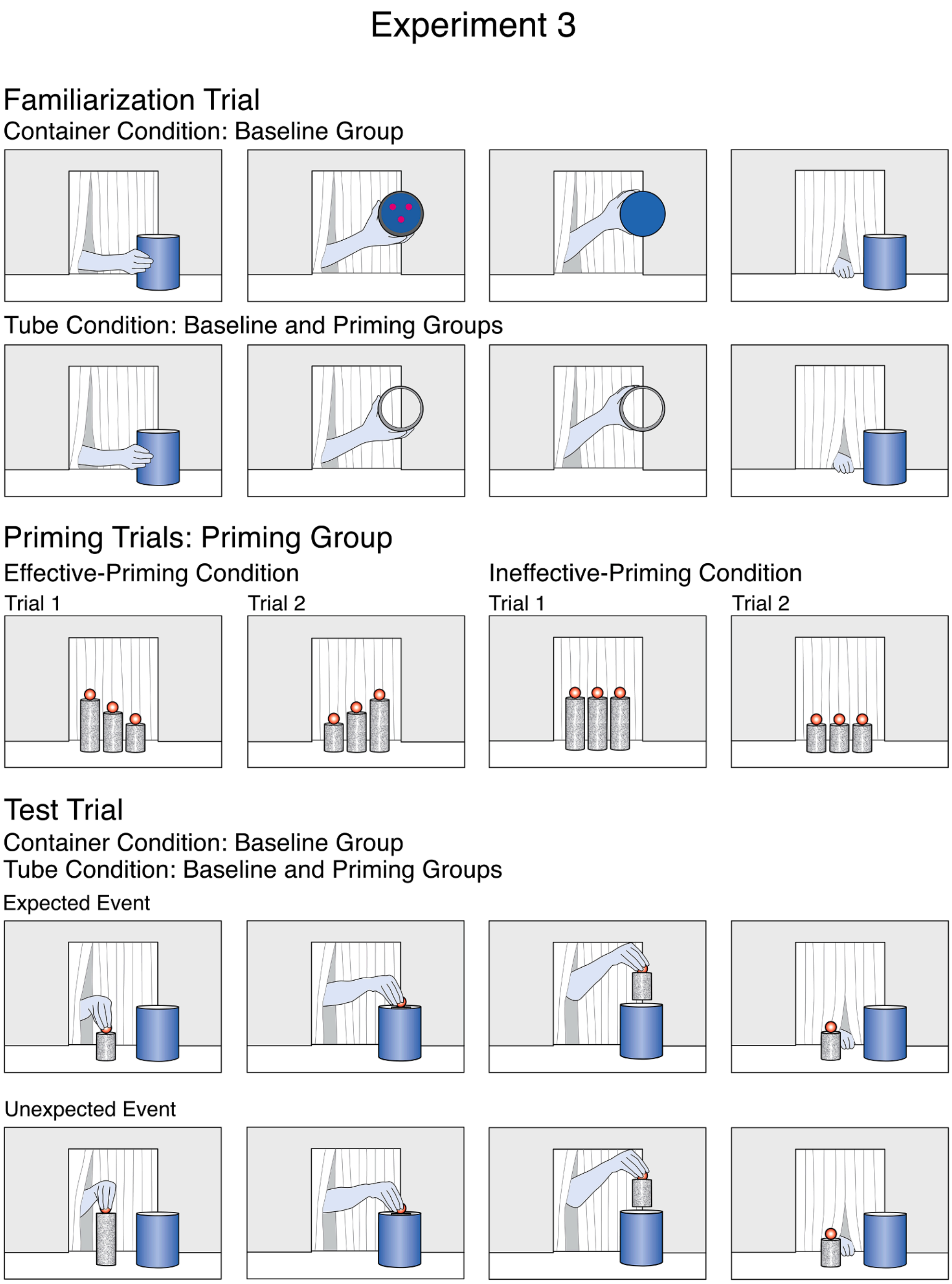

Fig. 4. Schematic depiction of the events in the familiarization, priming, and test trials of each condition in Experiment 3.

can detect a violation involving this feature *the very first time* they see it. In Experiments 1 and 2, infants received four test trials and saw up to four event cycles per trial; this left open the possibility that after they were primed to attend to color, infants still needed to see multiple repetitions of each containment event to be able to detect the violation in the unexpected event. (Analysis of the first test trial in the combined effective-priming conditions of Experiments 1 and 2 was inconclusive, because 27/32 infants looked the maximum allowed, 60 s, suggesting that they enjoyed seeing multiple repetitions of the unexpected or expected event presented in the trial). By showing infants in Experiment 3 a single unexpected or expected event followed by a paused scene, we could determine whether infants who were primed to attend to height information would be able to detect a height change violation in a tube event straight away, without having to see multiple repetitions of the event.

4.1. Design

Experiment 3 tested 8-month-olds and built on prior findings that height is identified as a causally relevant feature at about 7.5 months for containment events (Hespos & Baillargeon, 2001a), but only at about 14 months for tube events (Wang et al., 2005). Infants were randomly assigned to either a baseline or a priming group. The *baseline* group served to confirm that infants could detect a height change violation in a containment event, but not a tube event (to anticipate, results were as expected). Infants were randomly assigned to either a *container* or a *tube* condition, and they received one familiarization trial followed by one test trial in which they saw either an unexpected or an expected event. Each trial had an initial phase and a final phase. Only the familiarization trial differed between conditions: It introduced a (15-cm) tall cylinder whose bottom was either closed (container condition) or open (tube condition). When upright on the apparatus floor, the two cylinders looked identical (in fact, the cylinder with the open bottom was used in all test events, to make possible the change violation). In the (15-s) initial phase of the *familiarization* trial, the cylinder stood on the apparatus floor, and a primary experimenter's gloved hand rested at a window (filled with a fringe curtain) centered in the back wall of the apparatus. The hand grasped the cylinder, rotated it forward to show its open top, rotated it backward to show its bottom, and then repeated these rotations a second time. Finally, the hand returned the cylinder to the apparatus floor and then paused at the window. During the final phase of the trial, infants watched this paused scene until the trial ended. The (14-s) initial phase of the *test* trial began like the familiarization trial except that a cylindrical block, covered with a grey granite-patterned adhesive paper and decorated with a round red knob at the top, now stood 4 cm to the left of the cylinder. In the *unexpected* event, the cylindrical portion of the block was 15-cm tall. To start, the hand grasped the block's knob, lifted the block, moved it above the cylinder, and lowered it until only the knob remained visible. The hand then twisted the block clockwise and counterclockwise twice in each direction; this allowed a secondary experimenter to reach into the cylinder through a hidden opening in the apparatus floor, grasp the (7.5-cm) bottom half of the block, and twist it off, under the cover of the primary experimenter's twisting actions. The primary experimenter's hand then lifted the now much shorter block, lowered it to the apparatus floor, and paused at the window. The *expected* event was identical except that the cylindrical portion of the block was only 7.5-cm tall (the secondary experimenter still reached inside the tube to equate procedures across events). The final scenes of the two events were thus perceptually identical.

Infants in the *priming* group were randomly assigned to an effective- or an ineffective-priming condition. They received the same familiarization and test trials as in the tube condition, with two priming trials inserted between them. In the *effective-priming* condition, infants saw a static array of three different-height blocks, each with an identical knob, centered in front of the window; the blocks were arranged monotonically (1 cm apart) from tallest to shortest in one trial and from shortest to tallest in the other trial, with order counterbalanced across infants. The tall and short blocks were identical to those shown at the start of the unexpected and expected events, respectively, and the medium block was intermediate between them. The primary experimenter's hand was not present during these trials. In the *ineffective-priming* condition, infants saw three tall blocks in one trial and three short blocks in the other trial, with order counterbalanced.

We made two predictions. First, if exposure to three different-height blocks in the two priming trials was sufficient to highlight height information in the OF system, then information about the block's height might be sent to the PR system at the start of the test event, along with the usual categorical information. Once included in the PR system's event representation and interpreted by its domain knowledge, this height information should allow infants to detect the change violation in the unexpected event: An object cannot spontaneously change height when briefly lowered into a tube. Infants should thus look significantly longer if shown the unexpected as opposed to the expected event. Second, if exposure to three identical tall blocks in one priming trial and three identical short blocks in another trial was not sufficient to render height information salient in the OF system (e.g., height varied across but not within sets, and only two different heights were shown across sets), then no height information should be included in the test events' representations, and infants should fail to detect the violation in the unexpected event. From the PR system's perspective, it should appear as though a block was simply lowered into a tube and lifted back again. Infants should thus look equally at the unexpected and expected events.

4.2. Method

4.2.1. Participants

Participants were 64 healthy term infants (31 male, $M = 7$ months, 28 days; range = 7 months, 16 days to 8 months, 14 days). Half of the infants were randomly assigned to the baseline group and half to the priming group. Within each group, 16 infants were randomly assigned to each condition (container or tube in the baseline group, effective- or ineffective-priming in the priming group); within each condition, 8 infants were randomly assigned to see each test event (unexpected or expected). An additional 5 infants were excluded because they were distracted (1, in the tube condition), active (1, in the effective-priming condition), fussy (1, in the effective-priming condition), or inattentive (1, in the ineffective-priming condition), or because their test looking time was over 2.5 standard deviations from the cell mean (1, in the container condition).

4.2.2. Apparatus and stimuli

The apparatus was similar to that in Experiment 1 except that the window in the right wall was filled with a solid curtain and there was now a window (41 × 35 cm, filled with a white fringe curtain) centered in the back wall. The primary experimenter sat behind the window and wore a pale shirt and a pale glove on her right hand.

The container (15 × 13 cm in diameter) was a cylinder whose exterior was covered with dark blue adhesive paper; its interior was covered with blue-and-white seashell-patterned adhesive paper, and its interior bottom surface was also decorated with three large pink dots, to help attract infants' attention. The tube was identical to the container without its bottom and was used for both test events in all conditions. Stimuli also included eight cylindrical blocks (6 cm in diameter); each was covered with grey granite-patterned adhesive paper and had a red wooden knob (3 cm in diameter) on its top. There were three tall blocks, whose cylindrical portion was 15 cm tall, three short blocks (7.5 cm), and 1 medium block (11.25 cm). The final block, which was used in the unexpected event, appeared tall (15 cm) but actually consisted of two pieces that could be twisted apart. The top piece included the knob and the top 7.5 cm of the cylinder; its bottom was recessed by 0.5 cm, leaving a thin outer edge that could fit tightly over the top of the bottom piece. The bottom piece was 8 cm tall; its top 0.5 cm was slightly narrowed so that it could be inserted into the top piece. When stacked, the top and bottom pieces formed a continuous cylinder, with only a faint dividing line between them. In addition, magnets were inset into the bottom of the top piece and the top of the bottom piece, to ensure that the tall block did not come apart when moved. When on the apparatus floor, the container/tube stood 28 cm from the right wall and 9 cm from the back wall. In the test trial, the tube rested over a hidden hole in the apparatus floor; the secondary experimenter used this hole to reach into the tube during the events. During the familiarization and priming trials, the hole was hidden by a fake floor (43 × 100 cm) covered with the same adhesive paper as the apparatus floor.

4.2.3. Procedure

The procedure was similar to that in Experiment 1 except as follows. Infants in both groups received one familiarization and one test trial, each with an initial and a final phase; infants in the priming group also received two priming trials presenting static arrays of blocks after the familiarization trial. Infants were highly attentive during the initial phases of the familiarization and test trials; across groups and conditions, they looked, on average, for 94% of each initial phase. The criteria used to end the priming trials were the same as in Experiment 1 (which also presented infants with static arrays). The final phase of the familiarization trial ended when infants either (a) looked away from the paused scene for 2 consecutive seconds after looking at it for at least 2 cumulative seconds or (b) looked for 20 cumulative seconds. The final phase of the test trial ended when infants either (a) looked away from the paused scene for 2 consecutive seconds after looking at it for at least 8 cumulative seconds or (b) looked for 25 cumulative seconds. As the test event was shown only once, followed by the paused scene, the longer minimum and maximum values (compared to the familiarization trial) allowed infants to continue processing the event. Interobserver agreement during the final phase of the test trial was measured for 61/64 infants (only one observer was present for three infants) and averaged 94% per infant.

Preliminary analyses of the test data in each group revealed no significant interaction of condition and event with infants' sex or priming trial order (priming group only), all $ps \geq 0.336$; the data were therefore collapsed across these latter two factors.

4.3. Results

4.3.1. Baseline group

Looking times during the final phase of the familiarization trial were analyzed by a one-way ANOVA with condition (container, tube) as a between-subjects factor. This factor was not significant, $F(1, 30) = 0.51$, $p = .480$, suggesting that infants looked about equally whether they were shown the container ($M = 15.17$, $SD = 5.75$) or the tube ($M = 13.77$, $SD = 5.31$).

Looking times during the final phase of the test trial (Fig. 5) were analyzed using a 2 × 2 ANOVA with condition (container, tube) and test event (unexpected, expected) as between-subjects factors. The analysis yielded a significant main effect of event, $F(1, 28) = 5.14$, $p = .031$, as well as a significant Condition × Event interaction, $F(1, 28) = 8.01$, $p = .009$, $\eta_p^2 = 0.223$ (no such interaction was found when responses in the final phase of the familiarization trial were analyzed in the same way, $F(1, 28) = 2.96$, $p = .097$). Planned comparisons indicated that in the container condition, infants looked significantly longer if shown the unexpected event ($M = 22.03$, $SD = 4.67$) as opposed to the expected event ($M = 11.85$, $SD = 3.75$), $F(1, 28) = 13.01$, $p = .001$, $d = 2.40$, whereas in the tube condition, infants looked equally at the unexpected ($M = 18.18$, $SD = 6.52$) and expected ($M = 19.30$, $SD = 7.01$) events, $F(1, 28) = 0.16$, $p = .692$, $d = -0.17$. Non-parametric Wilcoxon rank-sum tests confirmed the results of the container ($Z = 3.02$, $p = .003$) and tube ($Z = -0.27$, $p = .784$) conditions.

4.3.2. Priming group

Looking times during the final phase of the familiarization trial (all infants saw the tube) were analyzed using a one-way ANOVA with condition (effective-priming, ineffective-priming) as a between-subjects factor. This factor was not significant, $F(1, 30) = 0.04$, $p = .843$, suggesting that infants in the effective-priming ($M = 11.54$, $SD = 6.41$) and ineffective-priming ($M = 11.11$, $SD = 5.59$) conditions looked about equally at the tube. Looking times during the priming trials were averaged and analyzed in the same manner. The analysis again yielded no significant results, $F(1, 30) = 0.00$, $p = .991$, suggesting that infants in the effective-priming ($M = 12.62$, $SD = 4.73$) and ineffective-priming ($M = 12.60$, $SD = 4.92$) conditions looked about equally at the static arrays of blocks they were shown.

Looking times during the final phase of the test trial (Fig. 5) were analyzed using a 2 × 2 ANOVA with condition (effective-priming, ineffective-priming) and test event (unexpected, expected) as between-subjects factors. The analysis yielded a significant Condition × Event interaction, $F(1, 28) = 8.29$, $p = .008$, $\eta_p^2 = 0.229$ (no such interaction was found when responses in the final phase of the familiarization trial, $F(1, 28) = 0.30$, $p = .587$, or in the priming trials, $F(1, 28) = 0.04$, $p = .848$, were analyzed in the same manner). Planned comparisons indicated that in the effective-priming condition, infants looked significantly longer if shown the unexpected event ($M = 17.89$, $SD = 4.27$) as opposed to the expected event ($M = 10.68$, $SD = 2.17$), $F(1, 28) = 8.98$, $p = .006$, $d = 2.13$, whereas in the ineffective-priming condition, infants looked about equally at the unexpected ($M = 14.60$, $SD = 6.38$) and expected ($M = 17.19$, $SD = 5.39$) events , $F(1, 28) = 1.16$, $p = .291$, $d = -0.44$. Wilcoxon rank-sum tests confirmed the results of the effective-priming ($Z = 2.79$, $p = .005$) and ineffective-priming ($Z = -1.16$, $p = .246$) conditions.

4.4. Discussion

In the baseline group, infants in the container condition looked significantly longer if shown the unexpected as opposed to the expected event, whereas infants in the tube condition looked equally at the events. These differential results are particularly striking because infants in the two conditions actually saw the *same* events. Infants detected the violation in the unexpected event when they believed—based on the familiarization trial—that the cylinder on the apparatus floor was a container, and they failed when they believed it was a tube. These results thus confirm prior findings that by 8 months, infants have identified height as a containment feature, but not a tube feature (Hespos & Baillargeon, 2001a; Wang et al., 2005).

In the priming group, infants saw the same test events as in the tube condition of the baseline group. Infants in the effective-priming condition now succeeded in detecting the violation they were shown, suggesting that exposure to three different-height blocks in each priming trial was sufficient to highlight height information in the OF system. As a result, this information was passed on to the PR system in the test trial, became part of the event's representation, and was interpreted by the PR system's domain knowledge, allowing infants to detect the persistence violation in the unexpected event: An object cannot spontaneously change height when briefly lowered into a tube. Strikingly, infants detected this violation even though (a) they saw the unexpected event only once and (b) they were about 6 months younger than the age at which this feature is typically identified (14 months; Wang et al., 2005).

In contrast to infants in the effective-priming condition, those in the ineffective-priming condition looked equally at the two test events. This negative result indicates that exposure to the priming trials in this condition was not sufficient to render height information salient in the OF system, either because the blocks did not vary in height within each priming trial or because only two different heights were shown across trials. This negative result also helps rule out the possibility that infants in the effective-priming condition succeeded simply because exposure to the tall and short blocks prior to the test events facilitated the processing of the events.

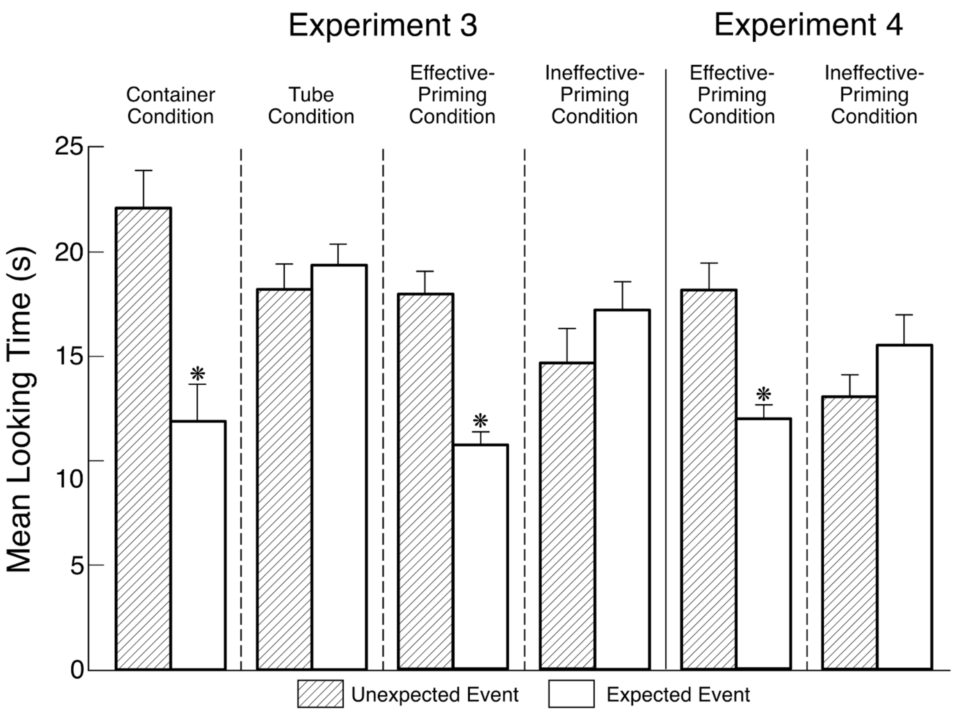

Fig. 5. Mean looking times in the final phase of the test trial in Experiments 3 and 4, by condition and event. Error bars represent standard errors, and an asterisk denotes a significant difference between events within a condition.

5. Experiment 4

Experiment 4 sought converging evidence for the results of Experiment 2 by using a verbal label to prime 13-month-olds to detect a persistence violation involving a feature they had not yet identified. As in Experiment 3, infants received a single test trial in which they saw an unexpected or an expected event involving height information in a tube event. To add to the generalizability of our findings, we showed infants a different type of persistence violation, an interaction rather than a change violation. In an interaction violation, objects interact in a way that is not physically possible given their respective featural properties. Here, using a design adapted from Wang et al. (2005), 13-month-olds saw a block being lowered into one of two tubes until it was fully hidden (infants were shown the tubes prior to the testing session). In the *expected* event, the tube was taller than the block, so this event was physically possible; in the *unexpected* event, the tube was shorter than the block, making it impossible for the block to become fully hidden inside it (Fig. 6).

Infants were randomly assigned to an effective- or an ineffective-priming condition. In the *effective-priming* condition, they received three priming trials in which the block was given a novel label using the nominal naming phrase, "A blicket!" Because the boundaries of this object category were unclear (e.g., were all blickets this size, shape, and color, or did they vary in these properties?), this might lead the OF system to bind some of the block's featural properties to its label (one of its identity categorical descriptors). If height information, in particular, was bound to the block's label, then it would be serendipitously passed on to the PR system at the start of the test event. As a result, it would be included in the event's representation and interpreted by the PR system's domain knowledge, allowing infants to detect the persistence violation in the unexpected event: An object cannot become fully hidden in a shorter tube. The *ineffective-priming* condition was identical except that infants heard the novel word alone, without the indefinite article "a" ("Blicket!"). If the word no longer unambiguously provided a label for the block, then height information was unlikely to be passed on to the PR system, and infants should thus fail to detect the violation in the unexpected event.

5.1. Design

Infants in each condition received three priming trials followed by a test trial. Each trial had an initial phase and a final phase. At the start of the (10-s) initial phase in each priming trial, a blue cylindrical block 16-cm tall and decorated with multicolored stars stood on the apparatus floor, centered in front of a window (filled with a white fringe curtain) in the back wall; the experimenter sat behind the window, with her gloved right hand on the ledge. The experimenter grasped the block, lifted it, tilted it left and right, and said either "A blicket!" (*effective-priming* condition) or "Blicket!" (*ineffective-priming* condition) each time she tilted it to the right. After labeling the block three times, the experimenter returned it to the apparatus floor and paused at her window; infants watched this paused scene until the trial ended. (The 12-month-olds in Experiment 2 heard the novel word five times per priming trial for a total of 15 times, but this may have been more repetitions than necessary: In an experiment by Waxman and Braun (2005), for example, 12–13-

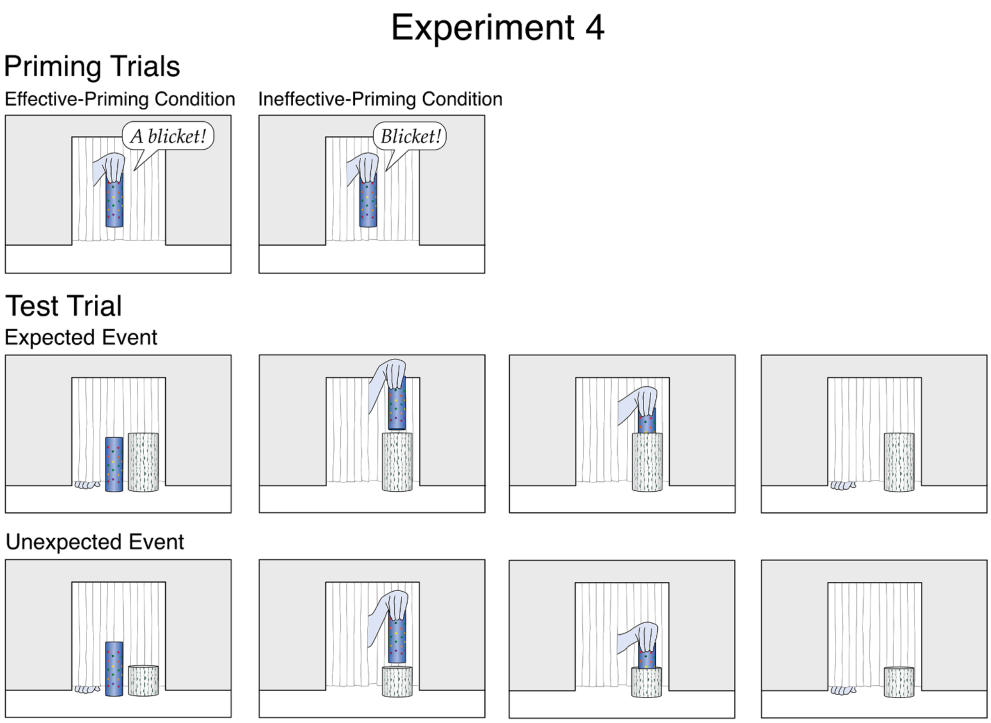

Fig. 6. Schematic depiction of the events in the priming and test trials of each condition in Experiment 4.

month-olds successfully formed a novel category with only six repetitions spread across three trials. Experiment 4 thus used fewer repetitions, three per trial for a total of nine in all.)

Infants in both conditions saw the same test events. At the beginning of the (16-s) initial phase in the *unexpected* event, a (9-cm) short white tube with green vertical stripes rested on the apparatus floor, 3 cm to the right of the block. The experimenter grasped the block, lifted it, moved it over the tube, lowered it about half-way, and then twisted it clockwise and counterclockwise five times in each direction (in Experiment 3, we used this twisting action mainly because it helped us produce our change violation; however, we realized over time that it might also have given the PR system time to process the information from the OF system, so we retained this twisting action in Experiment 4). Finally, the experimenter lowered the block inside the tube until it was fully hidden, and then she paused at her window. Infants watched this paused scene during the final phase of the trial. To produce this interaction violation, the short tube was placed over a hidden hole in the apparatus floor, and the experimenter lowered the block through the hole onto a platform under the floor. The *expected* event was identical except that the short tube was replaced by a (18-cm) tall container (recall that infants were shown a short and a tall tube prior to the testing session, so they believed they were facing a tube). The experimenter simply lowered the block to the bottom of the container.

5.2. Method

5.2.1. Participants

Participants were 32 healthy term infants (17 male, $M = 13$ months, 18 days; range = 13 months, 8 days to 13 months, 29 days). Half of the infants were randomly assigned to the effective-priming condition and half to the ineffective-priming condition; within each condition, 8 infants saw the unexpected event and 8 saw the expected event. An additional 5 infants were excluded because they were fussy (3, 2 in the effective-priming condition and 1 in the ineffective-priming condition) or had a test looking time over 2.5 standard deviations from the cell mean (2, in the effective-priming condition).

5.2.2. Apparatus, stimuli, and procedure

The apparatus was the same as in Experiment 3. The fake floor was used in the priming trials to cover the hole in the apparatus floor. The block (16 × 6 cm in diameter) was covered with dark blue adhesive paper and decorated with multicolored metallic stars. The cylinders were 10 cm in diameter and covered with white adhesive paper decorated with vertical green wavy stripes. They included a tall container (18 cm in height, used in the expected event), a short tube (9 cm, used in the unexpected event), and a tall tube (18 cm); the two tubes were shown to the infant prior to the testing session. A platform beneath the hole in the apparatus floor ensured that in the unexpected event, the block was lowered to the same depth (relative to the rim) as in the expected event.

The procedure was similar to that in Experiment 3. Infants were attentive during the initial phases of the priming and test trials; across conditions, they looked, on average, for 95% of each initial phase. The criteria used to end the final phase of each priming trial were the same as in Experiment 2 (where infants also heard labels), and the criteria used to end the final phase of the test trial were the same as in Experiment 3. Interobserver agreement for the test trial was calculated for 31/32 infants (only one observer was present for one infant) and averaged 91% per infant. Finally, preliminary analyses of the test data revealed no significant interactions of condition and event with infants' sex, $p = .113$; the data were therefore collapsed across this factor.

5.3. Results

Looking times during the final phases of the priming trials were averaged and analyzed using a one-way ANOVA with condition (effective-priming, ineffective-priming) as a between-subjects factor. This analysis revealed that infants in the effective-priming condition ($M = 15.88$, $SD = 3.17$) looked significantly longer than infants in the ineffective-priming condition ($M = 10.77$, $SD = 4.17$), $F(1, 30) = 15.20$, $p < .001$. Although infants apparently found the block more interesting when it was given an object label than when it was not (for similar differences, see e.g., Fulkerson & Waxman, 2007; LaTourette & Waxman, 2020), both conditions still demonstrated robust attention in the priming trials: Across the initial and final phases of the three trials, infants looked for a total of 76.02 s ($SD = 10.32$) in the effective-priming condition and 60.86 s ($SD = 13.06$) in the ineffective-priming condition. Moreover, analysis of infants' test responses using a multiple linear regression revealed no significant interaction of condition and event with infants' averaged looking time in the priming trials ($\beta = 1.32$, $SE = 1.20$, $t = 1.10$, $p = .282$), suggesting that this difference between conditions did not substantially affect our test results.

Looking times during the final phase of the test trial (Fig. 5) were analyzed using a 2 × 2 ANOVA with condition (effective-priming, ineffective-priming) and test event (unexpected, expected) as between-subjects factors. The analysis yielded a significant Condition × Event interaction, $F(1, 28) = 5.55$, $p = .026$, $\eta_p^2 = 0.165$ (no such interaction was found when responses in the final phases of the priming trials were analyzed in the same manner, $F(1, 28) = 0.01$, $p = .938$). Planned comparisons indicated that in the effective-priming condition, infants looked significantly longer if shown the unexpected event ($M = 16.81$, $SD = 5.00$) as opposed to the expected event ($M = 11.93$, $SD = 2.65$), $F(1, 28) = 4.91$, $p = .035$, $d = 1.22$, whereas in the ineffective-priming condition, infants looked equally at the unexpected ($M = 13.01$, $SD = 3.94$) and expected ($M = 15.46$, $SD = 5.49$) events, $F(1, 28) = 1.24$, $p = .275$, $d = -0.51$. Wilcoxon rank-sum tests revealed a marginally significant preference for the unexpected event in the effective-priming condition ($Z = 1.89$, $p = .059$) and no preference for either event in the ineffective-priming condition ($Z = -0.95$, $p = .344$).

5.4. Discussion

Infants who heard "A blicket!" as the block was tilted back and forth in the priming trials looked significantly longer if shown the unexpected as opposed to the expected event in the test trial; in contrast, infants who heard "Blicket!" looked equally at the events. These results suggest that when the word blicket was introduced in a nominal naming phrase, infants interpreted it as an object label. Because the boundaries of the object category referred to by this label were unclear, the OF system bound to the label a few of the block's featural properties, including its height. Consequently, when the OF system passed on the block's label (along with other categorical descriptors) to the PR system at the start of the test event, information about the block's height was serendipitously passed on as well. Once included in the PR system's event representation, this height information enabled infants to detect the persistence violation in the unexpected event: An object cannot become fully hidden inside a shorter tube. In contrast, infants who simply heard "Blicket!" did not interpret the word as an object label, and they therefore failed to detect the violation in the unexpected event.

These results thus confirm those of Experiment 2 and provide additional evidence that simply giving an object a novel label can lead infants to detect a persistence violation involving the object that they would otherwise not detect. In addition, the results extend those of Experiment 2 in four ways. First, Experiment 2 focused on a color violation in a containment event, whereas Experiment 4 focused on a height violation in a tube event. Together, these results suggest that when infants first hear a novel label applied to an object, standard features that are tied to the label (at least until more evidence is gathered; LaTourette & Waxman, 2020) include color (Experiment 2) and height (Experiment 4). Second, after (as-yet-unidentified) features are serendipitously passed on to the PR system, infants can detect at least two types of persistence violations involving the features, change violations (Experiment 2) and interaction violations (Experiment 4). Third, infants can detect violations involving either an object that goes out of view and becomes visible again (Experiment 2) or an object that goes out of view and remains hidden (Experiment 4). Infants' success following verbal priming manipulations is thus not limited to situations where they can mentally compare separate viewings of the labeled object. Finally, infants can detect violations when shown either multiple repetitions of an unexpected event (Experiment 2) or a single unexpected event (Experiment 4).

6. Experiment 5

Experiments 1–4 showed that following priming manipulations designed to induce the OF system to pass on information about an as-yet-unidentified feature to the PR system, infants succeed at VOE tasks involving the feature. In Experiment 5, we sought to broaden our findings by examining whether infants could also reason successfully about a primed feature in an *action* task, instead of a VOE task.

Prior research indicates that infants can demonstrate the same physical reasoning in VOE and action tasks, provided overall demands in the latter tasks do not overwhelm their limited information-processing resources (Hauf et al., 2012). To illustrate, consider a preferential-reaching task (Hespos & Baillargeon, 2006) that built on the VOE finding that by 7.5 months, most infants have identified height as a containment feature and detect a violation when a tall object is hidden in a short container (Hespos & Baillargeon, 2001a). In the task, 7.5-month-olds sat at a table across from an experimenter. A tall (21-cm) toy frog was removed from behind a large screen and seated in front of it; the frog had prominent legs and feet that jutted toward the infant. Next, the frog was returned behind the screen, which was then removed to reveal a platform on which rested a tall (23-cm) and a short (8.5-cm) container. Two legs, identical to the frog's, jutted from small holes at the bottom of each container; no frog protruded from the top of either container. The platform was pushed toward the infant, who was encouraged to search for the frog (the platform ensured that both containers were moved in tandem). Infants reached preferentially for the tall as opposed to the short container, and this effect was eliminated when infants did not first see the frog. Thus, when tested with a simple action task (e.g., the frog's legs remained visible, reducing representational demands, and the task required only a preferential reach, reducing action demands), 7.5-month-olds demonstrated that they had identified height as a containment feature: They realized that the frog could be hidden only in the tall container.

Experiment 5 tested 10-month-olds and was modeled after Experiment 3 except that it used a preferential-reaching task (Fig. 7). In the test trial of the *baseline* group, infants searched for a medium-height blue Cookie Monster toy in either a short and a tall container (*container* condition) or a short and a tall tube (*tube* condition). In line with prior findings (Hespos & Baillargeon, 2006; Wang & Kohne, 2007), infants in the container condition reached preferentially for the tall as opposed to the short container, whereas infants in the tube condition reached randomly for either tube. Infants in the *priming* group were tested using the same procedure as in the tube condition except that they first received a priming trial presenting an object array, a row of Cookie Monster toys. Infants saw three toys that differed only in height (short, medium, and tall; *effective-priming* condition), three identical medium toys (*ineffective-priming* condition), or three medium toys that differed only in color (green, blue, and orange; *irrelevant-priming* condition).

As before, we reasoned that if the priming trial in the effective-priming condition was sufficient to highlight height information in the OF system, then this information might be serendipitously passed on to the PR system in the test trial, leading infants (a) to realize, in accordance with the persistence principle, that the medium Cookie Monster toy could be hidden only in the tall tube and hence (b) to reach for that tube. In contrast, the priming trial in the ineffective-priming condition should be insufficient to render height information salient in the OF system (as the toys were identical in height), so infants should reach randomly for either tube, as in the tube condition of the baseline group. Finally, infants should also fail in the irrelevant-priming condition as the priming trial was likely to highlight color as opposed to height information in the OF system; passing on this color information to the PR system in the test trial should not help infants take into account the relative heights of the toy and tubes when determining where the toy was likely to be hidden. Such a negative result would rule out the possibility that it was seeing three different toys from the same category, as opposed to seeing three toys that differed only in height, that made possible infants' success in the effective-priming condition.

6.1. Design

The experimenter stood at a large table across from the infant, who sat on a parent's lap in a cut-out area of the table; a large screen rested on the table between them. Each trial began with the experimenter removing the screen (by bringing it closer to her end of the table) to reveal a wide platform; she then slid the platform slightly forward and paused for 3 s to allow the infant to inspect the stimuli on the platform ("Look at these!"). Each trial ended with the experimenter sliding the platform back and replacing the screen.

Infants in the *baseline* group were randomly assigned to a container or a tube condition, and they received one familiarization and one test trial. At the start of the familiarization trial in the *container* condition, a tall and a short brown rectangular container stood on either side of the platform; each container had an open slit at the bottom of its front wall (Fig. 7). The experimenter picked up one of the containers, moved it in front of the center of the platform, showed it to the infant (i.e., by tapping its front, opening, and bottom), and finally placed it upright in front of the infant. After about 10 s, the experimenter retrieved the container and put it back on the platform. She then repeated the same procedure with the other container, and the trial ended. At the start of the test trial, a medium-height blue Cookie Monster toy now stood between the two containers; the toy had four yellow pompoms attached to its feet. The experimenter placed the toy in front of the infant ("Do you want to play with it?"). After about 10 s, the experimenter retrieved the toy from the infant and replaced it on the platform. Next, she held the toy above one of the containers and said, "Look, I can hide it here!" while lowering

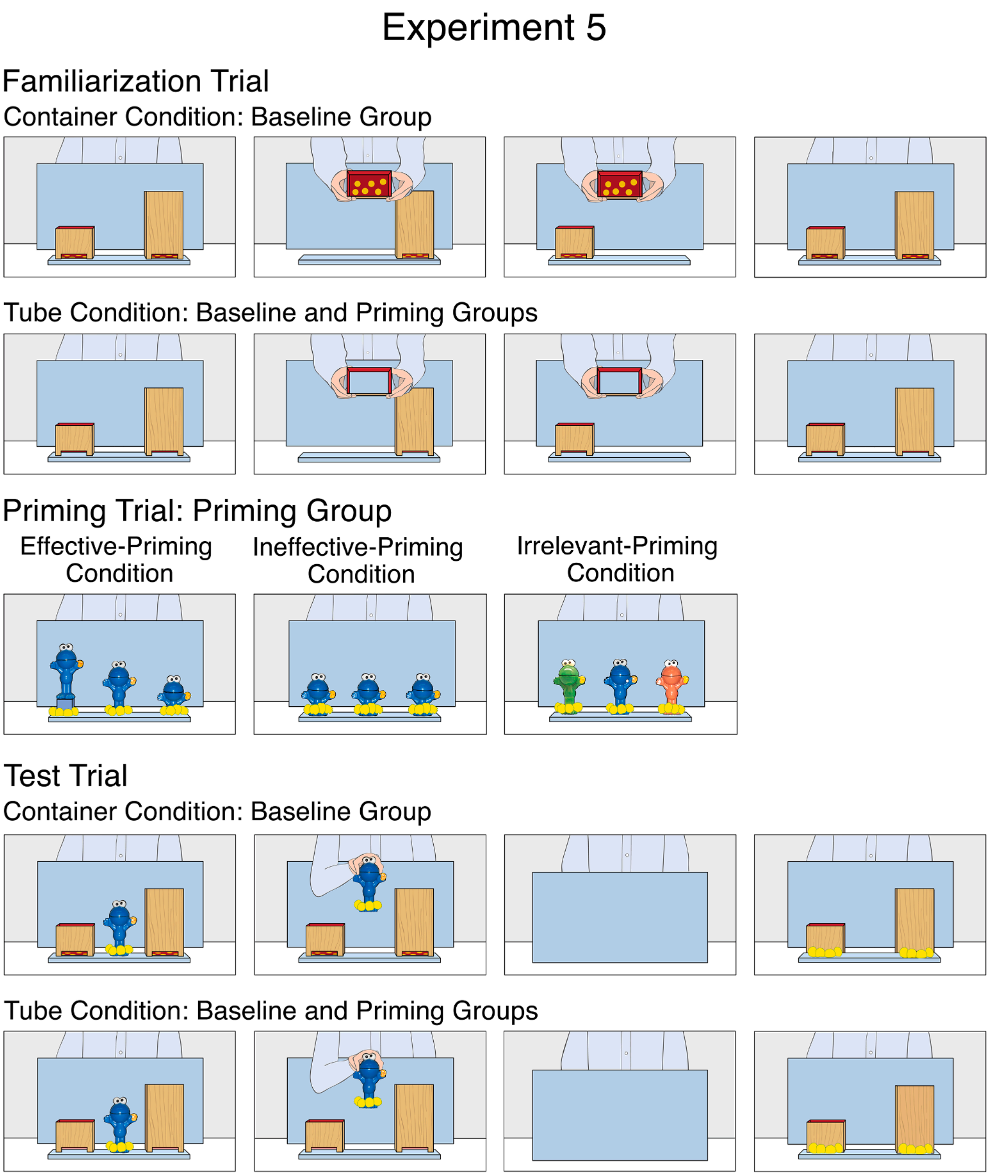

Fig. 7. Schematic depiction of the events in the familiarization, priming, and test trials of each condition in Experiment 5. (For information about the toys' colors in the priming trial of the irrelevant-priming condition, the reader is referred to the text or the web version of this article.)

its feet into the container three times. She then repeated these actions with the other container. After returning the toy to the platform, she replaced the screen and lifted the toy above it while saying, "Look, I am going to hide it!" twice. Next, she crouched behind the screen (to hide her actions from the infant), put away the toy, and inserted a set of pompons into the slit of each container. She then removed the screen one last time, pointed to each container in turn ("Look here!"), and finally pushed the platform toward the infant while saying, "Can you find it?" Infants were given 10 s to respond; during that time, the experimenter looked at the center of the platform, to avoid inadvertently cueing the infant. If the infant did not respond, the prompt was repeated, and infants were again given 10 s to respond. Infants who still did not respond were excluded and replaced (see Participants section). Trials in the *tube* condition were identical except that the tall and short tubes were used. Across infants, we counterbalanced whether the tall container/tube was on the left or the right side of the platform in the session, and whether the experimenter first acted on the tall or the short container/tube in the session.

Infants in the *priming* group received the same familiarization and test trials as in the tube condition, with a priming trial added between them. In the *effective-priming* condition, the screen was removed to reveal the platform with three, equally spaced, blue Cookie Monster toys that differed only in height (from left to right, these were short, medium, and tall for infants who saw the tall tube on the right side of the platform, and the reverse for infants who saw the tall tube on the left). The experimenter lifted and tilted each toy in turn (beginning with the tall toy if infants saw the experimenter act on the tall tube first, and with the short toy otherwise). Each toy was lifted a total of three times, and then the trial ended. Infants in the *ineffective-priming* condition saw three identical medium-height blue Cookie Monster toys, and infants in the *irrelevant-priming* condition saw three medium-height Cookie Monster toys that differed only in color (from left to right, these were green, blue, and orange).

6.2. Method

6.2.1. Participants

Participants were 80 healthy term infants (39 male, $M = 10$ months, 3 days, range = 9 months, 6 days, to 11 months, 15 days). Of these, 32 were assigned to the baseline group (with 16 each in the container and tube conditions), and 48 to the priming group (with 16 each in the effective-, ineffective-, and irrelevant-priming conditions). An additional 22 infants were tested but excluded. For 6 infants (1 in the container condition, 3 in the effective-priming condition, and 2 in the irrelevant-priming condition), the experimenter misjudged how far the infant could reach and did not push the platform close enough in the test trial (this often caused the infant to quickly reach back and forth for either container or tube in hopes of contacting one). Of the remaining infants, 5 failed to reach at all (1 in the container condition, 2 in the tube condition, 1 in the ineffective-priming condition, and 1 in the irrelevant-priming condition); 4 had reaches that were unclear or ambiguous (e.g., the infant was no longer centered in front of the platform; 1 in the tube condition, 2 in the effective-priming condition, and 1 in the irrelevant-priming condition); and the others were fussy (3, 1 in the tube condition, 1 in the ineffective-priming condition, and 1 in the irrelevant-priming condition), active (3, 1 in the container condition, 1 in the effective-priming condition, and 1 in the ineffective-priming condition), or distracted (1, in the effective-priming condition).

6.2.2. Apparatus and stimuli

The table (115.5 × 125.5 cm, with a cut-out area of 30 × 25 cm), platform (1 × 46.5 × 11 cm), and screen (30 × 58 cm) were covered with pastel adhesive paper. The experimenter wore a pale shirt and stood at one end of the table, across from the cut-out area. White ceiling-to-floor curtains behind the experimenter and on either side of the table hid the testing room.

The containers and tubes were rectangular, 12.5 cm wide, and 7.5 cm deep; they had a narrow slit (1.5 × 8.5 cm) at the bottom of their front walls; their exteriors were covered with brown wood-patterned adhesive paper; and their interiors were covered with red adhesive paper (the containers' interior bottoms were also decorated with yellow dots, which could be seen through the slits). The tall container and tube were 22 cm tall, and the short container and tube were 10 cm tall. The medium Cookie Monster toy used in the familiarization and test trials (19 × 10 × 7 cm at the widest points) was blue and had four bright yellow pompoms (each 3 cm in diameter) on its feet. Other Cookie Monster toys used in the priming trials also had pompoms and included four medium toys (19 cm), two blue, one green, and one orange; one short blue toy (14 cm; part of the toy's middle section was removed), and one tall blue toy (24 cm tall; it stood on a blue block). Finally, two magnetized sets of four pompoms could be inserted into the containers' or tubes' slits, to suggest that toys were in them (this also helped reduce representational demands, by keeping the toys partly visible).

Two video cameras recorded the infant's reaching behavior. One camera captured a side view of the infant, and the other captured an overhead view of the events and the infant's responses. The two images were combined and recorded to be checked offline for accuracy.

6.2.3. Procedure

Infants sat on a parent's lap in the cut-out area of the table. Parents were instructed to hold their infant centered on their lap (to keep the infant equidistant from either container or tube) and to close their eyes and remain silent during the trials. Each testing session began with a warm-up interaction between the experimenter and the infant; a penguin toy rested on the table in front of the screen, and the experimenter encouraged the infant to reach for it. Once the infant seemed comfortable, the toy was put away and the session proper began.

For each infant's test trial, once the platform was pushed within reach, the first touch to a container/tube or the pompoms protruding from its slit was coded as the infant's response. A second independent coder viewed videotaped sessions edited to include only the end of the test trial, when the platform was pushed within reach; this ensured that the coder was naïve about which group or condition the infant was assigned to. Only 76/80 infants could be rated by the naïve coder, as no video was available for the other infants due to technical difficulties. Coders assessed two issues. One was whether the infant made a clear reach for the tall or the short container/tube, and the two coders agreed on 74/76 infants. The other issue was whether the infant made contact with the container/tube or stopped short of it (a response that might be perceived as ambiguous), and the two coders agreed on 70/76 infants. Trials with discrepancies between the two coders were reviewed and resolved through frame-by-frame analysis if needed.

Finally, preliminary analyses of the test data in each group, carried out using logistic regressions, revealed no interaction of condition with infants' sex, side of the tall container/tube, and order in which the tall and short containers/tubes were acted on, all ps > 0.292; the data were therefore collapsed across these latter three factors.

6.3. Results

We first examined how many infants in each condition reached for the tall as opposed to the short container/tube using one-tailed cumulative binomial probabilities. The use of one-tailed tests was justified by the prior findings, reported earlier, that when faced with a tall and a short container, each with two frog feet jutting from small holes at the bottom, 7.5-month-olds reached for the tall container if they were first shown a tall frog, but reached randomly for either container otherwise (Hespos & Baillargeon, 2006; see also Wang & Kohne, 2007). Based on these findings, we predicted that infants in the container condition would reach for the tall container, infants in the effective-priming condition would reach for the tall tube, and infants in the other conditions would reach equally for either tube (tube, ineffective-priming, and irrelevant priming conditions). Results supported these predictions. In the baseline group (Fig. 8), 12/16 infants in the container condition reached for the tall container ($p = .038$, cumulative binomial probability), but only 9/16 infants in the tube condition reached for the tall tube ($p = .402$). In the priming group, 12/16 infants in the effective-priming condition reached for the tall tube ($p = .038$), but only 7/16 infants in the ineffective-priming condition ($p = .773$) and 5/16 infants in the irrelevant-priming condition ($p = .962$) did so.

To compare responses across conditions, we conducted two logistic regressions. The first compared two sets of conditions: the set where we predicted and found success (container and effective-priming conditions, with 24/32 infants reaching for the tall container/tube) and the set where we predicted and found random responding (tube, ineffective-priming, and irrelevant-priming conditions, with 21/48 infants reaching for the tall tube). The difference between the two sets was significant, $\beta = 1.35$, $SE = 0.50$, $z = 2.69$, $p = .007$.

The second logistic regression compared the three conditions in the priming group. Priming condition was entered as a pair of orthogonal contrast codes. The first contrast compared the effective-priming condition (1.00) to the average of the ineffective-priming (-0.50) and irrelevant-priming (-0.50) conditions, and this contrast was significant, $\beta = 1.08$, $SE = 0.46$, $z = 2.36$, $p = .018$. The second contrast compared the ineffective-priming condition (-0.50) to the irrelevant-priming condition (0.50), and this contrast was not significant, $\beta = 0.54$, $SE = 0.74$, $z = 0.73$, $p = .467$.

6.4. Discussion

In the baseline group, infants in the container condition were significantly more likely to reach for the tall as opposed to the short container. In contrast, infants in the tube condition reached randomly for either tube. Together, these results confirmed prior findings that by 10 months, most infants have identified height as a containment feature, but not a tube feature (Hespos & Baillargeon, 2001a; Wang et al., 2005).

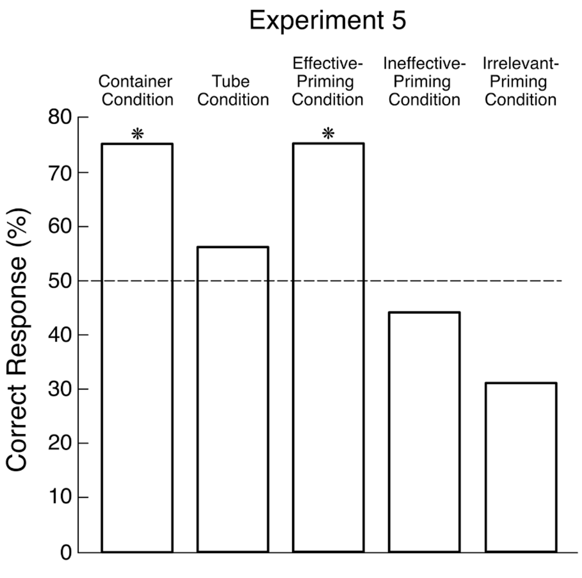

Fig. 8. Proportion of infants who reached for the tall container/tube in the test trial, by condition. An asterisk denotes a significant difference from chance performance (0.50).

In the priming group, infants in the effective-priming condition reached preferentially for the tall tube, suggesting that exposure to the small, medium, and tall Cookie Monster toys in the priming trial was sufficient to highlight height information in the OF system. This caused the OF system, at the start of the test trial, to spontaneously pass on information about the toy's height to the PR system. Once interpreted by the PR system's domain knowledge, this height information enabled infants to determine that the toy could be hidden only in the tall tube. In contrast, infants in the ineffective- and irrelevant-priming conditions tended to reach randomly for either tube, as in the tube condition of the baseline group. Because neither priming trial highlighted height information, this information was not available to the PR system, who therefore had no basis for determining which tube was likely to hold the medium Cookie Monster toy hidden by the experimenter.

Together, the results of Experiments 3–5 thus provided converging evidence that infants who have not yet identified height as a tube feature but are effectively primed, via object arrays or novel labels, to include height information in their representation of a tube event, can then (a) detect a change violation involving height (Experiment 3), (b) detect an interaction violation involving height (Experiment 4), or (c) succeed at a preferential-reaching task involving height (Experiment 5).

7. Experiment 6

Experiments 1–5 showed that if the OF system is primed, at the start of a containment or tube event, to send information about an as-yet-unidentified feature to the PR system, this information is then interpreted in accordance with the core principle of persistence, allowing infants to succeed at VOE and preferential-reaching tasks involving the feature. In a final, exploratory experiment, we asked whether the same conclusions would hold for a different event category, support events, and a different core principle, gravity (objects fall when inadequately supported). Specifically, if the OF system was primed, at the start of a support event, to send information about an as-yet-unidentified feature to the PR system, would this information then be interpreted by the gravity principle, allowing infants to succeed at a VOE task involving the feature?

Prior research on events involving passive support from below indicates that over the course of many months, infants identify a series of features that helps them predict outcomes more and more accurately; in this series, each new feature revises predictions from earlier features (Baillargeon & DeJong, 2017; Lin et al., in press). By about 6 months, most infants have identified *proportion-of-contact* as a causally relevant feature: When an object is released with one end on a base, they expect it to remain stable if half or more of its bottom surface lies on the base, and to fall otherwise (Baillargeon et al., 1992; Hespos & Baillargeon, 2008; Luo et al., 2009; Wang et al., 2016). This rule correctly predicts outcomes for symmetrical objects, but it can lead to errors for asymmetrical objects. For example, an L-shaped block will fall when released with half of its bottom surface on a base if it is the smaller, horizontal part of the block (as opposed to its larger, vertical part) that is supported. By about 13 months, most infants have acquired a more sophisticated *proportional-distribution* rule that applies to asymmetrical as well as symmetrical objects: When an object is released with one end on a base, they now attend to the whole object, not just its bottom surface, and they expect stability as long as half or more of the *whole object* is on the base (Baillargeon & DeJong, 2017). Though the exact explanation infants generate for this feature is still unknown, one proposal has been that infants mentally decompose the object into two parts, the supported part on the base and the unsupported part off the base; when the latter part is larger, they expect the object to tip off the base, in accordance with the gravity principle (Baillargeon & DeJong, 2017). Building on this proposal, Experiment 6 tested whether 7-month-olds who were induced to represent the relative sizes of the supported and unsupported parts of an asymmetrical object (i.e., to represent the proportional-distribution information about the object) would then find it unexpected if the object remained stable with its larger part unsupported.

Infants received a test trial in which they saw an experimenter in a back window place an asymmetrical test block on a base in such a way that only half of the block's bottom surface was on the base; the block always remained stable when released (Fig. 9). The block was shaped like an L from the experimenter's perspective and like a reverse L from the infant's perspective. In the *expected* event, the larger, vertical part of the block was on the base; in the *unexpected* event, the smaller, horizontal part of the block was on the base. Prior to the test trial, infants received two priming trials. In the *effective-priming* condition, they saw a row of four blocks, each with two parts: a yellow part and a white part with an orange pattern. In the first priming trial, the two parts of the leftmost block (block-1) were cubes identical in size. In each successive block, the yellow part grew in height by the equivalent of one additional cube and so was as tall as two (block-2), three (block-3), or four (block-4) stacked cubes; the test block was identical to block-4. The second priming trial was identical except that the row of blocks was reversed (from left to right, infants now saw block-4, block-3, and so on).

We reasoned that if exposure to the four different blocks in the priming trials highlighted information in the OF system about the relative sizes of the two parts in each block, then this information might be passed on to the PR system at the start of the test trial. As a result, when the test block was released with only one of its parts supported, infants might compare the size of that part to that of the unsupported part (instead of merely comparing their bottom surfaces, in line with the proportion-of-contact rule). Once interpreted by the PR system's domain knowledge, this proportional-distribution information should lead infants to expect the block to remain stable when its larger part was supported, but to fall when its larger part was unsupported. Infants should thus look significantly longer if shown the unexpected as opposed to the expected event.

Infants in the *ineffective-priming* condition saw similar priming trials except that all four blocks were identical to block-4. These trials still presented infants with asymmetrical objects with two distinct parts, but they no longer called attention to the parts' relative sizes, which did not vary across blocks. When infants brought to bear their proportion-of-contact rule in the test trial, they should view either event as expected because in each case half of the test block's bottom surface was supported. Infants should thus look equally whether they were shown the expected or the unexpected event. Such a negative finding would rule out alternative interpretations of positive results in the effective-priming condition, such as that exposure to the test block (block-4) prior to the test event facilitated infants' processing of the event.

7.1. Design

Infants were randomly assigned to an effective- or an ineffective-priming condition, and they received two priming trials, one familiarization trial, and one test trial. Each trial had an initial phase and a final phase. In the *effective-priming* condition, the (30-s) initial phase of the first priming trial began with block-1 to block-4 standing in front of a platform that was as wide as the apparatus and positioned against the back wall; an experimenter's gloved hand was resting on the ledge of a short wide window centered in the back wall. To start, the hand grasped block-1 (positioned 12 cm from the left wall), placed it in the same position on the platform, paused at the window, and repeated these actions until all four blocks were on the platform. In each case, the experimenter grasped the patterned part of the block and, when moving the block, kept its bottom parallel to the floor, so that her actions were similar across blocks. Finally, she withdrew her hand and closed the window. In the final phase, infants watched this paused scene until the trial ended. The second priming trial was identical except that the row of blocks was reversed, with block-4 now on the left; the hand again placed each block on the base, moving from left to right. The priming trials thus served both to present the four blocks and to show that the experimenter produced support events, as she moved each block to the platform.

The familiarization trial again showed a support event, this time involving the two objects that would be used in the test trial: block-4 and a base that was slightly wider than block-4. At the beginning of the (6-s) initial phase, block-4 stood in front of the platform in its rightmost position (as in the first priming trial); the base was centered in front of the platform, and the gloved hand rested at the window. The hand grasped the block, centered it on the base (so that it was fully supported), and released it. During the final phase, the experimenter withdrew her hand and closed the window, and infants watched this paused scene until the trial ended.

The (6-s) initial phase of the test trial was identical to that of the familiarization trial except that only one part of block-4 was placed on the base. In the *unexpected* event, the hand placed the patterned part on the right end of the base, leaving the yellow (taller) part unsupported. In the *expected* event, the hand placed the yellow part on the left end of the base, leaving the patterned (shorter) part unsupported. In the final phase of each event, the hand paused briefly, grasped the block again, lifted it about 10 cm, lowered it back onto the base, and released it; these actions (which lasted about 8 s) were repeated until the trial ended. Given the young age of our

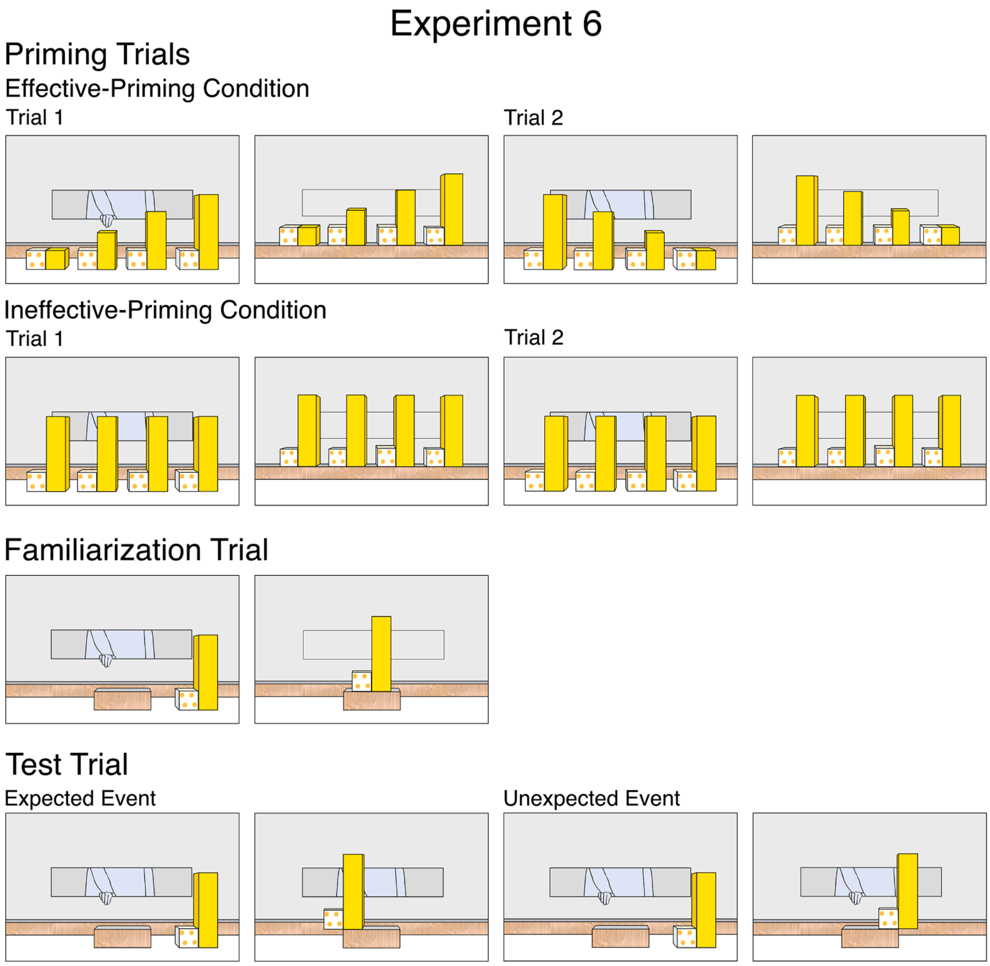

Fig. 9. Schematic depiction of the events in the priming, familiarization, and test trials of each condition in Experiment 6.

participants, we thought it better to give them more than one opportunity to observe that the block remained stable when released.

Trials in the *ineffective-priming* condition were identical with one exception: In the priming trials, all four blocks were identical to block-4. From infants' perspective, the two priming trials were thus identical as reversing the row of blocks had no perceptual effect. Nevertheless, reversing the row in the second trial equated procedures across conditions and kept the observers naïve about which condition the infant was assigned to.

One feature of our design deserves some explanation. In our previous experiments, infants who received priming trials did not see, prior to test, primed objects involved in events similar to the test events (i.e., primed infants in Experiments 1–2 did not see containment events with Boohbah toys prior to test, and primed infants in Experiments 3–5 did not see tube events with blocks or toys prior to test). In Experiment 6, however, infants did see support events involving block-4 in the priming and familiarization trials. Pilot data without these initial support events were less promising, so we adopted the design described here. We speculated that for our 7-month-old participants, seeing block-4 involved in support events in advance of the test trial might help reduce overall processing demands, making it easier for the PR system to interpret the novel proportional-distribution information passed on by the OF system. Crucially, the support events shown in the priming and familiarization trials did not involve proportional-distribution information, as all blocks were fully supported when released. In this respect, Experiment 6 was thus similar to our previous experiments and provided another test of the prediction that infants should reason successfully about an as-yet-unidentified feature in an event if the OF system was primed to send information about the feature to the PR system at the start of the event.

7.2. Methods

7.2.1. Participants

Participants were 32 healthy terms infants (17 male, $M = 7$ months, 16 days; range = 6 months, 24 days to 8 months, 6 days). Half of the infants were randomly assigned to the effective-priming condition and half to the ineffective-priming condition; within each condition, 8 infants saw the unexpected event and 8 saw the expected event. One additional infant was excluded due to fussiness (in the effective-priming condition).

7.2.2. Apparatus, stimuli, and procedure

The apparatus was similar to that in Experiment 4, with the following exceptions. The back window (15 × 74 cm) was centered in the back wall 18 cm above the floor; it had no curtain but could be closed by lifting a panel covered with the same adhesive paper as the back wall. The experimenter sat on a high stool and wore a pale shirt and long pale gloves; white ceiling-to-floor curtains behind her hid the testing room. A wide platform (7.5 × 100 × 7.5 cm) stretched across the apparatus, against the back wall; it was covered with a brown wood-patterned adhesive paper, and its top surface was also covered with black felt, to minimize noises when blocks were placed on it. Stimuli included a base similar to the platform but much narrower (7.5 × 22.5 × 7.5 cm), and seven L-shaped blocks. Each block had a white part with an orange pattern and a yellow part; the patterned part was always a cube 7.5 cm a side, and the yellow part consisted of one cube (block-1, 7.5 cm tall), two stacked cubes (block-2, 15 cm tall), three stacked cubes (block-3; 22.5 cm tall), or four stacked cubes (block-4; 30 cm tall); the stacked cubes' individual boundaries were not apparent under the adhesive paper. There were four identical blocks-4 (for the ineffective-priming condition priming trials), and one of them was used as the test block. In each priming trial, the blocks stood 5.5 cm apart, 16 cm in front of the platform, and the whole row was centered 12 cm from either wall.

The procedure was similar to that in Experiment 4, except as follows. Infants in each condition received two priming trials, one familiarization trial, and one test trial, each with an initial and a final phase. Infants were highly attentive during these initial phases and looked, on average, for 92% of each initial phase. The criteria used to end the final phases of the priming trials were the same as in Experiment 4; these criteria were also used to end the final phase of the familiarization trial. The criteria used to end the final phase of the test trial were also identical to those in Experiment 4, with one exception: The look-away criterion was reduced from 2 to 0.5 consecutive seconds. Because the final phase in Experiment 6 involved not a paused scene but a repeated 8-s event cycle (the test block was continually lifted and lowered back onto the base until the trial ended), infants were less likely to look away for as long as 2 consecutive seconds because the experimenter's repeated actions tended to recapture their attention (infants' young age might also have contributed to their enhanced attention). The minimum look of the final phase of the test trial, 8 s, was the same as before and corresponded to one event cycle in the final phase; infants thus had the opportunity to see that the test block remained stable when released at least twice before the trial could end, once in the initial phase and once in the final phase. Inter-observer agreement in the final phase of the test trial was calculated for all 32 infants and averaged 93% per infant. Preliminary analyses of the test data revealed no significant interaction of condition and event with infants' sex, $p = .676$; the data were therefore collapsed across this factor in subsequent analyses.

7.3. Results

Looking times in the final phases of the *priming* trials were averaged and analyzed as in Experiment 4. The main effect of condition was not significant, $F(1, 30) = 0.00$, $p = .971$, indicating that infants in the two conditions looked about equally (effective-priming: $M = 9.63$, $SD = 5.29$; ineffective-priming: $M = 9.57$, $SD = 4.11$). Looking times during the final phase of the *familiarization* trial were analyzed as above and again did not differ across conditions, $F(1, 30) = 0.53$, $p = .472$ (effective-priming: $M = 12.83$, $SD = 6.16$; ineffective-priming: $M = 11.22$, $SD = 6.30$).

Looking times in the final phase of the *test* trial (Fig. 10) were analyzed as in Experiment 4. The analysis yielded only a significant Condition × Event interaction, $F(1, 28) = 5.11$, $p = .032$, $\eta_p^2 = 0.154$ (no such interaction was found when responses in the final phases of the priming trials, $F(1, 28) = 0.68$, $p = .416$, and familiarization trial, $F(1, 28) = 0.72$, $p = .404$, were analyzed in the same manner). Planned comparisons indicated that in the effective-priming condition, infants who saw the unexpected event ($M = 20.48$, $SD = 2.26$) looked significantly longer than those who saw the expected event ($M = 14.91$, $SD = 5.28$), $F(1, 28) = 4.29$, $p = .048$, $d = 1.37$, whereas in the ineffective-priming condition, infants who saw the unexpected ($M = 15.14$, $SD = 6.74$) and expected ($M = 18.16$, $SD = 6.09$) events looked about equally, $F(1, 28) = 1.27$, $p = 0.270$, $d = -0.47$. Wilcoxon rank-sum tests confirmed the results of the effective-priming ($Z = 2.42$, $p = .016$) and ineffective-priming ($Z = -1.12$, $p = .265$) conditions.

7.4. Discussion

Infants in the effective-priming condition looked significantly longer if shown the unexpected as opposed to the expected event. This positive result suggests that exposure in the priming trials to four asymmetrical blocks that differed only in the relative sizes of their two parts highlighted this information in the OF system. This caused the OF system, at the beginning of the test trial, to send the PR system information about the relative sizes of the two parts of the test block (block-4). As a result, when the block was released with only one of its parts on the base, infants compared the sizes of the supported and unsupported parts, instead of merely comparing their bottom surfaces (in line with the proportion-of-contact rule). Once interpreted by the PR system's domain knowledge, this proportional-distribution information allowed infants to detect the gravity violation in the unexpected event: The test block could not remain stable when released with its larger part unsupported. This result is striking in that infants succeeded in detecting this violation *six months* before they would normally do so: Recall that proportional distribution is typically not identified as a support feature until about 13 months of age (Baillargeon & DeJong, 2017). This result thus extends that of Experiment 3, where primed infants succeeded in detecting a persistence violation in a tube event six months before they would normally do so. Together, these results provide strong evidence that when learning about simple event categories, infants are learning primarily *what* featural information to include in their event representations to better predict outcomes, and not *how* to interpret that information. Their core knowledge allows them to make sense of relevant featural information whenever it is included in their event representations—even if it is included serendipitously and as many as six months before they begin to routinely include it in their event representations.

In contrast to infants in the effective-priming condition, those in the ineffective-priming condition looked about equally at the expected and unexpected events. This negative result confirmed prior findings with infants under 13 months of age (Baillargeon & DeJong, 2017), and it also ruled out the possibility that infants in the effective-priming condition succeeded simply because they saw the test block, or support events involving the test block (in the priming and familiarization trials), prior to the test trial.

8. Mini meta-analyses of VOE experiments

To assess the robustness of the test data in our five VOE experiments (Experiments 1–4 and 6), we ran a meta-analysis for the effective- and ineffective-priming conditions separately (Cumming, 2014; Field & Gillett, 2010; Hedges & Vevea, 1998). There was no

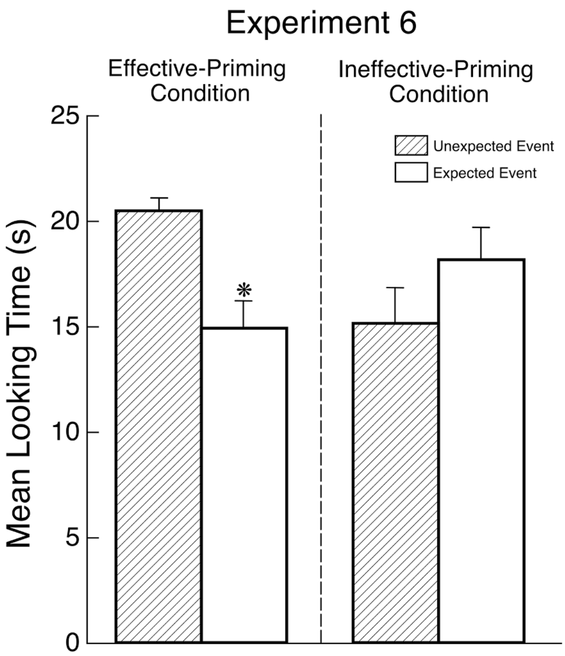

Fig. 10. Mean looking times in the final phase of the test trial in Experiment 6, by condition and event. Error bars represent standard errors, and an asterisk denotes a significant difference between events within a condition.

evidence of heterogeneity of effects across experiments (Cochran's Q tests, $ps > 0.623$), so we used a fixed-effects meta-analytic model. Infants in the effective-priming condition looked significantly longer at the unexpected than at the expected event, $d+ = 1.14$ [0.73, 1.55], $Z = 5.44$, $p < .0001$, whereas infants in the ineffective-priming condition looked equally at the two events, $d+ = -0.17$ [-0.55, 0.20], $Z = -0.91$, $p = .361$.

9. General discussion

The present research tested a prediction of the two-system model of early physical reasoning about simple event categories (Lin et al., in press; Stavans et al., 2019): Priming experiences that cause the OF system, at the start of an event, to pass on information about an as-yet-unidentified feature to the PR system should, serendipitously, allow infants to succeed at tasks involving this information. Across six experiments, we found strong support for this prediction, using two types of priming manipulations. First, after an object was given a novel label using a naming phrase (e.g., "A gormet!"), 13-month-olds detected an interaction violation when the object became fully hidden into a shorter tube (Experiment 4), and 12-month-olds detected a violation when the object surreptitiously changed color when briefly lowered into a container (Experiment 2). Second, following exposure to objects that were identical except for color, 12-month-olds detected a change violation when one of the objects surreptitiously changed color in a container (Experiment 1); following exposure to objects that were identical except for height, 10-month-olds searched for one of the objects in a tube tall enough to hide the object (Experiment 5), and 8-month-olds detected a change violation when one of the objects was much shorter after being briefly lowered into a tall tube (Experiment 3); finally, following exposure to objects that were identical except for the relative sizes of their two parts, 7-month-olds detected an interaction violation when one of the objects remained stable on a base with its larger part unsupported (Experiment 6). Because height is typically not identified as a tube feature until about 14 months (Wang et al., 2005), and proportional distribution is typically not identified as a support feature until about 13 months (Baillargeon & DeJong, 2017), these last three findings, in particular, demonstrate that primed infants successfully reasoned about these features as many as four to six months before they would normally do so.

We have suggested that these various findings can all be explained in the same general way: In each case, the priming manipulation served to expand what information the OF system passed on to the PR system at the start of an event. In the case of our novel-label manipulations, because the boundaries of the object category referred to by the label were unclear (e.g., were all gormets the same size, shape, and color as the labeled object?), the OF system bound to the label some of the featural properties of the object, including its size and color. Consequently, when the test event began and the OF system sent the object's label, along with other categorical descriptors, to the PR system, information about these features was passed on as well. Once included in the PR system's event representation and interpreted by its domain knowledge, this featural information allowed infants to detect the persistence violations they were shown. In the case of our object-array manipulations, because only one feature varied across the three or four objects in the array, this feature became salient and was passed on to the PR system, along with the usual categorical descriptors, at the start of the event. Here again, once included in the PR system's event representation and interpreted by its domain knowledge, this featural information allowed infants to detect the persistence or gravity violation they were shown.

The evidence reported in this article confirms and extends several prior findings in the physical-reasoning literature. In particular, it provides new evidence that infants may reason correctly about a feature when presented with events from one category but not another—even when these events are similar or indeed perceptually identical (as was the case for the containment and tube events in Experiment 3). Thus, in Experiments 1–2, 12-month-olds correctly reasoned about color information in an occlusion but not a containment event, and in Experiments 3–5, 8- to 13-month-olds correctly reasoned about height information in a containment but not a tube event.

In addition, our research extends prior attempts to induce infants to succeed at physical-reasoning tasks using various experimental manipulations. As was briefly discussed in the introduction, at least three types of manipulations have been used to date with good effect. First, in *categorical-encoding* experiments, infants succeeded at detecting individuation violations when their OF system was induced, via linguistic or functional manipulations, to assign the two objects shown in alternation to distinct basic-level categories; the OF and PR systems could then agree on how many objects were involved in the event (i.e., two objects), resulting in successful individuation (Futó et al., 2010; Stavans & Baillargeon, 2018; Xu, 2002). For example, after hearing a distinct label (e.g., "Look, a duck!", "Look, a ball!") as each object came into view, 9-month-olds detected a violation when the screen was lowered to reveal only one of the objects. In other categorical-encoding experiments, infants succeeded at detecting change violations when their OF system was induced to assign the pre- and post-change objects to distinct pattern- or color-based categories formed via arbitrary associations with specific functions (Wilcox & Chapa, 2004; Wilcox et al., 2011). For example, after seeing three different dotted cups being used to pound pegs and three different striped cups being used to pour salt, 4-month-olds detected a violation when a dotted ball changed into a striped ball behind a narrow occluder (Wilcox & Chapa, 2004). Second, in *teaching* experiments, infants succeeded at detecting interaction violations involving a particular feature after their PR system was helped to identify this feature using observations designed to support EBL (Baillargeon & DeJong, 2017; Wang & Baillargeon, 2008a; Wang & Kohne, 2007). For example, after being "taught" the feature height in covering events, 9-month-olds detected a violation in new events in which a tall object became fully hidden under a short cover (Wang & Baillargeon, 2008a). Third, in *carryover* experiments, infants succeeded at detecting change or interaction violations involving a particular feature if they saw the same objects in two successive events from different categories, one in which the feature had been identified followed by one in which the feature had not yet been identified (Baillargeon et al., 2009; Wang & Baillargeon, 2005; Wang & Onishi, 2017; Wang, 2011). For example, 4.5-month-olds detected an interaction violation when a tall object was first slid in front of a short container (occlusion event), returned to its initial position, and then lifted and lowered into the container until it became almost fully hidden (containment event; Wang, 2011). The idea here is that the PR system, having tapped the OF system for the appropriate featural information in the occlusion event, simply carried over (or re-used) that information for the containment event, without bothering to tap the OF system again. Once included in the PR system's representation of the containment event and interpreted by its domain knowledge, this featural information enabled infants to detect the persistence violation they were shown. Our experiments thus extend these various strands of research by showing yet another way in which infants can be induced to succeed at physical-reasoning tasks.

Given these various findings, could our results be explained in other ways? For example, one suggestion (inspired by the categorical-encoding findings described above) might be that infants succeeded at our tasks because our priming manipulations led them to form novel object categories. Although it seems likely that infants in many of our experiments did form novel object categories, we doubt that these categorical encodings alone were responsible for their success. Recall that the 10-month-olds in Experiment 5 succeeded when shown three Cookie Monster toys that differed in height, but not when shown three Cookie Monster toys that differed in color. It seems likely that in each case infants formed an object category—but they succeeded only when height was highlighted as a feature.

Another suggestion (inspired by the carryover findings described above) might be that our manipulations primed not the OF system but the PR system, leading it to request information about the primed, as-yet-unidentified feature. This possibility also seems unlikely. Recall, for example, that the 12-month-old infants in Experiment 1 simply saw a static array of Boohbah toys; the OF system would be engaged in representing this array, but the PR system would not, as no event (i.e., no physical interaction) took place (see Footnote 3). It thus seems more parsimonious to assume that the system that was engaged in representing the array was the one that was primed.

Yet another suggestion might be that our priming manipulations led the OF system to pass on information about the primed as-yet-unidentified feature not when passing on the usual categorical information at the start of an event, as we have been arguing, but when supplying the featural information requested by the PR system. In other words, perhaps the OF system simply tacked on the primed feature(s) to those requested by the PR system. One way to decide between these two possibilities might be as follows. If the information the OF system passes on at the start of an event is generally the same information it uses to individuate the objects in the event and to track them past its endpoint (arrow 3 from the OF system in Fig. 1), then any primed featural information the OF system passes on at the start of the event should also be used for these purposes. This predicts that infants should succeed at an individuation task following exposure to an appropriate object array. After seeing four different-color Boohbah toys, for example, infants should detect an individuation violation if two of the toys were next brought out in alternation from behind a screen that was then removed to reveal only one of the toys. With the OF and PR systems both using categorical and color information to individuate and track the toys, the systems should agree on the number of toys present, leading to a successful performance. Conversely, in the alternative model outlined above in which the primed color information is simply tacked on to the other featural information requested by the PR system, the OF system should still use only categorical information to individuate and track the toys, resulting in a disagreement with the PR system and hence a catastrophic individuation failure.

Beyond the prediction just described, our findings raise many questions for future research. For example, how long does our priming effect last? If a brief interval was inserted between our priming and test trials, would infants revert to failing at the tasks? In Wang's (2011) carryover experiment, for example, the positive effect found was eliminated if a 20-s delay was inserted between the occlusion and containment events (infants either were turned away from the apparatus or saw the tall object being moved back and forth next to the container). There thus appear to be sharp temporal limits to carryover effects, and it will be interesting to find out whether similar limits apply to the priming effects shown here.

Another question for future research concerns the range of manipulations that might be effective in priming infants to represent and reason about as-yet-unidentified features. One interesting prediction is that if we were to *combine* the visual and verbal manipulations used here, our priming effect would disappear. To see why, imagine that infants received four priming trials similar to those in Experiment 2, except that a Boohbah toy of a different color was labeled as "a gormet" in each trial. We suspect that under these conditions, the OF system would be unlikely to bind color information to this novel label; unlike oranges, which are typically orange, and strawberries, which are typically red, gormets would come in many colors, with no particular color serving as a defining, shared feature of objects in the category. Such a result would echo recent findings by LaTourette and Waxman (2020). In a training phase, 12-month-olds first heard the same novel label ("Look at the boff!") applied to four distinct members of the same object category, presented one at a time. At test, each training object was presented in silence, paired with a novel member of the category; the rationale was that if infants remembered which objects they had seen in the training phase, they would look significantly longer at the novel category member. However, results with the last two training objects (which were presented in the first two test trials) indicated that infants failed to recognize these objects, suggesting that applying the same label to all of the training objects led infants to focus on and remember their *common* features rather than their unique, individuating features. In line with these results, we are suggesting that labeling each of four different-color Boohbah toys as "a gormet" in a modified version of Experiment 2 would make it unlikely that color would be bound to the new label "gormet" in the OF system, leading infants to fail to detect the change violation in the unexpected event.

Yet another important direction for future research will be to determine how and when the OF system begins to spontaneously use featural information to track objects. We saw earlier that around the first birthday, a major development occurs in the OF system, which begins to include basic-level categories (e.g., 'ball') in its categorical descriptors for objects. Prior to 12 months, as mentioned above, the OF system can be induced via categorical-encoding manipulations to include basic-level categories in the information it uses to track objects from event to event (Futó et al., 2010; Stavans & Baillargeon, 2018; Xu, 2002). In an analogous way, our research suggests that before the OF system begins to use featural information to track objects, it can be primed to do so using visual or verbal manipulations designed to highlight this information.

10. Conclusion

One of our goals in the present article was to show that the two-system model (Lin et al., in press; Stavans et al., 2019) not only helps explain prior findings in the physical-reasoning literature but also makes novel predictions. Here we reported converging evidence from six experiments for one such prediction: Infants aged 7 to 13 months were induced to succeed at physical-reasoning tasks involving as-yet-unidentified features, sometimes as many as six months before they would normally do so, when their OF system was primed via visual or verbal means to pass on information about these features to the PR system. These results provide new evidence that understanding the nature and development of each system and teasing out the ways in which they interact can substantially advance our understanding of early physical reasoning.

Acknowledgments

This research was supported by a grant from NICHD to R.B. (HD-021104) and by grants from NICHD (HD-054448) and NSF (BCS-1348522) to C.L.F. We thank Lin Bian and Sandy Waxman for helpful comments and suggestions; the UIUC Infant Cognition Lab for their help with the data collection; Steve Holland for his help with the figures; Amélie Bernard for her help with the design and implementation of Experiment 6; and the families who participated in the research.

References

Aguiar, A., & Baillargeon, R. (1999). 2.5-month-old infants' reasoning about when objects should and should not be occluded. *Cognitive Psychology, 39*, 116–157.
Baillargeon, R. (1987). Object permanence in 3.5- and 4.5-month-old infants. *Developmental Psychology, 23*, 655–664.
Baillargeon, R. (1995). A model of physical reasoning in infancy. In C. Rovee-Collier, & L. P. Lipsitt (Eds.), *Advances in Infancy Research* (Vol. 9, pp. 305–371). Norwood, NJ: Ablex.
Baillargeon, R. (2008). Innate ideas revisited: For a principle of persistence in infants' physical reasoning. *Perspectives on Psychological Science, 3*, 2–13.
Baillargeon, R., & DeJong, G. F. (2017). Explanation-based learning in infancy. *Psychonomic Bulletin & Review, 24*, 1511–1526.
Baillargeon, R., & DeVos, J. (1991). Object permanence in young infants: Further evidence. *Child Development, 62*, 1227–1246.
Baillargeon, R., Li, J., Gertner, Y., & Wu, D. (2011). How do infants reason about physical events? In U. Goswami (Ed.), *The Wiley-Blackwell Handbook of Childhood Cognitive Development, 11* (2nd ed., pp. 11–48). Oxford, England: Blackwell.
Baillargeon, R., Li, J., Ng, W., & Yuan, S. (2009). An account of infants' physical reasoning. In A. Woodward, & A. Needham (Eds.), *Learning and The Infant Mind* (pp. 66–116). New York, NY: Oxford University Press.
Baillargeon, R., Needham, A., & DeVos, J. (1992). The development of young infants' intuitions about support. *Early Development and Parenting, 1*, 69–78.
Baillargeon, R., Stavans, M., Wu, D., Gertner, Y., Setoh, P., Kittredge, A. K., & Bernard, A. (2012). Object individuation and physical reasoning in infancy: An integrative account. *Language Learning and Development, 8*(1), 4–46.
Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology, 64*(1), 3–26.
Bonatti, L., Frot, E., Zangl, R., & Mehler, J. (2002). The human first hypothesis: Identification of conspecifics and individuation of objects in the young infant. *Cognitive Psychology, 44*, 388–426.
Booth, A. E., & Waxman, S. R. (2009). A horse of a different color: Specifying with precision infants' mappings of novel nouns and adjectives. *Child Development, 80*(1), 15–22.
Cacchione, T., Schaub, S., & Rakoczy, H. (2013). Fourteen-month-old infants infer the continuous identity of objects on the basis of nonvisible causal properties. *Developmental Psychology, 49*, 1325–1329.
Carey, S. (2011). *The origin of concepts*. New York, NY: Oxford University Press.
Casasola, M. (2008). The development of infants' spatial categories. *Current Directions in Psychological Science, 17*, 21–25.
Cumming, G. (2014). The new statistics: Why and how. *Psychological Science, 25*, 7–29.
Dewar, K., & Xu, F. (2007). Do 9-month-old infants expect distinct words to refer to kinds? *Developmental Psychology, 43*(5), 1227.
Fennell, C. T., & Waxman, S. R. (2010). What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development, 81*(5), 1376–1383.
Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology, 63*(3), 665–694.
Fischer, J., Mikhael, J. G., Tenenbaum, J. B., & Kanwisher, N. (2016). Functional neuroanatomy of intuitive physical inference. *Proceedings of the National Academy of Sciences, USA, 113*, E5072–E5081.
Fulkerson, A. L., & Waxman, S. R. (2007). Words (but not tones) facilitate object categorization: Evidence from 6-and 12-month-olds. *Cognition, 105*(1), 218–228.
Futó, J., Téglás, E., Csibra, G., & Gergely, G. (2010). Communicative function demonstration induces kind-based artifact representation in preverbal infants. *Cognition, 117*, 1–8.
Gerken, L., & Bollt, A. (2008). Three exemplars allow at least some linguistic generalizations: Implications for generalization mechanisms and constraints. *Language Learning and Development, 4*(3), 228–248.
Gordon, R. D., & Irwin, D. E. (1996). What's in an object file? Evidence from priming studies. *Perception & Psychophysics, 58*, 1260–1277.
Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research, 41*, 1409–1422.
Hauf, P., Paulus, M., & Baillargeon, R. (2012). Infants use compression information to infer objects' weights: Examining cognition, exploration, and prospective action in a preferential-reaching task. *Child Development, 83*, 1978–1995.
Havy, M., & Waxman, S. R. (2016). Naming influences 9-month-olds' identification of discrete categories along a perceptual continuum. *Cognition, 156*, 41–51.
Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological Methods, 3*(4), 486–504.
Henderson, A. M., & Woodward, A. L. (2012). Nine-month-old infants generalize object labels, but not object preferences across individuals. *Developmental Science, 15*(5), 641–652.
Hespos, S. J., & Baillargeon, R. (2001a). Infants' knowledge about occlusion and containment events: A surprising discrepancy. *Psychological Science, 12*, 141–147.
Hespos, S. J., & Baillargeon, R. (2001b). Reasoning about containment events in very young infants. *Cognition, 78*, 207–245.
Hespos, S. J., & Baillargeon, R. (2006). Décalage in infants' knowledge about occlusion and containment events: Converging evidence from action tasks. *Cognition, 99*, B31–B41.
Hespos, S. J., & Baillargeon, R. (2008). Young infants' actions reveal their developing knowledge of support variables: Converging evidence for violation-of-expectation findings. *Cognition, 107*, 304–316.
Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review, 98*, 352–376.
Huttenlocher, J., & Lourenco, S. F. (2007). Using spatial categories to reason about location. In J. Plumert, & J. Spencer (Eds.), *The emerging spatial mind* (pp. 3–24). New York, NY: Oxford University Press.
Jin, K. S., Houston, J. L., Baillargeon, R., Groh, A. M., & Roisman, G. I. (2018). Young infants expect an unfamiliar adult to comfort a crying baby: Evidence from a standard violation-of-expectation task and a novel infant-triggered-video task. *Cognitive Psychology, 102*, 1–20.
Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*, 175–219.
Káldy, Z., & Leslie, A. M. (2003). Identification of objects in 9-month-old infants: Integrating 'what' and 'where' information. *Developmental Science, 6*, 360–373.

Káldy, Z., & Leslie, A. M. (2005). A memory span of one? Object identification in 6.5-month-old infants. *Cognition, 97*, 153–177.

Kibbe, M. M., & Leslie, A. M. (2011). What do infants remember when they forget? Location and identity in 6-month-olds' memory for objects. *Psychological Science, 22*, 1500–1505.

Kibbe, M. M., & Leslie, A. M. (2019). Conceptually rich, perceptually sparse: Object representations in 6-month-old infants' working memory. *Psychological Science, 30*, 362–375.

Kotovsky, L., & Baillargeon, R. (1998). The development of calibration-based reasoning about collision events in young infants. *Cognition, 67*, 311–351.

LaTourette, A. S., & Waxman, S. R. (2020). Naming guides how 12-month-ol infants encode and remember objects. *Proceedings of the National Academy of Sciences, 117* (35), 21230–21234.

Leslie, A. M. (1995). A theory of agency. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal Cognition: A Multidisciplinary Debate* (pp. 121–149). Oxford, England: Oxford University Press.

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition, 25*(3), 265–288.

Leslie, A. M., Xu, F., Tremoulet, P. D., & Scholl, B. J. (1998). Indexing and the object concept: Developing "what" and "where" systems. *Trends in Cognitive Sciences, 2*, 10–18.

Levine, S. C., & Baillargeon, R. (2016). Different faces of language in numerical development: Exact number and individuation. In D. Barner, & A. S. Baron (Eds.), *Core knowledge and conceptual change* (pp. 127–150). New York, NY: Oxford University Press.

Lin, Y., & Baillargeon, R. (2018). Infants individuate objects with distinct prior event roles. Poster presented at the Biennial International Congress of Infant Studies, Philadelphia, PA.

Lin, Y., Stavans, M., & Baillargeon, R. (2020). Infants' physical reasoning and the cognitive architecture that supports it. In O. Houdé & G. Borst (Eds.), Cambridge Handbook of Cognitive Development. Cambridge, England: Cambridge University Press (in press).

Lin, Y., Stavans, M., & Baillargeon, R. (2019). Infants can use many types of categories to individuate objects. *Paper presented at the Biennial Meeting of the Society for Research in Child Development*.

Luo, Y., & Baillargeon, R. (2005). When the ordinary seems unexpected: Evidence for incremental physical knowledge in young infants. *Cognition, 95*, 297–328.

Luo, Y., Kaufman, L., & Baillargeon, R. (2009). Young infants' reasoning about physical events involving inert and self-propelled objects. *Cognitive Psychology, 58*, 441–486.

Margoni, F., Baillargeon, R., & Surian, L. (2018). Infants distinguish between leaders and bullies. *Proceedings of the National Academy of Sciences, 115*(38), E8835–E8843.

Mou, Y., & Luo, Y. (2017). Is it a container? Young infants' understanding of containment events. *Infancy, 22*, 256–270.

Needham, A., & Baillargeon, R. (1993). Intuitions about support in 4.5-month-old infants. *Cognition, 47*, 121–148.

Needham, A., & Baillargeon, R. (2000). Infants' use of featural and experiential information in segregating and individuating objects: A reply to Xu, Carey and Welch (2000). *Cognition, 74*(3), 255–284.

Needham, A., Dueker, G., & Lockhead, G. (2005). Infants' formation and use of categories to segregate objects. *Cognition, 94*(3), 215–240.

Newcombe, N., Huttenlocher, J., & Learmonth, A. (1999). Infants' coding of location in continuous space. *Infant Behavior and Development, 22*, 483–510.

Newcombe, N. S., Sluzenski, J., & Huttenlocher, J. (2005). Pre-existing knowledge versus on-line learning: What do infants really know about spatial location? *Psychological Science, 16*, 222–227.

Pauen, S. (2002). The global-to-basic level shift in infants' categorical thinking: First evidence from a longitudinal study. *International Journal of Behavioral Development, 26*, 492–499.

Piaget, J. (1954). *The construction of reality in the child*. New York, NY: Basic Books.

Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition, 32*, 65–97.

Pylyshyn, Z. W. (2007). *Things and places: How the mind connects with the world*. Cambridge, MA: MIT Press.

Quinn, P. C., & Bhatt, R. S. (2005). Learning perceptual organization in infancy. *Psychological Science, 16*(7), 511–515.

Rips, L. J., Blok, S., & Newman, G. (2006). Tracing the identity of objects. *Psychological Review, 113*, 1–30.

Schaub, S., Bertin, E., & Cacchione, T. (2013). Infants' individuation of rigid and plastic objects based on shape. *Infancy, 18*(4), 629–638.

Scott, R. M., & Baillargeon, R. (2009). Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development, 80*(4), 1172–1196.

Setoh, P., Wu, D., Baillargeon, R., & Gelman, R. (2013). Young infants have biological expectations about animals. *Proceedings of the National Academy of Sciences, USA, 110*, 15937–15942.

Smith, K., Mei, L., Yao, S., Wu, J., Spelke, E., Tenenbaum, J., & Ullman, T. (2019). Modeling expectation violation in intuitive physics with coarse probabilistic object representations. In H. Wallach, L. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems* (Vol. 32, pp. 8983–8993). Red Hook, NY: Curran Associates Inc.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review, 99*, 605–632.

Spelke, E. S., Phillips, A., & Woodward, A. L. (1995). Infants' knowledge of object motion and human action. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 44–78). Oxford, England: Oxford University Press.

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science, 348*(6230), 91–94.

Stavans, M., & Baillargeon, R. (2018). Four-month-old infants individuate and track simple tools following functional demonstrations. *Developmental Science, 21*, Article e12500.

Stavans, M., Lin, Y., Wu, D., & Baillargeon, R. (2019). Catastrophic individuation failures in infancy: A new model and predictions. *Psychological Review, 126*, 196–225.

Surian, L., & Caldi, S. (2010). Infants' individuation of agents and inert objects. *Developmental Science, 13*, 143–150.

Téglás, E., & Bonatti, L. L. (2016). Infants anticipate probabilistic but not deterministic outcomes. *Cognition, 157*, 227–236.

Téglás, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J. B., & Bonatti, L. L. (2011). Pure reasoning in 12-month-old infants as probabilistic inference. *Science, 332* (6033), 1054–1059.

Van de Walle, G. A., Carey, S., & Prevor, M. (2000). Bases for object individuation in infancy: Evidence from manual search. *Journal of Cognition and Development, 1*, 249–280.

Wang, S. (2011). Priming 4.5-month-old infants to use height information by enhancing retrieval. *Developmental Psychology, 47*, 26–38.

Wang, S. (2019). Regularity detection and explanation-based learning jointly support learning about physical events in early infancy. *Cognitive Psychology, 113*, Article 101219.

Wang, S., & Baillargeon, R. (2005). Inducing infants to detect a physical violation in a single trial. *Psychological Science, 16*, 542–549.

Wang, S., & Baillargeon, R. (2006). Infants' physical knowledge affects their change detection. *Developmental Science, 9*, 173–181.

Wang, S., & Baillargeon, R. (2008a). Can infants be "taught" to attend to a new physical variable in an event category? The case of height in covering events. *Cognitive Psychology, 56*, 284–326.

Wang, S., & Baillargeon, R. (2008b). Detecting impossible changes in infancy: A three-system account. *Trends in Cognitive Sciences, 12*, 17–23.

Wang, S., Baillargeon, R., & Brueckner, L. (2004). Young infants' reasoning about hidden objects: Evidence from violation-of-expectation tasks with test trials only. *Cognition, 93*, 167–198.

Wang, S., Baillargeon, R., & Paterson, S. (2005). Detecting continuity violations in infancy: A new account and new evidence from covering and tube events. *Cognition, 95*, 129–173.

Wang, S., & Goldman, E. J. (2016). Infants actively construct and update their representations of physical events: Evidence from change detection by 12-month-olds. *Child Development Research*. article 3102481.

Wang, S., & Kohne, L. (2007). Visual experience enhances infants' use of task-relevant information in an action task. *Developmental Psychology, 43*, 1513–1522.

Wang, S., & Mitroff, S. R. (2009). Preserved visual representations despite change blindness in infants. *Developmental Science, 12*, 681–687.

Wang, S., & Onishi, K. H. (2017). Enhancing young infants' representations of physical events through improved retrieval (not encoding) of information. *Journal of Cognition and Development, 18*, 289–308.
Wang, S., Zhang, Y., & Baillargeon, R. (2016). Young infants view physically possible support events as unexpected: New evidence for rule learning. *Cognition, 157*, 100–105.
Waxman, S. R., & Booth, A. E. (2001). Seeing pink elephants: Fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology, 43*(3), 217–242.
Waxman, S. R., & Braun, I. (2005). Consistent (but not variable) names as invitations to form object categories: New evidence from 12-month-old infants. *Cognition, 95*(3), B59–B68.
Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12-to 13-month-old infants. *Cognitive Psychology, 29*, 257–302.
Wilcox, T. (1999). Object individuation: Infants' use of shape, size, pattern, and color. *Cognition, 72*, 125–166.
Wilcox, T., & Baillargeon, R. (1998). Object individuation in infancy: The use of featural information in reasoning about occlusion events. *Cognitive Psychology, 37*, 97–155.
Wilcox, T., & Chapa, C. (2002). Infants' reasoning about opaque and transparent occluders in an individuation task. *Cognition, 85*(1), B1–B10.
Wilcox, T., & Chapa, C. (2004). Priming infants to attend to color and pattern information in an individuation task. *Cognition, 90*, 265–302.
Wilcox, T., Nadel, L., & Rosser, R. (1996). Location memory in healthy preterm and full-term infants. *Infant Behavior and Development, 19*, 309–323.
Wilcox, T., & Schweinle, A. (2003). Infants' use of speed information to individuate objects in occlusion events. *Infant Behavior and Development, 26*(2), 253–282.
Wilcox, T., Schweinle, A., & Chapa, C. (2003). Object individuation in infancy. In F. Fagan, & H. Hayne (Eds.), *Progress in Infancy Research* (Vol. 3, pp. 193–243). Mahwah, NJ: Lawrence Erlbaum Associates.
Wilcox, T., Smith, T., & Woods, R. (2011). Priming infants to use pattern information in an object individuation task: The role of comparison. *Developmental Psychology, 47*, 886.
Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition, 85*, 223–250.
Xu, F. (2007). Sortal concepts, object individuation, and language. *Trends in Cognitive Sciences, 11*(9), 400–406.
Xu, F. (2019). Towards a rational constructivist theory of cognitive development. *Psychological Review, 126*(6), 841–864.
Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology, 30*, 111–153.
Xu, F., & Carey, S. (2000). The emergence of kind concepts: A rejoinder to Needham and Baillargeon (2000). *Cognition, 74*(3), 285–301.
Xu, F., Carey, S., & Quint, N. (2004). The emergence of kind-based object individuation in infancy. *Cognitive Psychology, 49*, 155–190.
Xu, F., Cote, M., & Baker, A. (2005). Labeling guides object individuation in 12-month-old infants. *Psychological Science, 16*(5), 372–377.
Xu, F., & Denison, S. (2009). Statistical inference and sensitivity to sampling in 11-month-old infants. *Cognition, 112*, 97–104.
Xu, F., & Tenenbaum, J. B. (2007). Sensitivity to sampling in Bayesian word learning. *Developmental Science, 10*(3), 288–297.
Yin, J., & Csibra, G. (2015). Concept-based word learning in human infants. *Psychological Science, 26*(8), 1316–1324.
Zosh, J. M., & Feigenson, L. (2015). Array heterogeneity prevents catastrophic forgetting in infants. *Cognition, 136*, 365–380.