Object Individuation and Physical Reasoning in Infancy:

An Integrative Account

Renée Baillargeon

Maayan Stavans

Di Wu

Yael Gertner

Peipei Setoh

Audrey K. Kittredge

Amélie Bernard

University of Illinois

Correspondence concerning this article should be sent to: Renée Baillargeon, Department of Psychology, University of Illinois, 603 East Daniel, Champaign, IL 61820, USA (email: rbaillar@illinois.edu).

Abstract

Much of the research on object individuation in infancy has used a task in which two different objects emerge in alternation from behind a large screen, which is then removed to reveal either one or two objects. In their seminal work, Xu and Carey (1996) found that it is typically not until the end of the first year that infants detect a violation when a single object is revealed. Since then, a large number of investigations have modified the standard task in various ways and found that young infants succeed with some but not with other modifications, yielding a complex and unwieldy picture. In this article, we argue that this confusing picture can be better understood by bringing to bear insights from a related subfield of infancy research, physical reasoning. By considering how infants reason about object information within and across physical events, we can make sense of apparently inconsistent findings from different object-individuation tasks. In turn, object-individuation findings deepen our understanding of how physical reasoning develops in infancy. Integrating the insights from physical-reasoning and object-individuation investigations thus enriches both subfields and brings about a clearer account of how infants represent objects and events.

Object individuation refers to the ability to determine how many objects are present in a scene. Beginning with the seminal work of Xu and Carey (1996), much of the research on object individuation in infancy has used the following task: one object emerges on one side of a large screen and returns behind it, and then a different object emerges on the opposite side of the screen and again returns behind it; this event sequence is repeated several times, and finally the screen is removed to reveal either both objects (*two-object* event) or only one of the objects (*one-object* event). When this task was first introduced, its rationale was that if infants could determine that two objects were present behind the screen, they would succeed at detecting the violation in the one-object event. As we shall see, this rationale has turned out to be incorrect: infants in the first year of life often realize that two objects are present behind the screen and yet fail to detect the violation in the one-object event.

Over the past 15 years, a large number of investigations have introduced variations in the original task of Xu and Carey (1996) and demonstrated that infants succeed with some but not with other variations, yielding a complex and unwieldy picture. In this article, we argue that this confusing picture can be better understood by bringing to bear insights from a related subfield of infancy research, physical reasoning. By considering how infants represent object information within and across physical events, we can make sense of apparently inconsistent findings from different object-individuation tasks. In turn, object-individuation findings deepen our understanding of how physical reasoning develops in infancy. Integrating the insights from physical-reasoning and object-individuation investigations thus enriches both subfields and brings about a clearer account of how infants represent objects and events.

In what follows, we first discuss physical reasoning in infancy and then examine how this research sheds light on infants' successes and failures in various object-individuation tasks.

I. Physical Reasoning in Infants

Our account of physical reasoning in infancy (e.g., Baillargeon, Li, Gertner, & Wu, 2011; Baillargeon, Li, Ng, & Yuan, 2009a; Baillargeon, Wu, Yuan, Li, & Luo, 2009b; Wang & Baillargeon, 2008b) builds on work by numerous researchers in the infant and adult cognition literatures. According to this account, infants are born with a *physical-reasoning* (PR) system—an abstract computational system that provides them with a skeletal causal framework for reasoning and learning about the physical interactions of objects and other physical entities (e.g., Baillargeon et al., 2009a; Gelman, 1990; Leslie, 1995; Spelke, Breinlinger, Macomber, & Jacobson, 1992). It should be understood from the outset that the PR system operates without conscious awareness: infants are no more aware of the causal framework they use to reason about physical events than they are aware of the grammar they use to comprehend sentences (e.g., Fisher, Gertner, Scott, & Yuan, 2010; Yuan & Fisher, 2009).

The PR system's initial causal framework includes several core principles and concepts and is gradually elaborated as infants acquire specific causal rules for predicting and interpreting the outcomes of physical events (e.g., Baillargeon & Carey, in press; Leslie, 1995; Luo, Kaufman, & Baillargeon, 2009; Spelke et al., 1992). Of most relevance to the research reviewed in this article is the core principle of *persistence*, which states that, all other things being equal, objects persist, as they are, in time and space (e.g., Baillargeon, 2008; Baillargeon et al., 2009a). The persistence principle has multiple corollaries: an object cannot spontaneously appear or disappear (*continuity*), occupy the same space as another object (*solidity*), break apart (*cohesion*), fuse with another object (*boundedness*), or change size, shape, pattern, or color (*unchangeableness*) (e.g., Baillargeon, 2008; Baillargeon et al., 2009a; Spelke et al., 1992; Spelke, Phillips, & Woodward, 1995b).

A. How do infants learn about events?

When infants observe a physical event, the PR system builds a specialized *physical representation* of the event. Though initially very sparse, physical representations become progressively more precise and detailed, like rough blueprints that are gradually filled in, as infants learn what information is helpful for predicting and interpreting outcomes. Because infants' physical knowledge critically affects the content of their physical representations, we begin with a discussion of how infants learn about events.

1. Event categories

In the first weeks of life, infants begin to identify *event categories* that correspond to distinct kinds of causal interactions between objects (e.g., Casasola, Cohen, & Chiarello, 2003; Hespos & Baillargeon, 2001a; McDonough, Choi, & Mandler, 2003; Quinn, 2007). Early event categories (all familiar to our evolutionary ancestors) include *occlusion* (i.e., events in which an object occludes another object), *support* (i.e., events in which an object supports another object), *collision* (i.e., events in which an object hits another object), and *containment* (i.e., events in which a container contains an object). In each event category, objects play specific roles: an occlusion event will involve an 'occluder' and an 'occludee', a support event will involve a 'support' and a 'supportee', a collision event will involve a ‘hitter’ and a ‘hittee’, and so on.

The process by which infants identify event categories is one of *explanation-based learning* (EBL) (e.g., Baillargeon, 2002; Bernard, Baillargeon, & DeJong, 2011; DeJong, 1997; Wang & Baillargeon, 2008a). EBL has both empirical (statistical) and analytical (explanatory) components and typically involves four main steps. First, learning is triggered when infants build similar physical representations for two or more events—and notice that the events have *contrastive outcomes*. Second, infants search for the *conditions* that map onto these outcomes until they detect a possible condition-outcome regularity (we assume that infants’ impressive statistical-learning ability plays a key role in this process; e.g., Fiser & Aslin, 2002; Saffran, 2009). Third, infants construct an *explanation* for the observed regularity, using their prior causal knowledge; although shallow and abstract (e.g., Baillargeon et al., 2009b; Keil, 1995; Luo et al., 2009; Wilson & Keil, 2000), this explanation nevertheless helps them make sense of the regularity. Finally, the explanation suggests a general *causal rule* that describes a type of causal interaction in which objects play specific roles—in other words, an event category.

To illustrate the EBL process, consider how infants identify the category of *occlusion* events, prior to 2.5 months of age (e.g., Aguiar & Baillargeon, 1999; Luo & Baillargeon, 2005). To start, infants notice that the objects they are tracking sometimes remain continuously visible and sometimes do not. Infants then search for the conditions associated with these contrastive outcomes and eventually notice that the objects cease to be visible when they move *behind* other objects and remain visible otherwise. Next, infants engage in a causal analysis of this regularity: by bringing to bear their prior causal knowledge (and especially the persistence principle), infants realize that the objects cease to be visible when the nearer objects hide them from view. This explanation leads to a general *occlusion* rule that describes the causal interaction between *occluders* and *occludees*: specifically, when an occludee is behind an occluder, the occluder hides the occludee from view. From that point on, whenever infants see an event in which an object moves or is placed behind another object (e.g., a parent steps behind a door, a ball is pushed behind a doll, or a bowl is lowered behind a cereal box), the PR system categorizes the event as an occlusion event and assigns appropriate roles to the objects in the event.[1]

2. Variables

In most cases, the initial causal rule that is generated for an event category (e.g., “occludees are hidden when behind occluders”, “supportees are stable when released on top of supports”, and “hittees are displaced when hit by hitters”) is only a coarse, all-or-none rule that leads to incorrect as well as correct predictions (e.g., Luo & Baillargeon, 2005). As infants come across some of these incorrect predictions, they revise their flawed rule through the identification of relevant *variables* (e.g., Baillargeon, Needham, & DeVos, 1992; Hespos & Baillargeon, 2008; Kotovsky & Baillargeon, 1998; Wang, Baillargeon, & Paterson, 2005). A variable both calls attention to a certain type of information in an event (e.g., features of objects or their arrangements) and provides a causal rule for interpreting that information. Like event categories, variables are identified through EBL.

For example, consider infants’ identification of the occlusion variable *lower-edge-discontinuity*, at about 3 months of age (see Figure 1A; e.g., Aguiar & Baillargeon, 2002; Luo & Baillargeon, 2005). To start, infants notice that, when an occludee passes behind an occluder, sometimes the occludee remains continuously hidden, as expected, but sometimes it becomes temporarily visible beneath the occluder, contrary to expectation (e.g., Luo & Baillargeon, 2005). Infants then search for the conditions that map onto these outcomes; eventually, they notice that the occludee becomes temporarily visible when it passes behind an occluder whose lower edge is not continuous with the surface on which it rests (e.g., like an inverted-U-shaped screen), creating a gap between the occluder and the surface. Next, infants build an explanation for this observed condition-outcome regularity (the persistence principle dictates that an occludee must become visible when passing behind a gap beneath an occluder). Finally, a revised causal rule is generated: an occludee will remain continuously hidden when passing behind an occluder as long as no gap exists between the bottom of the occluder and the surface on which it rests. From that point on, whenever infants see an occlusion event, the PR system includes information about the occluder’s lower edge in the physical representation of the event.

A few weeks later, at about 3.5 months of age, infants identify another occlusion variable, *height* (see Figure 1A; e.g., Baillargeon & DeVos, 1991; Hespos & Baillargeon, 2001a). To start, infants notice that, when an occludee is behind an occluder whose lower edge is continuous with the surface on which it rests, sometimes the occludee is hidden, as expected, but sometimes it remains partly visible above the occluder, contrary to expectation (e.g., Luo & Baillargeon, 2005). Infants then search for the conditions that map onto these outcomes and eventually notice that occludees remain visible when they are taller, but not shorter, than the occluders. Next, infants build an explanation for this observed condition-outcome regularity (the persistence principle dictates that a tall occludee must protrude above a short occluder). Finally, a revised causal rule is generated: the height of an occludee relative to that of an occluder determines whether the occludee will be fully or only partly hidden when behind the occluder. From that point on, whenever infants see an occlusion event, the PR system also includes information about the heights of the occluder and occludee in the physical representation of the event.

In addition to lower-edge-discontinuity and height, infants identify several other variables during the first year that help them better predict whether occludees will remain fully and continuously hidden when behind occluders. These variables include *width*, which is identified at about 3.5 or 4 months (Wang, Baillargeon, & Brueckner, 2004), and *occluder transparency*, which is identified at about 7.5 months (Luo & Baillargeon, 2011). As shown in Figure 1A, each new variable revises predictions from earlier variables, as in a decision tree. We refer to such decision trees as *vectors.* A vector in an event category focuses on a specific question pertinent to the category (e.g., will the occludee be fully and continuously hidden behind the occluder, or will it not?) and specifies under which conditions each of two contrastive outcomes (e.g., hidden, visible) is likely to obtain. With each new variable that is added along the vector—or with each new partition in the decision tree—infants' predictions slowly begin to approximate those of older children and adults (Luo & Baillargeon, 2005).

3. Additional vectors

As infants experience events in a category, they will often notice contrastive outcomes having to do with other facets of the events, and this will lead to the formation of additional vectors for the category. For example, in the case of occlusion events, infants must learn to predict not only whether the occludee will remain fully and continuously hidden when behind the occluder, as discussed above, but also where and when the occludee will reappear from behind the occluder, and whether the occludee that reappears from behind the occluder is the same occludee that disappeared (e.g., Aguiar & Baillargeon, 2002; Baillargeon & DeVos, 1991; Kochukhova & Gredebäck, 2007; von Hofsten, Kochukhova, & Rosander, 2007; Wilcox, 1999; Wilcox & Schweinle, 2003). Similarly, in the case of containment events, infants must learn to predict whether the containee will fit through the opening of the container, whether the containee will protrude above the container, whether the containee will remain visible through the sidewalls of the container, and whether the containee that is removed from the container is the same containee that was placed inside it (e.g., Hespos & Baillargeon, 2001a; Luo & Baillargeon, 2011; Ng & Baillargeon, in Baillargeon et al., 2009a; Wang et al., 2004). In order to fully predict outcomes within an event category, infants must identify the variable(s) relevant to each category vector.

In addition to the occlusion vector in Figure 1A, discussed in the last section, consider as another example of an occlusion vector some of the variables infants identify to predict whether the occludee that reappears from behind the occluder is the same occludee that disappeared (see Figure 1B). During the first year, infants identify several variables along this vector, including shape at about 4.5 months, pattern at about 7.5 months, and color at about 11.5 months (e.g., Wilcox, 1999; Wilcox & Chapa, 2004). Note that none of these variables would be helpful in predicting whether the occludee should be fully or continuously hidden when behind the occluder, or how soon the occludee should reappear from the occluder; variables are often (but not always) relevant to a single category vector.

Across event categories, vectors and variables can be organized quite differently—it all depends on what the variables (and their associated causal rules) are used to predict. For example, in the case of occlusion events, width and height are attached to the same vector—both are helpful for predicting whether the occludee will be fully or only partly hidden when behind the occluder—and they are identified at about the same age, 3.5 to 4 months (width and height may actually be identified as a single occlusion variable, *size*; e.g., Baillargeon & DeVos, 1991; Wang et al., 2004). In the case of containment events, however, width and height are attached to different vectors and are identified at different ages: width (identified at about 4 months) predicts whether the containee will fit through the opening of the container, whereas height (identified several months later, at about 7.5 months) predicts whether the containee will protrude above the container (e.g., Hespos & Baillargeon, 2001a; Wang et al., 2004). Mirroring these infant data, recent experiments indicate that when adults are presented with challenging displays depicting multiple occlusion or containment events, they are equally good at detecting width and height changes in occlusion events, but they are better at detecting width as opposed to height changes in containment events (Strickland & Scholl, 2011). Some aspects of infants' physical knowledge—having to do with the specific event categories, vectors, and variables they create—thus seem to persist into adulthood.

4. Décalages

The process by which infants learn about early event categories has at least two adaptive advantages. First, infants are unlikely to acquire spurious variables (e.g., "occludees become fully hidden when behind green occluders, but remain partly visible when behind red occluders"), because only condition-outcome regularities for which the PR system can build explanations become variables (e.g., Baillargeon, 2002; Wang & Baillargeon, 2008a). Second, any new variable in an event category is immediately, and appropriately, generalized to all events within the category, because the variable is broadly framed to apply to any objects in the category roles (e.g., "the height of an occludee relative to that of an occluder determines whether the occludee will be fully or only partly hidden when behind the occluder") (e.g., Setoh & Baillargeon, 2011; Wang & Baillargeon, 2008a).

Despite its impressive advantages, however, the process by which infants acquire their early physical knowledge has one important limitation: because variables are not generalized or transferred across event categories, even when equally relevant, striking lags or *décalages* occasionally arise in infants' responses to similar events from different categories (e.g., Hespos & Baillargeon, 2001a, 2006; Wang et al., 2005; Wang & Baillargeon, 2006). According to the EBL account, in order to identify a variable in an event category, infants must be exposed to appropriate contrastive observations from which to abstract the variable. Thus, whether infants successfully reason about the same variable in different event categories depends on whether they have had the opportunity to identify the variable in each category. In some cases, infants acquire the same variable at about the same age in different categories (e.g., Wang et al., 2004). In other cases, however, weeks or months separate the acquisition of the same variable in different categories, resulting in marked décalages in infants' responses (e.g., Hespos & Baillargeon, 2006; Wang & Baillargeon, 2006). For example, although the variable height is identified at about 3.5 months in occlusion events, as we saw earlier (e.g., Baillargeon & DeVos, 1991), it is not identified until about 7.5 months in containment events (e.g., Hespos & Baillargeon, 2001a), until about 12 months in covering events (e.g., Wang et al., 2005), and until about 14.5 months in tube events (e.g., Wang et al., 2005) (containment, covering, and tube events all constitute different event categories for infants, because they represent different kinds of causal interactions between objects). Could these décalages be reversed, so that infants identify the variable height in tube events, say, before they do so in containment or covering events? Prior findings suggest that the answer to this question is yes: a number of 'teaching' experiments in which infants were exposed to contrastive outcomes consistent with EBL have confirmed that the age at which infants identify a variable in different event categories depends primarily on experience (e.g., Baillargeon, 2002; Setoh & Baillargeon, 2011; Wang & Baillargeon, 2008a; Wang & Kohne, 2007). Each event category functions as a mini domain, and infants' knowledge about each domain is closely tailored to their daily experiences, making them well adapted to their physical environment.

5. Learning what as opposed to learning how

Implicit in our account of how infants learn about early event categories is one fundamental and arresting generalization: infants learn *what information to add* to their initial, sparse physical representations—not *how to reason about* that information. The PR system's skeletal causal framework can immediately interpret most of the variable information infants learn to attend to during the first year of life. One way to think about early physical reasoning is that it involves the interaction of two distinct processes: an *interpretive* process (the PR system's causal framework), which interprets any information that happens to be included in infants' physical representations, and a *predictive* process (the EBL mechanism), which serves to enrich infants' initial, sparse physical representations through the addition of new information, variable by variable, vector by vector, and category by category. In the first year of life, the causal rules associated with variables are all implications of the skeletal causal framework (e.g., the rules associated with the variables height, width, shape, pattern, and color in occlusion events all directly follow from the persistence principle), so that the focus of learning is on what information should be included in physical representations, rather than how to interpret that information once represented.

The fact that the PR system's skeletal causal framework can immediately interpret most, if not all, of the variable information included in physical representations of early event categories leads to a striking prediction: infants who have not yet identified a variable as relevant for an event category, but are temporarily induced through priming or other contextual manipulations to include information about the variable in their physical representation of an event, should immediately succeed at violation-of-expectation and preferential-reaching tasks involving the variable. Recent experiments support this prediction (e.g., Baillargeon et al., 2009a; Gertner, Baillargeon, Fisher, & Simons, 2009; Li & Baillargeon, 2011; Wang & Baillargeon, 2005); we return to this point later when discussing positive carryover effects.

B. How do infants represent a physical event?

We have seen that when infants observe a physical event, the PR system builds a physical representation of the event. This representation typically has two layers, a *structural* and a *variable* layer. The structural layer includes generic information that is encoded about any event, whereas the variable layer includes more specific information that differs across events. Each layer is discussed in turn below.

1. The structural layer

Spatiotemporal and categorical information. The structural layer in the physical representation of an event includes *spatiotemporal* and *categorical* information. The spatiotemporal information describes the arrangement of the objects in the event and specifies how this arrangement changes over the course of the event. Though initially rather coarse, the spatiotemporal information infants represent about events becomes gradually more precise with the development of perceptual capacities (e.g., Valenza & Bulf, 2011; Yonas & Granrud, 1984). The categorical information specifies what kinds of objects are involved in the event: categorical descriptors are provided for each object and are updated as necessary as the event unfolds. From a very early age, infants spontaneously categorize objects along a number of abstract dimensions (contrary to claims that they possess a single initial category of physical objects, sometimes referred to as 'Spelke' objects; e.g., Xu & Carey, 1996; Xu, 2007). Early categorical descriptors include *ontological* descriptors such as whether objects are inert or self-propelled (e.g., Luo et al., 2009; Saxe, Tzlenic, & Carey, 2007) and non-agentive or agentive (e.g., Csibra, 2008; Johnson, Shimizu, & Ok, 2007), as well as *functional* descriptors such as whether objects are closed, open at the top to form containers, open at the bottom to form covers, or open at both ends to form tubes (e.g., Hespos & Baillargeon, 2001b; Wang et al., 2005). As we shall see, findings from the object-individuation literature have helped shed further light on the categorical descriptors in the structural layer; specifically, these findings indicate that, in the first year of life, categorical descriptors (1) also include the ontological categories humanlike versus non-humanlike objects (e.g., Bonatti, Frot, Zangl, & Mehler, 2002; Bonatti, Frot, & Mehler, 2005); (2) also include functional categories such as objects that rattle as opposed to jingle when shaken (e.g., Mareschal & Johnson, 2003; Wilcox, Woods, Tuggy, & Napoli, 2006); but (3) typically do not include specific object or *taxonomic* categories (e.g., such as whether an object is a ball or a toy duck; e.g., Xu & Carey, 1996; Rivera & Zawaydeh, 2006).

Both the spatiotemporal and the categorical (i.e., ontological, functional, and taxonomic) information about an event can help specify how many objects are involved in the event.[2] Although infants can approximately represent large sets of objects (e.g., McCrink & Wynn, 2004; Xu & Spelke, 2000), they can precisely track only about three objects per event or per set within an event (e.g., Cheries, Wynn, & Scholl, 2006; Feigenson & Carey, 2005; Feigenson & Halberda, 2004; Leslie & Chen, 2007).

Event categories and category roles. As infants identify event categories, changes take place in the structural layer of physical representations. When infants encounter an event from a known category, the PR system categorizes the event and assigns appropriate roles to the objects in the event. After repeatedly watching an event in which two identical objects play distinct roles (e.g., object-A is the occluder and object-B is the occludee), infants dishabituate if the two objects exchange roles, even though the outcomes of the events are perceptually identical (e.g., if object-B becomes the occluder and object-A the occludee) (e.g., Leslie & Keeble, 1987; Onishi, 2011).

An example. Consider the situation depicted in Figure 2A: 10-month-old infants see three objects standing apart on an apparatus floor; from left to right, these are a yellow toy duck, a red block decorated with white dots, and a large green screen. After a pause, the duck and block move behind the screen. What information would the PR system include in the structural layer of the event's physical representation (see Figure 2B)? The spatiotemporal information at the start of the event specifies that three objects are present: the duck, block, and screen all occupy separate locations in space. The categorical descriptors for the duck and block (which move on their own) are self-propelled, non-agentive, and closed, whereas those for the screen are inert, non-agentive, and closed (in general, novel objects are categorized as inert and non-agentive by default, until they provide unambiguous evidence to the contrary; e.g., Johnson et al., 2007; Luo et al., 2009). When the duck and block move behind the screen, the event is categorized as an occlusion event, with the screen as the occluder and the duck and block as the occludees.

Taken as a whole, the information in the structural layer of our event's physical representation captures the essence of the event: it specifies *how many objects* are involved in the event (three), *what kinds of objects* they are (self-propelled, non-agentive, and closed for the duck and block; inert, non-agentive, and closed for the screen), *what kind of causal interaction* the objects are engaged in (an occlusion event), and *what role* each object plays in the event (occluder for the screen; occludees for the duck and block). Although substantial, the information in the structural layer of our event's physical representation is still limited: in particular, it says nothing about the height, width, shape, pattern, and color of the objects in the event. This kind of information will be included in the variable layer of the physical representation, as explained below.

2. The variable layer

We saw earlier that, for each event category, infants identify variables (organized in vectors) that help them better predict and interpret events' outcomes. With the identification of variables, physical representations become progressively richer and more detailed. When infants watch an event, the PR system first represents the structural information about the event and uses this information to categorize it. Next, the PR system accesses its knowledge of the event category selected, and more specifically taps the list of variables that have been identified as relevant for predicting outcomes within the category. The PR system then gathers information about each of these variables and includes this information in the variable layer of the event's physical representation.[3] Information about variables not yet identified is typically not included in the physical representation. Thus, the décalages in infants' identification of variables across event categories mean that, at a given age, information about a variable (e.g., height) may be included in the physical representation of one event (e.g., an occlusion event) but not another event (e.g., a containment event), even though this information is causally relevant for both events.

Let us now return to our event and consider what variable information 10-month-olds would include in their physical representation of the event (see Figure 2B). Because by this age most infants have identified height, width, shape, and pattern as occlusion variables, information about all of these variables would be included in the variable layer of the physical representation. In contrast, color information would not be included, because at this age most infants have not yet identified color as an occlusion variable.

C. How do infants represent a sequence of two physical events?

Our discussion so far has focused on how infants represent a *single* physical event. In everyday life, infants observe some isolated events, but they also observe many *sequences* of events. What happens when the same objects are involved first in an event from one event category and then in an event from a different event category (e.g., when a spoon is placed first *behind* and then *inside* a bowl)? Does the PR system build an entirely separate physical representation for each event? This would seem disadvantageous: because the PR system must operate rapidly (to keep up with events as they unfold in the world), *carrying over* the object information from the first physical representation to the second physical representation would be far more efficient.

Recent research indicates that infants do, indeed, carry over variable information across physical representations: when the same objects are involved in a sequence of two events from different event categories, the PR system carries over the variable information from the first to the second physical representation (e.g., Wang & Baillargeon, 2005; Li, Sigmon-Hernandez, & Baillargeon, 2010). No carryover occurs when different objects are involved in the two events (even if the objects in the two events are perceptually identical; Li et al., 2010). The evidence for the carryover of variable information includes both *positive* and *negative* carryover effects. Positive carryover effects occur when infants see an event in which a variable *has* been identified followed by an event in which the same variable has *not* yet been identified: because information about the variable is carried over from the first to the second physical representation (where it can immediately be interpreted by the PR system's causal framework), infants succeed in detecting violations in the second event that they would have failed to detect had they seen only the second event. Conversely, negative carryover effects occur when infants see an event in which a variable has *not* yet been identified followed by an event in which the same variable *has* been identified: because information about the variable is not included in the first physical representation, it cannot be carried over to the second physical representation; as a result, infants fail to detect violations in the second event that they would have readily detected had they seen the second event by itself.

One way to explain the carryover of object information from one event to the next might be in terms of a *binding* process (e.g., Leslie et al., 1998). As infants watch the event in Figure 2A, the PR system binds together, for each object, the information from the structural layer (e.g., categorical descriptors such as self-propelled, non-agentive, and closed) and from the variable layer (e.g., height, width, and shape information). Each object representation thus constitutes a tight object bundle that can move as a whole from event to event.[4]

Let us assume that an experimenter's hand now removes the screen from the apparatus (see Figure 3). As the screen is removed, the PR system begins a new physical representation: because the event is no longer an occlusion event (the screen is no longer an occluder, and the duck and block are no longer occludees), a new physical representation is required. The object representations for the duck and block—each with its structural and variable information neatly bundled together—are carried over from the first to the second physical representation, and guide infants' expectations about what objects should be present in this new, post-occlusion event.

D. How do infants reason about physical events?

In the preceding section, we focused on the question of what information infants represent when watching simple physical events. We have seen that the physical representation the PR system builds for an event has two layers: a structural layer, which includes spatiotemporal and categorical information about the event, and a variable layer, which includes more detailed information about the features and arrangement of the objects in the event. But how do infants *reason* about this information, once represented? We suggested earlier that any information that is included in a physical representation (either directly or through carryover and other contextual manipulations) becomes subject to the PR system's causal framework, which includes the persistence principle. Let us return to the event in Figure 3 and consider what persistence violations infants should detect when the screen is removed; we consider first structural and then variable persistence violations.

1. Detecting structural persistence violations

Given the structural object information carried over from the first to the second physical representation in Figure 3, infants should expect to see *two* objects when the screen is removed, and they should therefore detect a violation if shown only one object (infants typically look reliably longer at outcomes that violate, as opposed to confirm, their expectations). Such an outcome constitutes a persistence and more specifically a continuity violation: objects cannot magically disappear. Infants should also expect to see two *closed* objects when the screen is removed, and they should thus look reliably longer if shown two open objects. Such an outcome constitutes a persistence and more specifically an unchangeableness violation: objects cannot magically change their physical properties.

Prior evidence indicates that infants aged 2.5 months and older detect structural persistence violations in occlusion, containment, covering, and other events (for reviews, see Baillargeon et al., 2009a, 2011). For example, infants are surprised (as indexed by longer looking times) if objects magically appear or disappear (e.g., Cheries et al., 2006; Luo et al., 2009; Wang et al., 2005; Wynn, 1992), pass through other objects (e.g., Hespos & Baillargeon, 2001b; Luo, Baillargeon, Brueckner, & Munakata, 2003; Saxe, Tzelnic, & Carey, 2006; Spelke et al., 1992), and change from being closed to open or vice-versa (e.g., Wu, 2011).

2. Detecting variable persistence violations

Exactly what variable persistence violations infants detect in Figure 3 will depend on what variable information was included in the physical representation of the occlusion event and carried over to the new physical representation.

As we saw earlier, most 10-month-olds have identified height, width, shape, and pattern—but not color—as occlusion variables. Accordingly, 10-month-olds should detect a violation if the duck and block, when revealed, are much smaller than before, have changed shape, or have become striped; in contrast, infants should fail to detect a violation if the duck and block have changed color. All of these outcomes constitute persistence and more specifically unchangeableness violations: objects cannot magically change their properties.

Prior evidence indicates that infants detect a variable persistence violation in an event if they have identified the variable as relevant to the event's category (information about the variable is then included in the variable layer of the event's physical representation); conversely, infants fail to detect a variable persistence violation in an event if they have not yet identified the variable as relevant to the event's category (no information about the variable is then included in the event's physical representation) (for reviews, see Baillargeon et al., 2009a, 2011). For example, 4.5- to 6.5-month-olds detect a violation if an object changes shape when passing behind an occluder, but not if an object changes shape when briefly buried (e.g., Kittredge & Baillargeon, 2011; Newcombe, Huttenlocher, & Learmonth, 1999; Wilcox, 1999; Wilcox & Baillargeon, 1998b). Similarly, 8- to 11-month-olds detect a violation if a tall object changes height when briefly hidden inside a tall container, but not if an object changes height when briefly hidden under a tall cover or inside a tall tube (e.g., Li & Baillargeon, 2011; Wang & Baillargeon, 2006). Likewise, 11.5-month-olds detect a violation if an object changes color when passing behind an occluder, but not if an object changes color when briefly hidden inside a container (e.g., Ng & Baillargeon, in Baillargeon et al., 2009a; Wilcox, 1999). In each case, infants detect the violation if information about the variable is included in the physical representation of the event, but they fail to detect the violation otherwise.

3. Generating explanations for apparent persistence violations

Let us assume that the screen in our event is removed to reveal the duck, the block—and a toy car. Could infants generate a plausible explanation for this three-object event, by inferring that the car was already hidden behind the screen at the start of the event? Prior evidence indicates that when the number of objects revealed in an event is greater or smaller than expected, infants often posit hidden objects (e.g., Baillargeon, 1994; Spelke et al., 1995a) or hidden displacements (e.g., Luo et al., 2009; Wang et al., 2005) to make sense of the event, as long as these explanations are consistent with the physical layout in which the event takes place.

For example, to examine whether infants could posit a hidden object to make sense of an otherwise impossible event, Baillargeon, Miller, and Constantino (cited in Baillargeon, 1994) showed 10-month-olds a test event in which an experimenter's hand placed two Ernie dolls, one at a time, behind a large screen; the screen was then lowered to reveal three Ernie dolls. Infants showed no surprise at this three-doll outcome if the screen stood upright at the start of the event, because they could infer that a third Ernie doll had already been hidden behind the screen. However, infants did show surprise at the three-doll outcome if the screen lay flat on the apparatus floor, with no object behind it, at the start of the event, because they could then no longer make sense of the outcome: they realized that a third Ernie doll could not have magically appeared in the scene (e.g., Simon, Hespos, & Rochat, 1995; Wynn, 1992).

Similarly, to examine whether infants could posit a hidden displacement to make sense of an otherwise impossible event, Luo et al. (2009) first familiarized 6-month-olds with a self-propelled box and then showed them the following test event: a screen was first raised to hide the box and was then lowered to reveal no box. Infants showed no surprise at the box's disappearance when the box could have slipped, out of sight, behind a second screen standing nearby. Infants did show surprise at the box's disappearance, however, if the second screen stood farther away, so that the box could not have reached it without becoming temporarily visible.

Thus, although infants may expect to see a specific number of objects when a screen is removed (as they carry over the object information from the occlusion event to the new, post-occlusion event), additional reasoning processes may be brought to bear, post-hoc, if the number of objects revealed is greater or smaller than expected. Infants posit hidden objects and hidden displacements to resolve these discrepancies, as long as such explanations are plausible given the physical constraints of the scene.

II. Object Individuation in Infants

We now turn to the research on object individuation in infancy that has been carried out over the past 15 years. As we will see, by applying and extending the account of early physical reasoning presented in the first portion of this article, we can make sense of apparently inconsistent findings from different object-individuation tasks.

A. The seminal findings of Xu and her colleagues

In the event shown in Figure 2A, infants could use spatiotemporal information to determine that at least three objects were present: the duck, block, and screen could all be seen (at the same time) to occupy distinct locations in space. In a seminal series of experiments, Xu and her colleagues (e.g., Xu & Carey, 1996; Xu, Carey, & Quint, 2004) adapted a task devised by Spelke et al. (1995a) to examine whether 10- and 12-month-olds could use *categorical* or *featural* information, as opposed to spatiotemporal information, to determine how many objects were involved in an occlusion event.

1. Findings with 10-month-olds

In one experiment, Xu and Carey (1996) tested 10-month-olds in a *categorical/featural* condition. The infants were first given two pairs of baseline trials; in each pair, an experimenter's hand removed a screen to reveal either two objects (e.g., a toy truck and a toy camel) or only one of the objects (e.g., the truck). Next, the infants received two pairs of test trials; one pair involved a baby bottle decorated with small blue bears and a ball decorated with green and pink stripes; the other pair involved a red sippy cup and a yellow baby book. In each test trial, one object (e.g., the bottle) emerged from one side of a large green screen and then returned behind it; next, another object (e.g., the ball) emerged from the other side of the screen and again returned behind it. This event was repeated several times, and then the infants were shown a new event: the screen was removed to reveal either both objects (two-object event) or only one of the objects (one-object event). Infants looked reliably longer at the two- than at the one-object event in both the baseline and the test trials, and thus gave no evidence that they expected to see two objects when the screen was removed.

In another experiment (Xu & Carey, 1996), 10-month-olds were tested in a *single-object* condition similar to the categorical/featural condition with two exceptions: the baseline trials involved identical cups, and the same object (e.g., a ball) was seen on either side of the screen in the test trials. In this condition, both the one-object (e.g., one ball) and the two-object (e.g., two balls) test events were plausible: even if infants initially expected to see one object when the screen was removed, they could readily generate an explanation for the two-object event by inferring that an additional object had been hidden behind the screen (e.g., Baillargeon, 1994; Spelke et al., 1995a). In line with this analysis, Xu and Carey found that the infants tended to look equally at the one- and two-object events in the baseline trials (presumably because all of the trials involved identical cups), but showed a weak preference for the two-object event in the test trials.

Finally, in another experiment, infants were tested in a *spatiotemporal* condition similar to the categorical/featural condition except that the two objects used in each test trial were brought out *simultaneously* at the start of the trial, thus providing infants with unambiguous spatiotemporal information that at least two objects were present behind the screen. Infants in this condition looked reliably longer at the two- than at the one-object event in the baseline trials but looked about equally at the two events in the test trials, suggesting that that they expected to see two objects when the screen was removed and that this expectation served to reduce their baseline preference for two objects.

These results have been replicated, with minor variations, in multiple laboratories (e.g., Bonatti et al., 2002; Futó, Téglás, Csibra, & Gergely, 2010; Krøjgaard, 2000; Leslie et al., 1998; Rivera & Zawaydeh, 2006; Surian & Caldi, 2010; Wilcox & Baillargeon, 1998a; Wilcox & Chapa, 2002). For example, Krøjgaard (2000) found that 10-month-olds in the categorical/featural condition failed (i.e., merely showed their baseline preference for two objects) even when one of the two objects used in the test trials was a favorite object from their own homes. In other variations that did not elicit a baseline preference for two objects (e.g., Rivera & Zawaydeh, 2006; Surian & Caldi, 2010), 10-month-olds in the categorical/featural condition looked about equally at the one- and two-object test events. Together, these various results provide robust evidence that, when two distinct objects emerge in alternation from behind a screen, 10-month-olds have no clear expectation as to whether one or two objects should be revealed when the screen is removed.

2. Findings with 12-month-olds

In additional experiments, Xu and her colleagues (Xu & Carey, 1996; Xu et al., 2004) found that 12-month-olds tested in the categorical/featural condition looked reliably longer at the two- than at the one-object event in the baseline trials, but looked about equally at the two events in the test trials. This last result suggested that these older infants expected to see two objects when the screen was removed and that this expectation partly overcame their baseline preference for two objects.

Interestingly, this result was found only when the two objects used in the test trials belonged to different taxonomic (object) categories. Infants failed (i.e., merely showed their baseline preference for two objects) when the two objects belonged to the same taxonomic category, even if the objects differed in color (e.g., a ball with green and pink stripes and a ball with purple and orange stripes), in size (e.g., a small and a large red ball), or in color, size, and pattern (e.g., a large red ball covered with glitter and a small soccer ball with orange, green, and white hexagons). When the two objects used in the test trials belonged to the same taxonomic category, 12-month-olds succeeded (i.e., overcame their baseline preference for two objects) only if they were tested in a spatiotemporal condition and saw the two objects simultaneously, on either side of the screen, at the start of each trial. Kingo and Krøjgaard (2011) recently confirmed that, when tested in a categorical/featural condition, 12-month-olds succeeded when tested with two objects from different taxonomic categories (e.g., a crocodile and a bed, a giraffe and a cupboard), but not when tested with two objects from the same taxonomic category (e.g., a crocodile and a giraffe, or a bed and a cupboard).

B. A carryover account

The preceding findings with 10- and 12-month-olds are, at first blush, inconsistent with our discussion of physical reasoning in infancy. In particular, we saw earlier that (1) most 10-month-olds have identified height and width (henceforth size), shape, and pattern as occlusion variables, and most 12-month-olds have also identified color as an occlusion variable; (2) infants interpret events in accordance with a principle of persistence which states that objects cannot spontaneously change their physical properties; and (3) infants posit additional objects to make sense of events that would otherwise constitute persistence violations, as long as the physical constraints of the situation make such explanations plausible. In line with these findings, we would expect 10- and 12-month-olds who see two different objects emerge in alternation from behind a screen (1) to attend to each object's featural properties (i.e., size, shape, and pattern at 10 months, also color at 12 months); (2) to realize that an object cannot magically change its properties when passing behind a screen; and (3) to infer that two different objects are involved in the event, as long as the screen is large enough to hide multiple objects (e.g., Wilcox, 1999; Wilcox & Baillargeon, 1998b). But if infants believe that two objects are present behind the screen, why do they not expect to see two objects when the screen is removed? Why are infants not surprised when a single object is revealed? In order to make sense of these results, we need to posit *three assumptions* that extend the account of early physical reasoning presented earlier.

1. First assumption: The carryover of object information breaks down when inconsistencies exist between the structural and variable layers of an event's physical representation

In our discussion of physical reasoning, we saw that when infants see a sequence of two events involving the same objects, the PR system carries over the object information from the physical representation of the first event to that of the second event. The results of Xu and her colleagues suggest that this carryover of object information proceeds smoothly when the *same number* of objects is specified in the structural and variable layers of the first event's physical representation, but breaks down when *different numbers* of objects are specified in the two layers.

To see what we mean, consider first the responses of the 10-month-olds in the *spatiotemporal* condition (we use as our example an occlusion event involving a baby bottle and a ball; see Figure 4A). As the bottle and ball emerge first simultaneously and then in alternation from behind the screen, what object information does the PR system include in the structural and variable layers of the event's physical representation? With respect to the structural layer, the spatiotemporal information specifies that two objects are present, since the bottle and ball are seen to occupy distinct locations in space. Each object receives the same categorical descriptors, namely, self-propelled, non-agentive, and closed. The PR system categorizes the event as an occlusion event and accesses the list of variables that have been identified as relevant for this event category. Because by 10 months most infants have identified size, shape, and pattern as occlusion variables, information about these variables is then included in the variable layer and linked to each object representation in the structural layer. The object information in the structural and variable layers thus yields two separate object bundles that are carried over, when the screen is removed, to the next physical representation, leading infants to expect two objects.

Next, consider the responses of the 10-month-olds in the *single-object* condition (see Figure 4B). As the ball emerges alternately on the left and right sides of the screen, what object information does the PR system represent? With respect to the structural layer, the spatiotemporal information is *ambiguous*: it is insufficient to determine whether the same object is moving back and forth behind the screen, or whether different objects are emerging in alternation from behind the screen. By contrast, the categorical information suggests that *one object* is present, since the ball receives the same categorical descriptors—self-propelled, non-agentive, and closed—each time it comes into view. The PR system thus assumes that a single object is moving back and forth behind the screen. After the PR system categorizes the event as an occlusion event, it gathers size, shape, and pattern information about the object and binds this information to the object representation in the structural layer. Together, the object information in the structural and variable layers thus yields a single object bundle that is carried over, when the screen is removed, to the next physical representation, leading infants to expect one object.

Finally, and most crucially, consider the responses of the 10-month-olds in the *categorical/featural* condition (see Figure 4C). As the bottle and ball emerge in alternation from behind the screen, what object information does the PR system represent? With respect to the structural layer, the situation is the same as in the single-object condition: the spatiotemporal information is ambiguous and therefore of no use in determining whether one or more objects are involved in the event; and the categorical information suggests that only one object is present, since each object, as it comes into view, receives the same categorical descriptors, namely, self-propelled, non-agentive, and closed. As in the single-object condition, the object information in the structural layer thus suggests that *one object* is traveling back and forth behind the screen. The situation is very different in the variable layer, however: the variable information gathered by the PR system signals two different sizes, shapes, and patterns. According to the persistence principle, objects cannot spontaneously change their physical properties; since the screen is large enough to hide multiple objects, the plausible inference is that *two objects* are present behind the screen. Overall, the object information in the structural and variable layers of the occlusion event's physical representation is thus *inconsistent*: the information in the structural layer points to *a single object*, whereas the information in the variable layer points to *two objects*. As a result, when the screen is removed, the carryover process breaks down: *no object information is carried over* to the next physical representation, so infants have no clear expectation about how many objects they should see.

2. Second assumption: By the end of the first year, the categorical information in the structural layer begins to include taxonomic descriptors

In our discussion of early physical reasoning, we saw that the categorical information the PR system routinely includes in the structural layer of physical representations includes abstract categorical descriptors such as inert/self-propelled, agentive/non-agentive, and open/closed. The findings by Xu and her colleagues with 12-month-olds suggest that, by the end of the first year, this categorical information also includes more concrete *taxonomic* descriptors—object categories such as 'bottle', 'ball', 'cup', and 'book' (Xu & Carey, 1996; Xu et al., 2004).

Consider the responses of the 12-month-olds when tested in a categorical/featural condition contrasting two objects from *different object categories* (e.g., a baby bottle and a ball; see Figure 5A). Although the spatiotemporal information available is still ambiguous, the categorical information for each object is now distinct: one is self-propelled, non-agentive, closed, and a ‘bottle’, whereas the other one is self-propelled, non-agentive, closed, and a ‘ball’. As a result, *two objects* are specified in the structural layer of the event’s physical representation. The PR system categorizes the event as an occlusion event and accesses the list of variables that have been identified as relevant for this event category; because by 12 months most infants have identified size, shape, pattern, and color as occlusion variables, information about these variables is then included in the variable layer. Since the bottle and ball differ in size, shape, pattern, and color, the object information in the variable layer specifies that *two objects* are present. Overall, the object information in the structural and variable layers is thus consistent in suggesting that two objects are involved in the event. When the screen is removed, beginning a new event, the two object bundles are carried over to the next physical representation, leading infants to expect two objects.

When the objects in the occlusion event belong to the *same object category* and differ only in their featural properties (e.g., two balls that differ in size; see Figure 5B), the situation for the 12-month-olds is similar to that depicted in Figure 4C. In the structural layer, the spatiotemporal information is insufficient to determine whether one or more objects are present, and each object receives identical categorical descriptors, namely, self-propelled, non-agentive, closed, and a ‘ball’. Because the number of objects specified in the structural layer (one) is inconsistent with that specified in the variable layer (two), no carryover of object information occurs, so that infants have no clear expectation about how many objects should be present. Infants expect to see two objects when the screen is removed only if they see the two objects simultaneously during the trials, so that spatiotemporal information specifies that two objects are present (see Figure 5C).

Why is it that, beginning at the end of the first year, the categorical information in the structural layer of physical representations begins to include taxonomic descriptors such as ‘bottle’ and ‘ball’? One explanation for this developmental change, suggested by Xu and her colleagues (e.g., Xu, 2002; Xu & Carey, 1996; Xu, Cote, & Baker, 2005), has to do with the fact that, by their first birthday, infants begin to acquire labels for objects. Because a naming label such as “a ball” refers to an object category (e.g., Balaban & Waxman, 1997; Graham, Kilbreath, & Welder, 2004; Waxman, 1999; Welder & Graham, 2001), infants who are in the process of learning labels may begin to encode objects taxonomically even when no labels are used.

Part of the evidence that language plays a role in the developmental transition between 10 and 12 months of age is that, in the experiments of Xu and Carey (1996), infants in the categorical/featural condition performed better if they knew the labels for the four objects used in the two pairs of test trials (“bottle”, “ball”, “cup”, “book”). Parental reports indicated that the majority of 12-month-olds understood at least two of the four labels, whereas the majority of 10-month-olds did not. Moreover, those few 10-month-olds who knew two or more of the four labels performed like 12-month-olds (i.e., looked equally at the one- and two-object events). Rivera and Zawaydeh (2006) subsequently confirmed these findings: they found that 10- and 11-month-olds who were reported by their parents to know the labels for the two objects shown in the occlusion event performed reliably better than those who knew no label or only one of the labels.

Xu (2007) has suggested that “learning count nouns that map onto kinds of objects can play a causal role” in the formation of object categories (p. 403). We think it unlikely that labels are required for the *formation* of object categories, however, since even young infants form and use object categories (e.g., Eimas & Quinn, 1994; Needham, Cantlon, & Ormsbee Holley, 2006; Needham, Dueker, & Lockhead, 2005; Quinn, Eimas, & Rosenkrantz, 1993). Instead, as discussed later, our claims are that (1) although young infants who encounter an artifact in a non-functional context typically do not encode its object category, they can readily be primed to do so do by various non-linguistic manipulations, and (2) toward the end of the first year, perhaps as a result of language acquisition, infants begin to spontaneously encode artifacts taxonomically, even in non-functional contexts and without priming manipulations.

3. Third assumption: Reconciling an inconsistency between the structural and variable layers of an event’s physical representation requires creating distinct spatiotemporal descriptors.

We have argued that infants fail at object-individuation tasks when the information in the structural layer of the occlusion event’s physical representation suggests that *a single object* is present behind the screen, whereas the information in the variable layer suggests that in fact *two objects* are present. As a result of this inconsistency, when the screen is removed, no object information is carried over to the next physical representation, so that infants have no expectation about how many objects they should see. This carryover account naturally gives rise to the following questions: when such inconsistencies arise, does the PR system attempt to revise the object information in the structural layer to render the two layers consistent? If yes, what does this revision process entail, and under what conditions is it successful?

Research by Wilcox and her colleagues (e.g., Wilcox, 2003, 2007; Wilcox, Alexander, Wheeler, & Norvell, in press; Wilcox & Schweinle, 2002) suggests answers to these questions. Specifically, it appears that when the information in the structural layer of an occlusion event's physical representation points to one object, whereas the information in the variable layer points to two objects, the PR system attempts to reconcile this inconsistency by creating *distinct spatiotemporal descriptors* for the two objects in the structural layer. Recall that, in the structural layer, only categorical and spatiotemporal descriptors—or only 'what' and 'where' descriptors —can be used to keep track of objects (e.g., Leslie et al., 1998; Mareschal & Johnson, 2003). Thus, when two objects receive the same categorical ('what') descriptors in the structural layer, but are assigned distinct featural properties in the variable layer, the PR system has only one recourse to resolve this discrepancy: it must produce distinct spatiotemporal ('where') descriptors for the two objects in the structural layer, by specifying each object's individual trajectory. When the 'what' descriptors cannot be of help (because they are identical for the two objects), the PR system has no choice but to fall back on the 'where' descriptors.

From an adult perspective, the process of producing distinct spatiotemporal descriptors for the two objects in an occlusion event, by specifying each object's individual trajectory, seems quite trivial (e.g., 'object-A, which initially stands behind the left half of the screen, emerges to the left of the screen and then returns behind it'; 'object-B, which initially stands behind the right half of the screen, emerges to the right of the screen and then returns behind it'). However, the many failures that have been obtained with categorical/featural conditions at 10 and 12 months indicate that this process is actually far from trivial for infants. Why should that be the case? The answer has to do, in large part, with infants' limited working memory and spatial-reasoning skills (e.g., Moore & Johnson, 2008; Quinn & Liben, 2008; Schweinle & Wilcox, 2004). Spatial-reasoning skills, in particular, develop slowly during the first year of life, though somewhat faster in males than in females.

Evidence for the preceding analysis comes from three key findings obtained by Wilcox and her colleagues with categorical/featural conditions: (1) when the occlusion event is very brief, even young infants give evidence that they expect to see two objects when the screen is removed (e.g., Wilcox & Baillargeon, 1998a; Wilcox & Schweinle, 2002); (2) up to a point, the longer the occlusion event, the older the age at which infants succeed, with males typically succeeding before females (e.g., Wilcox, 2007; Wilcox et al., in press); and (3) when the occlusion event becomes too long for infants to handle on their own, they can still succeed if first given ‘piecemeal’ trials that help them isolate each object’s individual trajectory (e.g., Wilcox, 2003, 2007).

To illustrate these findings, consider the three occlusion event sequences in Figure 6, which are ordered from easiest to hardest; for ease of comparison, a self-propelled box and a self-propelled ball serve as the occludees in all the sequences. In the shortest sequence (Figure 6A), the box moves behind the left edge of the screen, the ball appears at the right edge of the screen, and then the screen is lowered to reveal no object; only the ball is visible to the right of the screen. Infants aged 5.5 to 9 months have been shown to succeed (i.e., to give evidence that they expect two objects) with such brief occlusion sequences, though 4.5-month-olds fail (e.g., Wilcox & Baillargeon, 1998a; Wilcox & Schweinle, 2002). The sequence in Figure 6B begins the same way but then continues on in reverse: the ball returns behind the screen, the box returns to its starting position to the left of the screen, and finally the screen is lowered to reveal no object; only the box is visible to the left of the screen. At 9 months, males succeed with this sequence, but females do not; at 4.5 months, both males and females fail (Wilcox et al., in press). In the sequence shown in Figure 6C, the box emerges to the left of the screen and returns behind it, the ball emerges to the right of the screen and returns behind it, and then the screen is lowered to reveal only the ball. At 10.5 months, males succeed with this sequence, but females do not; at 9.5 months, both males and females fail (e.g., Wilcox, 2007; Wilcox & Baillargeon, 1998a; Wilcox & Chapa, 2002). In additional experiments, Wilcox (2003, 2007) found that younger infants succeed with the sequence in Figure 6C if they first receive two piecemeal trials that help them isolate each object’s trajectory. In the first piecemeal trial, the box emerges to the left of the screen and returns behind it, and nothing else happens until the infant looks away and the trial ends. In the second piecemeal trial, the ball emerges to the right of the screen and returns behind it, and again nothing else happens until the trial ends. After receiving these two piecemeal trials, 10.5-month-old females now succeed at the task (Wilcox, 2007), as do 7.5-month-old males (but not females; Wilcox, 2003). Interestingly, piecemeal trials are effective only if the objects become occluded: younger infants still fail with the sequence in Figure 6C if they receive control-piecemeal trials in which they see the box to the left of the screen throughout the first trial and the ball to the right of the screen throughout the second trial (Wilcox & Baillargeon, 1998a; Wilcox, 2003). These negative results make clear that piecemeal trials enhance infants' performance, not because they allow infants to become familiar with the box and ball, but because they help infants reason about each occludee's trajectory.

Our interpretation of these various findings is as follows. In each occlusion sequence, the object information in the structural layer initially specifies that only one object is present: the box and ball receive identical categorical descriptors, namely self-propelled, non-agentive, and closed. However, when the object information in the variable layer specifies that two objects are in fact present (recall that size and shape are identified as occlusion variables by about 3.5 to 4.5 months), a discrepancy is detected. Next, the PR system attempts to resolve this discrepancy by assigning distinct spatiotemporal descriptors to the objects. When the occlusion event is very brief, so that it is relatively easy to determine each object's trajectory, the PR system assigns a distinct spatiotemporal descriptor to each object; the information in the variable layer is then bound to these descriptors, forming two object bundles that can be carried over to the next physical representation. When the occlusion event is slightly longer, it becomes harder for infants to determine each object's trajectory, with females often lagging behind males in this process. Finally, in those cases where infants cannot spontaneously determine each object's trajectory, piecemeal trials can help by presenting only one object's trajectory per trial.

III. Infants' Use of Spatiotemporal and Categorical Information

According to the carryover account described in the preceding section, infants in object-individuation tasks hold a clear expectation about how many objects should be present when the screen is removed as long as the *same number* of objects is specified in the structural and variable layers of the occlusion event's physical representation. Thus, if one object is specified in both layers, infants expect to see one object and are surprised to see none; and if two objects are specified in both layers, infants expect to see two objects and are surprised to see only one (in each case, as noted earlier, infants may not be surprised to see *more* objects than expected, if the physical layout makes it possible for them to assume that additional objects were already hidden behind the screen). However, (1) if *different numbers* of objects are specified in the structural layer (one object) and the variable layer (two objects), and (2) the PR system is unable to revise the structural layer by creating distinct spatiotemporal descriptors for the two objects, then the carryover of object information breaks down, and infants show *no clear expectation* about how many objects should be revealed when the screen is removed. (Of course, in most physical events infants see every day, there will be agreement between the number of objects specified in the structural and variable layers; it is only in unusual circumstances, such as those uncovered by Xu and Carey (1996), that inconsistencies arise, creating an opportune window through which researchers can study the development of infants' representations of objects and events).

In the next sections, we examine recent findings from the object-individuation literature in terms of the carryover account. In particular, we look at infants' responses in situations where either *spatiotemporal* or *categorical* information in the structural layer specifies that two objects are involved in the occlusion event. Because there is then agreement between the structural and variable layers about the number of objects present (i.e., two objects), this object information can be carried over to the next physical representation, leading infants to expect to see two objects when the screen is removed.

A. Spatiotemporal information

As we saw earlier, Xu and her colleagues found that, if two objects emerge first simultaneously and then in alternation from behind a screen, 10- and 12-month-olds expect to see two objects when the screen is removed (e.g., Xu & Carey, 1996; Xu et al., 2004). A number of investigations have shown that, in addition to *location* information, the PR system can use several other kinds of spatiotemporal information to set up two object representations in the structural layer of events' physical representations (see Figure 7). For example, infants attend to *path* information: if an object disappears behind one screen and reappears from behind another screen without appearing in the gap between them, infants as young as 4 months expect to see two objects when the screens are removed (e.g., Spelke et al., 1995a; Xu & Carey, 1996; see also Aguiar & Baillargeon, 2002). Infants also attend to *speed* information: if an object disappears behind the left edge of a wide screen and immediately reappears at the right edge (i.e., too soon to have traveled, at its present speed, from one edge to the other), infants as young as 4.5 months of age expect to see two objects when the screen is removed (e.g., Wilcox & Schweinle, 2003). In line with the sex differences reported in the last section, if the event is made slightly longer by having the object return to its starting position to the left of the screen, then 4.5-month-olds now fail at the task; infants do not succeed until about 7.5 months for males and 9.5 months for females (Schweinle & Wilcox, 2004).

B. Categorical information

According to the account of physical reasoning described earlier, the structural layer in the physical representation of an event includes categorical information about the objects in the event; early categorical descriptors include *ontological* and *functional* descriptors (e.g., Csibra, 2008; Hespos & Baillargeon, 2001b; Johnson et al., 2007; Luo et al., 2009; Saxe et al., 2007; Wang et al., 2005). This account predicts that infants should succeed at object-individuation tasks as long as the two objects in the occlusion event receive different ontological of functional descriptors, leading the PR system to set up two object representations in the structural layer of the event's physical representation. Below, we review findings consistent with this prediction (see Figure 7).

1. Ontological categories

Surian and Caldi (2010) tested 10-month-olds in a categorical/featural condition using computer-animated events that contrasted an *inert* object and a *self-propelled agentive* object. For example, in one test pair, an experimenter's hand brought out a blue block from behind the left edge of a large screen and then returned the block behind the screen. Next, a black and yellow bee emerged from behind the right edge of the screen and flew toward a hand; the hand initially withdrew but then began to chase the bee, which flew back behind the screen. This sequence was repeated several times and then the screen was lowered to reveal, as usual, one or both of the objects. Results indicated that infants looked reliably longer at the one- than at the two-object event (infants in a baseline condition looked about equally at the two events). These data suggest that, because the two objects shown in alternation received different categorical descriptors (e.g., inert, non-agentive, and closed for the block; self-propelled, agentive, and closed for the bee), two object representations were created in the structural layer of the occlusion event's physical representation; variable information was then bound to each object representation. When the screen was removed, the two object bundles were carried over to the next physical representation, leading infants to expect two objects. Additional results indicated that, when *both* of the objects shown in alternation were self-propelled and agentive (e.g., a bee and a rabbit), infants looked about equally at the one- and two-object events. Because the objects now received identical categorical descriptors (self-propelled, agentive, and closed), one object representation was created in the structural layer but two in the variable layer (because of the differences in the objects' sizes, shapes, and so on). The PR system was not able to resolve this inconsistency, since the occlusion event was quite long, so no object information was carried over to the next physical representation. As a result, infants had no clear expectation about how many objects should be present when the screen was removed.

Bonatti et al. (2002, 2005) reported data that lend themselves to a similar analysis. They tested 10- and 12-month-olds in a categorical/featural condition contrasting *humanlike* and *non-humanlike* objects. At both ages, infants expected to see two objects when the screen was removed if the objects used in the test trial were (1) a humanlike doll and a non-humanlike object (e.g., a red-haired doll and a toy dog), or (2) two different dolls, one with its head right-side-up (humanlike) and one with its head upside-down on its shoulders (non-humanlike). In contrast, infants had no specific expectation about how many objects they should see when the screen was removed if the test trial involved two different humanlike dolls, both with their heads right-side-up. As mentioned earlier, we take these results to indicate that the categorical information in the structural layer of physical representations includes descriptors (at least by 10 months) that specify whether objects are humanlike or not. Thus, when a humanlike and a non-humanlike object emerge in alternation from behind a screen, the objects receive different categorical descriptors, leading the PR system to set up two object representations in the structural layer of the event's physical representation; variable information is then bound to each object representation, and the two object bundles are carried over to the next physical representation, leading infants to expect two objects. In contrast, when two different humanlike objects emerge in alternation from behind a screen, both objects receive the same categorical descriptors (e.g., self-propelled, humanlike). As a result, only one object representation is set up in the structural layer whereas two are set up in the variable layer. If the PR system is unable to resolve this inconsistency (because the event is too long for infants to determine each object's trajectory), no object information is carried over to the next physical representation, and infants have no clear expectation about how many objects they should see when the screen is removed.

2. Functional categories

As mentioned above, part of the categorical information young infants routinely encode in the structural layer of physical representations includes simple functional descriptors, such as whether objects are closed (e.g., like balls or blocks) or have internal openings (e.g., like rings, containers, or tubes). A few laboratories are testing whether young infants succeed (as predicted) at object-individuation tasks contrasting closed and open objects (e.g., a block and a container), but no data have yet been published.

Research by Wilcox et al. (2006) suggests that young infants also assign different functional descriptors to rattles whose contents produce distinct sounds when shaken. In one experiment, 4.5-month-olds were tested in an object-individuation task contrasting two different rattles (see Figure 7). Infants in an experimental condition received two test trials. At the start of each trial, infants saw an experimenter's hand move behind a large screen; the hand and its actions were then fully hidden. Next, infants heard (but could not see) two different rattles: one rattle (e.g., a papier-mâché egg lined with plastic and partially filled with uncooked rice) was shaken for a few seconds, and then the other rattle (e.g., a similar egg partially filled with small jingle bells) was shaken for a few seconds. Finally, the hand emerged from behind the screen, which was then lowered to reveal either one egg (one-object event) or two identical eggs (two-object event). Infants in a control condition heard two similar rattles (e.g., two eggs filled with rice). Results indicated that in the experimental condition infants who saw the one-object event looked reliably longer than those who saw the two-object event, whereas in the control condition infants looked about equally at the two events. In additional experiments, the positive result of the experimental condition disappeared (1) when infants heard two different non-rattling sounds (e.g., two different notes on an electronic keyboard), or (2) when infants heard the two different rattles but did not see a hand move behind the screen. One interpretation of these results is that infants in the experimental condition construed the sounds they heard as coming from two objects from different functional categories: each object held different contents that produced different sounds when shaken (one rattled whereas the other jingled). As a result, the PR system set up two object representations in the structural layer of the occlusion event's physical representation, leading infants to expect two objects when the screen was removed. (Support for the present interpretation comes from an experiment by Dewar and Xu (2009). After being familiarized to two identical objects that produced different sounds when shaken (e.g., one rattled and one jingled), 10-month-olds were surprised when the two objects were labeled with the same naming phrase (e.g., "A zav!"). Thus, like the 4-month-olds tested by Wilcox et al., the 10-month-olds tested by Dewar and Xu treated the two rattles as belonging to two different functional categories, like a guitar and a drum).

What happens when infants are presented with novel artifacts whose functions are less familiar or more opaque? Futó et al. (2010) obtained evidence in an object-individuation task that 10-month-olds could be induced with pedagogical support to encode two such artifacts as belonging to two distinct functional categories (i.e., as two different kinds of artifacts used to produce different effects, like a lamp and a radio; see Figure 7). Infants in a function condition were tested with videotaped events contrasting two small objects: one flashed small lights in its front surface when its handle was pulled, and the other played a melody when its dial was rotated. In each familiarization event, infants heard “Hi, baby, hi!” and then they saw an experimenter’s hand bring one of the objects out from behind a large screen; the hand demonstrated the object’s function several times and then returned the object behind the screen. Next, infants heard “Watch this!” and then they saw the hand demonstrate the other object’s function on the opposite side of the screen. The test events were similar except that, after returning the second object behind the screen, the hand removed the screen to reveal one or both objects. Infants in a baseline condition saw abbreviated events in which the screen was simply removed to reveal the object(s). In the baseline condition, infants looked reliably longer at the two- than at the one-object event; in the function condition, in contrast, infants looked about equally at the two events. This last result suggests that the PR system assigned different categorical descriptors to the two objects based on their different functions; as a result, two object representations were created in the structural layer, the featural information in the variable layer was linked to each representation, and the two object bundles were carried over to the next physical representation when the screen was removed, leading infants to expect two objects. Interestingly, the result of the function condition was eliminated if either (1) no communicative cues were used or (2) the hand was no longer involved in demonstrating the objects’ functions (e.g., the dial rotated by itself). The objects’ functions were obviously not sufficiently salient or ‘transparent’ to be encoded categorically by infants without substantial contextual support.

3. Taxonomic categories

As was discussed earlier, the evidence reported by Xu and her colleagues (Xu & Carey, 1996; Xu et al., 2004; see also Kingo & Krøjgaard, 2011) suggests that, by their first birthday, infants begin to include object categories (e.g., ‘bottle’ or ‘ball’) among the categorical descriptors for objects in the structural layer of physical representations. Younger infants typically do not include these taxonomic descriptors, although they can easily be induced to do so through contextual manipulations. Below, we describe two such manipulations (see Figure 7).

In a series of experiments, Xu (2002) found that 9-month-olds tested in a standard categorical/featural condition gave evidence that they expected to see two objects when the screen was removed if the objects were *labeled with different naming phrases* as they emerged in alternation from behind the screen. For example, infants who saw an occlusion event involving a toy duck and a ball heard “Look, [baby’s name], a duck!” and “Look, [baby’s name], a ball!" as the objects came into view. Infants looked about equally at the one- and two-object test events, and this result (relative to that of a baseline condition) indicated that infants expected to see two objects when the screen was removed. The same result was also obtained with novel objects and labels ("a fendle", "a toma"), but was eliminated when the two objects received the same label ("a toy") or were paired with different tones or emotional expressions ("Ah", "Ewy") instead of with labels. Together, these results suggest that, by 9 months of age, a naming phrase, even when novel, induces infants to represent an object in terms of its *object category* (e.g., Balaban & Waxman, 1997; Graham et al., 2004; Waxman, 1999; Welder & Graham, 2001). Thus, when two objects that emerge in alternation from behind a screen receive different labels, these different taxonomic descriptors lead the PR system to set up two object representations in the structural layer of the event's physical representation; the information in the variable layer is then bound to each object representation, and the two object bundles are carried over to the next physical representation, leading infants to expect two objects when the screen is removed.

X. Li, Stavans, Baillargeon, Carey, and Bonatti (2011) used a different, non-verbal manipulation to induce 9.5-month-olds to encode objects' categories. In one experiment, infants were assigned to a two-category or a mixed-objects condition. In the two-category condition, infants first received two display trials designed to highlight two object categories: in one trial, infants saw three different cups resting apart on the apparatus floor; in the other trial, infants saw three different baby shoes. In the test trial, an experimenter's hand brought out a cup and a shoe in alternation from behind a large screen, and then the screen was removed to reveal only one of the objects (e.g., the shoe). The mixed-objects condition was similar except that the objects were not sorted by category in the display trials: infants saw two cups and one shoe in one trial and two shoes and one cup in the other trial. Infants in the two-category condition looked reliably longer at the one-object event than did those in the mixed-objects condition, suggesting that (1) the display trials induced the infants in the two-category condition to focus on two object categories, cups and shoes; (2) when the infants next saw one of the cups and shoes brought out from behind the screen, they encoded each object categorically ('cup', 'shoe'); and (3) these different taxonomic descriptors led the PR system to set up two distinct object representations in the structural layer of the event's representation. As a result, infants expected to see two objects when the screen was removed, and they detected a violation when they did not. Similar results were obtained with novel object categories (brightly patterned blocks and cylinders). Finally, negative results were found when infants were shown three *identical* objects in each display trial (i.e., three identical cups, three identical shoes), suggesting that it was not prior exposure to the test objects, but rather categorical encoding of the test objects, that contributed to infants' success at the task.

IV. Do Infants Realize that Two Objects Are Present during the Occlusion Event?

According to the carryover account, young infants fail in standard categorical/featural conditions because no object information is carried over to the next physical representation when (1) the information in the structural layer of the occlusion event's physical representation specifies that one object is present, whereas the information in the variable layer specifies that two objects are present, and (2) the PR system is unable to revise the object information in the structural layer by assigning distinct spatiotemporal descriptors to the two objects (infants' spatial-reasoning skills are overwhelmed so they are unable to specify each object's unique trajectory). This account explains why infants have no clear expectation about how many objects should be present *after* the occlusion event, when the screen is removed. But it leaves unclear what expectations, if any, infants possess *during* the occlusion event.

One possibility is that infants have no clear expectation about how many objects are present behind the screen during the occlusion event: since there is a conflict between the object information in the structural layer (one object) and in the variable layer (two objects), infants may have a confused or ambiguous picture of the event. Another possibility, however, is that inconsistencies between the structural and variable layers matter *only* when it comes to the carryover of object information between physical representations (i.e., only object information that is properly bound or bundled can be carried over). On this view, during the occlusion event, the object information in the variable layer would be correctly interpreted by the PR system (and the persistence principle) to indicate that two objects are present.

Which of the two possibilities outlined above is correct? Standard object-individuation tasks cannot address this question, because they examine how many objects infants expect to see *after* the occlusion event is ended. Fortunately, two-screen object-individuation tasks have been developed that allow us to assess infants' expectations about how many objects are present *during* the occlusion event; these tasks are described next.

1. Two-screen tasks with an opaque and a transparent screen

When young infants are tested in a categorical/featural condition and see the two objects emerge in alternation from behind the screen, do they hold a clear expectation that two objects are present? Dramatic evidence that they do comes from two-screen experiments by Wilcox and Chapa (2002). In these experiments, 9.5-month-olds were tested in a condition similar to a standard categorical/featural condition, with one critical exception: a *transparent screen* stood behind the opaque screen and was revealed when the opaque screen was removed.

In one experiment, for example, 9.5-month-olds were assigned to a box-ball or a ball-ball condition (see Figure 8A). In the box-ball condition, infants were first allowed to inspect the transparent screen and its dark frame. Next, infants received a test trial with an initial and a final phase. In the initial phase, an experimenter's hand brought out a box and a ball in alternation from behind a large opaque screen; in the final phase, the screen was lowered to reveal the transparent screen, behind which stood only the ball. In the ball-ball condition, a ball was shown on both sides of the screen. Infants in the box-ball condition looked reliably longer during the final phase than did those in the ball-ball condition, suggesting that they expected to see two objects behind the transparent screen. This effect was eliminated when no transparent screen stood behind the opaque screen, as in the standard categorical/featural condition of Xu and Carey (1996).

Our interpretation of these results is as follows. The addition of the transparent screen meant that, when the opaque screen was removed, the infants were *still faced* with an occlusion event, albeit one now involving a transparent rather than an opaque occluder (recall that an occlusion event is an event in which an object moves or is placed behind another object, or occluder; the occluder may be short or tall, wide or narrow, opaque or transparent, and so on). Because the occlusion event continued, with the transparent screen as the occluder and the box and ball as the occludees, no new physical representation was required; the PR system simply updated the physical representation of the ongoing occlusion event to include the second, transparent occluder. Because by 9.5 months most infants have identified size, shape, and pattern information as occlusion variables, all of this information was included in the variable layer of the physical representation and (in accordance with the principle of persistence and its corollary of unchangeableness) led the infants to expect two objects behind the transparent screen. Accordingly, infants detected a violation when only one object was revealed behind the transparent screen.

2. Two-screen tasks with two opaque screens

Hespos and Baillargeon (cited in Needham & Baillargeon, 2000) obtained similar findings in a different two-screen task with younger infants. In one experiment, 6.5-month-olds were assigned to a ball-box or a box-box condition (see Figure 8B). In the ball-box condition, infants received a test trial with an initial and a final phase. During the initial phase, an experimenter's hand lifted a ball above the right end of a tall screen and then lowered it back behind the screen; next, the hand lifted a box above the left end of the screen and again returned it behind the screen (this whole sequence was repeated three times). During the final phase, the screen was lowered to reveal a short opaque screen, and the hand now lifted and lowered the box at both ends of the screen (i.e., the ball was no longer shown; infants thus did not see an impossible outcome, merely a novel one). Infants in the box-box condition received a similar test trial except that the box was lifted and lowered at either end of the screen in both phases. Infants in the ball-box condition looked reliably longer during the final phase than did those in the box-box condition, suggesting that they noticed that the hand was no longer showing the ball. This effect was eliminated when the tall screen was lowered to reveal a short container instead of a short screen (the front of the container was identical to the short screen). Together, these and control results suggested that, as long as infants were presented with an ongoing occlusion situation, (1) they retained access to the object information in the variable layer, which specified that two different objects (a ball and a box) were shown, and (2) they detected the change in the final phase when a single object was shown instead. In contrast, when the tall screen was lowered to reveal a container, the PR system was forced to build a new physical representation for this new, containment situation. The object information in the physical representation of the initial, occlusion event was inconsistent (one object was specified in the structural layer, but two in the variable layer) and could not be carried over to the physical representation of the containment event; as a result, infants did not detect the object change introduced in this event.

Together, the findings from the various two-screen tasks described above point to two important conclusions. First, when tested in a standard categorical/featural condition, it is only *after* the occlusion event is ended, when the screen is removed, that young infants hold no clear expectation about how many objects should be present; *during* the occlusion event, infants are led by the information in the variable layer to realize that two objects are involved in the event. Second, standard categorical/featural conditions provide a poor test of young infants' individuation abilities: although infants often realize that two objects are present while the occlusion event is occurring, they fail to detect a violation when the screen is removed to reveal only one object. During the occlusion event, infants can use the featural differences between the two objects to individuate them—but under standard task conditions they are rarely able to resolve the discrepancy between the variable and structural layers. As a result, no carryover of object information occurs, and infants fail to detect the violation in the one-object event.

V. Predictions

Our account of object-individuation findings makes several interesting predictions; three are listed below.

First, and perhaps most strikingly, our account predicts that 10-month-olds who are tested in a standard categorical/featural condition with two objects from different taxonomic categories (e.g., a toy duck and a ball) should fail to detect a violation not only when the screen is removed to reveal *one object*, as shown previously, but also when the screen is removed to reveal *no object*. Similarly, 12-month-olds who are tested in a standard categorical/featural condition with two different objects from the same taxonomic category (e.g., a large and a small ball) should fail to detect a violation not only when the screen is removed to reveal *one object*, as shown before, but also when the screen is removed to reveal *no object*. In either case, if there is no carryover of object information when the screen is removed, then infants should have no expectation about what they should see, and they should detect no violation if shown either one object or no object. Experiments (by Stavans and Baillargeon) are under way to test this prediction, with promising results. Such results would provide robust support for the carryover account proposed here and for our claim that, when the screen is removed, the PR system produces what is essentially an error message: it was not able to compute a specific expectation about what should be seen, and so it simply starts over with the next event.

Second, our account predicts that young infants who are tested in a standard categorical/featural condition with two simple artifacts from different functional categories (e.g., a cup and a shoe, or a key and a spoon) might succeed if first given a brief demonstration of each artifact's canonical function (more complex or opaque functions would also require pedagogical support, as in Futó et al., 2010). It may be that young infants readily form category representations of everyday artifacts and their functions but that these representations tend to be more flexible, less stable, or more diffuse than those of older children and adults (e.g., German & Defeyter, 2000): after all, shoes, cups, keys, and markers can all be chewed, banged, shaken, thrown, and so on. Thus, when presented with objects in a minimal context (e.g., a shoe and a cup emerging in alternation from behind a screen), young infants may not spontaneously focus on the objects' conventional functional categories—but they may do so following an appropriate hint. The PR system would then include different categorical (functional) descriptors for the two objects in the structural layer of the occlusion event's physical representation, leading infants to expect two objects when the screen is removed. Again, experiments (by Stavans and Baillargeon) are under way to test this prediction, with promising results. These positive results would make clear that there are multiple ways of inducing young infants to encode objects categorically: by labeling each object as it comes into view (Xu, 2002), by presenting displays of different objects from each category (e.g., three different cups and three different shoes; X. Li et al., 2011), and by demonstrating each object's function prior to test (above).

Finally, a general prediction from our review of object-individuation findings is as follows: whether infants succeed or fail at a standard categorical/featural task that is modified to bypass carryover difficulties—as in a brief-event, piecemeal, or two-screen task—will depend on whether the PR system (1) has identified the variable(s) necessary to distinguish the two objects used in the task and thus (2) includes the relevant information in the variable layer of the occlusion event's physical representation. In line with this prediction, 10-month-olds *succeeded* at a two-screen task with two human dolls that differed in hair and skin color, hairstyle, and hair ornament (Wu & Baillargeon, 2008): because by 10 months infants have identified size, shape, and pattern (though not color) as occlusion variables, and the dolls differed in multiple features, infants responded correctly when a single doll was revealed behind the transparent screen. Conversely, 8.5-month-olds *failed* at a two-screen task with two cylinders that differed only in color: because by 8.5 months have not yet identified color as an occlusion variable, infants mistakenly expected a single cylinder to be present behind the transparent screen (Ng, Baillargeon, & Wilcox, 2007).

VI. Concluding Remarks

The findings and interpretations from the various categorical/featural object-individuation tasks reviewed in this article can be summarized in five main points.

1. When two distinct objects emerge in alternation from behind a wide screen, young infants recognize *during* the event that two objects are present, as long as the objects are assigned distinct categorical (ontological, functional, or taxonomic) descriptors and/or the objects differ in featural properties that have already been identified as occlusion variables. This expectation can be demonstrated in two-screen tasks as well as in other physical-

reasoning tasks (e.g., Wilcox, 1999; Wilcox & Baillargeon, 1998b).

2. After the screen is removed, young infants give no evidence that they expect to see two objects if, during the occlusion event, the categorical information in the structural layer specified that a single object was present whereas the featural information in the variable layer specified that two objects were present, *and* the PR system was unable to resolve this discrepancy (due to working-memory and spatial-reasoning limitations) by creating a distinct spatiotemporal descriptor for each object in the structural layer. Object information can be carried over to the next physical representation only when it is internally consistent; when it is not, no expectation is carried over about the number of objects that should be present.

3. After the screen is removed, young infants expect to see two objects if, during the occlusion event, (1) the spatiotemporal and/or categorical information in the structural layer specified that two objects were present, *or* (2) the categorical information in the structural layer pointed to a single object whereas the featural information in the variable layer pointed to two objects, but the PR system was able to resolve this discrepancy by creating a distinct spatiotemporal descriptor for each object in the structural layer, either spontaneously (as in brief-event tasks) or following trials designed to isolate each object's trajectory (as in piecemeal tasks).

4. By the end of the first year, infants are more likely to encode objects' taxonomic categories (perhaps as a result of language acquisition, as suggested by Xu and her colleagues; e.g., Xu, 2002; Xu & Carey, 1996; Xu et al., 2005). As a result, older infants often succeed when tested with objects from different taxonomic categories (even with long occlusion sequences), because the categorical information in the structural layer and the featural information in the variable layer then both point to two objects. This object information can then be carried over to the next physical representation, leading infants to expect two objects.

5. Finally, a speculation: In object-individuation tasks, infants expect to see two objects when the screen is removed only when the information in the structural layer of the occlusion event's physical representation includes two distinct categorical ('what') or spatiotemporal ('where') descriptors. Why should this be true? One possibility is that 'what' and 'where' descriptors provide the various systems that underlie infants' responses to objects and events with common *labels* for communicating about objects. For example, 'what' descriptors may allow the physical-reasoning and object-representation systems to more easily exchange information about objects, to compare and

align representations, and so on. With the advent of language acquisition, taxonomic labels (e.g., “a cup”) would naturally be recruited into this process, as they provide even more precise descriptors for objects.

Xu and Carey (1996) were the first to suggest that, like spatiotemporal information, categorical information plays a critical role in infants’ representations of physical events, and everything in this article supports this seminal intuition. Of course, this intuition had to be elaborated as many new (and often seemingly contradictory) findings have come to light over the past 15 years. In this article, we have offered a new account for these various findings, by placing them within a broader account of how infants learn, represent, and reason about physical events.

Acknowledgements

The preparation of this article was supported by a grant from the National Institute of Child Health and Human Development to Renée Baillargeon (HD-21104). We would like to thank Luca Bonatti, Andrei Cimpian, Jerry DeJong, Cindy Fisher, Jane Hedges, Amy Needham, Dan Swingley, and Teresa Wilcox for helpful comments and discussions.

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.

Aguiar, A., & Baillargeon, R. (2002). Developments in young infants' reasoning about occluded objects. *Cognitive Psychology*, *45*, 267-336.

Baillargeon, R. (1994). Physical reasoning in young infants: Seeking explanations for unexpected events. *British Journal of Developmental Psychology, 12*, 9-33.

Baillargeon, R. (2002). The acquisition of physical knowledge in infancy: A summary in eight lessons. In U. Goswami (Ed.), *Blackwell handbook of childhood cognitive development, first edition* (pp. 46-83). Oxford: Blackwell.

Baillargeon, R. (2008). Innate ideas revisited: For a principle of persistence in infants' physical reasoning. *Perspectives on Psychological Science*, *3*, 2-13.

Baillargeon, R., & Carey S. (in press). Core cognition and beyond: The acquisition of physical and numerical knowledge. In S. Pauen & M. Bornstein (Eds.), *Early childhood development and later achievement*. London: Cambridge University Press.

Baillargeon, R., & DeVos, J. (1991). Object permanence in 3.5- and 4.5-month-old 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, second edition (*pp. 11-48). Oxford: Blackwell.

Baillargeon, R., Li, J., Ng, W., & Yuan, S. (2009a). An account of infants' physical reasoning. In A. Woodward & A. Needham (Eds.), *Learning and the infant mind* (pp. 66-116). New York: 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., Spelke, E. S., & Wasserman, S. (1985). Object permanence in 5-month-old infants. *Cognition*, *20*, 191-208.

Baillargeon, R., Wu, D., Yuan, S., Li, J., & Luo, Y. (2009b). Young infants' expectations about self-propelled objects. In B. Hood & L. Santos (Eds.), *The origins of object knowledge* (pp. 285-352). Oxford: Oxford University Press.

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology*, *64*, 3-26.

Bernard, A., Baillargeon, R., & DeJong, G. F. (2011, March). Going beyond statistical regularities: Evidence for causal reasoning in infants. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Montreal, Canada.

Bonatti, L., Frot, E., & Mehler, J. (2005). What face inversion does to infants' numerical abilities*. Psychological Science, 16*, 506-510.

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.

Casasola, M., Cohen, L.B., & Chiarello, E. (2003). Six-month-old infants' categorization of containment spatial relations. *Child Development*, *74*, 679-693.

Cheries, E., Wynn, K., & Scholl, B. (2006). Interrupting infants' persisting object representations: An object-based limit? *Developmental Science, 9*, F50-F58.

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition*, *107*, 705-717.

DeJong, G. F. (1997). Explanation-based learning. In A. Tucker (Ed.), *Encyclopedia of computer science* (pp. 499-520). Boca Raton, FL: CRC Press.

Dewar, K., & Xu, F. (2009). Do early nouns refer to kinds or distinct shapes? *Psychological Science, 20*, 252-257.

Eimas, P. D., & Quinn, P. C. (1994). Studies on the formation of perceptually based basic-level categories in young infants. *Child Development, 65,* 903-917.

Feigenson, L., & Carey, S. (2005). On the limits of infants' quantification of small object arrays. *Cognition*, *97*, 295-313.

Feigenson, L., & Halberda, J. (2004). Infants chunk object arrays into sets of individuals. *Cognition, 91,* 173-190.

Fiser, J., & Aslin, R. N. (2002) Statistical learning of new visual feature combinations by infants. *Proceedings of* *the National Academy of Sciences, 99,* 15822-15826.

Fisher, C., Gertner, Y., Scott, R., & Yuan, S. (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science, 1,* 143-149.

Futó, J., Téglás, E., Csibra, G., & Gergely, G. (2010). Communicative function demonstration induces kind-based artifact representation in preverbal infants. *Cognition, 117,* 1-8.

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

German, T. P., & Defeyter, M. A. (2000). Immunity to functional fixedness in young children. *Psychonomic Bulletin & Review, 7,* 707-712.

Gertner, Y., Baillargeon, R., Fisher, C. L., & Simons, D. J. (2009, April). Language facilitates infants' physical reasoning. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Denver, CO.

Graham, S. A., Kilbreath, C. S., & Welder, A. N. (2004). 13-month-olds rely on shared labels and shape similarity for inductive inferences. *Child Development*, *75,* 409-427.

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). Knowledge 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., Duffy, S., & Levine, S.C. (2002). Infants and toddlers discriminate amount: Are they measuring? *Psychological Science, 13,* 244-249.

Johnson, S. C., Shimizu, Y. A., & Ok, S-J. (2007). Actors and actions: The role of agent behavior in infants' attribution of goals. *Cognitive Development*, *22*, 310-322.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24,* 174-219.

Keil, F. C. (1995). The growth of causal understandings of natural kinds. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 234-262). Oxford: Clarendon Press.

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

Kingo, O. S., & Krøjgaard, P. (2011). Object manipulation facilitates kind-based object individuation of shape-similar objects. *Cognitive development, 26,* 87-103.

Kittredge, A. K., & Baillargeon, R. (2011, March). Do infants attend to object shape in a novel physical event? Paper presented at the Biennial Meeting of the Society for Research in Child Development, Montreal, Canada.

Kochukhova, O., & Gredebäck, G. (2007). Learning about occlusion: Initial assumptions and rapid adjustments. *Cognition, 105,* 26-46.

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

Krøjgaard, P. (2000). Object individuation in 10-month-old infants: Do significant objects make a difference? *Cognitive Development, 15,* 169-184.

Leslie, A. M. (1995). A theory of agency. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 121-141). Oxford: Clarendon Press.

Leslie, A. M., & Chen, M.L. (2007). Individuation of pairs of objects in infancy. *Developmental Science,* 10, 423-430.

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition*, *25*, 165-288.

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

Li, J., & Baillargeon, R. (2011). Evidence for the persistence principle: Static priming effects in violation-of-expectation and preferential-reaching tasks. Manuscript in preparation.

Li, J., Baillargeon, R., & Needham, A. (2011). How do young infants form event categories? The case of support events. Manuscript in preparation.

Li, J., Sigmon-Hernandez, R., & Baillargeon, R. (2010, March). Infants' reasoning about a sequence of events from distinct categories: Spatiotemporal-based carryover effects. Paper presented at the Biennial International Conference on Infant Studies, Baltimore, MD.

Li, X., Stavans, M., Baillargeon, R., Carey, S., & Bonatti, L. (2011). Object individuation in 9.5-month-old infants: Category-based priming effects. Manuscript in preparation.

Luo, Y., & Baillargeon, R. (2005). When the ordinary seems unexpected: Evidence for rule-based physical reasoning in young infants. *Cognition, 95,* 297-328.

Luo, Y., & Baillargeon, R. (2011). Infants' reasoning about transparent occluders and containers. Manuscript in preparation.

Luo, R., Baillargeon, R., Brueckner, L., & Munakata, Y. (2003). Reasoning about a hidden object after a delay: Evidence from robust representations in 5-month-old infants. *Cognition*, *88*, B23-B32.

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

Mareschal, D., & Johnson, M. H. (2003). The “what” and “where” of object representations in infancy. *Cognition*, *88*, 259-276.

McCrink, K., & Wynn, K. (2004). Large-number addition and subtraction in infants. *Psychological Science, 15,* 776-781.

McDonough, L., Choi, S., & Mandler, J.M. (2003). Spatial categorization: Flexible infants, Lexical adults. *Cognitive Psychology*, *46*, 229-259.

Moore, D. S., & Johnson, S. P. (2008). Mental rotation in human infants: A sex difference. *Psychological Science, 19,* 1063-1066.

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 (1999). *Cognition*, *74*, 255-284.

Needham, A., Cantlon, J. F., & Ormsbee Holley, S. M. (2006). Infants' use of category knowledge and object attributes when segregating objects at 8.5 months of age. *Cognitive Psychology*, *53*, 345-360.

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

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

Ng, W., Baillargeon, R., & Wilcox, T. (2007, March). 8.5-month-old infants use pattern but not color information to individuate objects in occlusion events. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Boston, MA.

Oakes, L. M., Ross-Sheehy, S., & Luck, S. J. (2006). Rapid development of feature binding in visual short-term memory. *Psychological Science, 17,* 781-787.

Onishi, K. H. (2011). Perception of event roles in infancy. Manuscript in preparation.

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

Quinn, P. C. (2007). On the infant's prelinguistic conception of spatial relations: Three developmental trends and their implications for spatial language learning. In J. M. Plumert & J. P. Spencer (Eds.), *The emerging spatial mind* (pp. 117-141*)*. New York: Oxford University Press.

Quinn, P. C., Eimas, P. D., & Rosenkrantz, S. L. (1993). Evidence for representations of perceptually similar natural categories by 3- and 4-month-old infants. *Perception, 22,* 463-475.

Quinn, P. C., & Liben, L. S. (2008). A sex difference in mental rotation in young infants. *Psychological Science, 19,* 1067-1070.

Rivera, S. M., & Zawaydeh, A. N. (2006). Word comprehension facilitates object individuation in 10- and 11-month-old infants. *Brain Research, 1146,* 146-157.

Rose, S. A., Feldman, J., & Jankowski, J. J. (2001). Visual short-term memory in the first year of life: Capacity and recency effects. *Developmental Psychology, 37,* 539-549.

Rose, S. A., Gottfried, A. W., Melloy-Carminar, P., & Bridger, W. H. (1982). Familiarity and novelty preferences in infant recognition memory: Implications for information processing. *Developmental Psychology*, *18,* 704-713.

Saffran, J. R. (2009). What can statistical learning tell us about infant learning? In A. Woodward & A. Needham

(Eds.), *Learning and the infant mind* (pp. 29-46). New York: Oxford University Press.

Saxe, R., Tzelnic, T., & Carey, S. (2006). Five-month-old infants know humans are solid, like inanimate objects. *Cognition, 101,* B1-B8.

Saxe, R., Tzelnic, T., & Carey, S. (2007). Knowing who-dunnit: Infants identify the causal agent in an unseen causal interaction. *Developmental Psychology*, *43*, 149-158.

Schweinle, A., & Wilcox, T. (2004). Sex differences in infants' ability to represent complex event sequences. *Infancy, 6,* 333-359.

Setoh, P., & Baillargeon, R. (2011, March). How do infants generalize new physical expectations? Paper presented at the Biennial Meeting of the Society for Research in Child Development, Montreal, Canada.

Simon, T., Hespos, S. J., & Rochat, P. (1995). Do infants understand simple arithmetic? A replication of Wynn (1992). *Cognitive Development, 10*, 253–269.

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

Spelke, E. S., Kestenbaum, R., Simons, D. J., & Wein, D. (1995a). Spatiotemporal continuity, smoothness of motion, and object identity in infancy. *British Journal of Developmental Psychology, 13,* 113-142.

Spelke, E. S., Phillips, A., & Woodward, A. L. (1995b). 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: Clarendon Press.

Strickland, B., & Scholl, B. J. (2011). Visual perception involves "event type" representations: The case of containment vs. occlusion. Manuscript under review.

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

Valenza, E., & Bulf, H. (2011). Early development of object unity: Evidence for perceptual completion in newborns. *Developmental Science, 14,* 799-808.

von Hofsten, C., Kochukova, O., & Rosander, K. (2007). Predictive tracking over occlusions by 4-month-old infants. *Developmental Science, 10,* 625-640.

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., & 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 11-month-old infants. *Developmental Science, 12,* 681-687.

Waxman, S. R. (1999). Specifying the scope of 13-month-olds' expectations for novel words. *Cognition, 70,* B35-B50.

Welder, A. N., & Graham, S. A. (2001). The influence of shape similarity and shared labels on infants' inductive inferences about nonobvious object properties. *Child Development, 72, 1653-1673*.

Wilcox, T. (1999). Object Individuation: Infants use of shape, size, pattern, and color. *Cognition, 72*, 125-166.

Wilcox, T. (2003). Event-mapping tasks: Investigating the effects of prior information and event complexity on performance. *Infant Behavior and Development, 26,* 568-587.

Wilcox, T. (2007). Sex differences in infants' mapping of complex occlusion sequences: Further evidence. *Infancy, 12,* 1-25.

Wilcox, T., Alexander, A. M., Wheeler, L., & Norvell, J. M. (in press). Sex differences during visual scanning of occlusion events. *Developmental Psychology*.

Wilcox, T., & Baillargeon, R. (1998a). Object individuation in infancy: The use of featural information in reasoning about occlusion events. *Cognitive Psychology, 37*, 97-155.

Wilcox, T., & Baillargeon, R. (1998b). Object individuation in young infants: Further evidence with an event-monitoring task. *Developmental Science*, *1*, 127-142.

Wilcox, T., & Chapa, C. (2002). Infants’ reasoning about opaque and transparent occluders in an individuation task. *Cognition, 85*, 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., & Schweinle, A. (2002). Object individuation and event mapping: Developmental changes in infants' use of featural information. *Developmental Science*, *5*, 87-105.

Wilcox, T., & Schweinle, A. (2003). Infants use of speed of motion to individuate objects in occlusion events. *Infant Behavior and Development*, *26*, 253-282.

Wilcox, T., Woods, R., Tuggy, L., & Napoli, R. (2006). Shake, rattle, and….one or two objects? Young infants’ use of auditory information to individuate objects. *Infancy, 9*, 97-123.

Wilson, R. A., & Keil, F. C. (2000). The shadows and shallows of explanation. In F. C. Keil & R. A. Wilson (Eds.), *Explanation and cognition* (pp. 87-114). Cambridge, MA: MIT Press.

Wu, D. (2011, March). Young infants detect open-closed but not variable changes to objects in covering events. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Montreal, Canada.

Wu, D., & Baillargeon, R. (2008, March). One or two humans? 10-month-olds use of ontological and featural information to individuate objects. Paper presented at the Biennial International Conference on Infant Studies, Vancouver, Canada.

Wu, D., Li, X., & Baillargeon, R. (2011). Inferring the presence of a hidden object: Converging evidence for the 100-day milestone. Manuscript in preparation.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, *358*, 749-750.

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*, 400-406.

Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology*, *30*, 111-153.

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*, 372-377.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*, B1-B11.

Yuan, S., & Fisher, C. (2009). "Really? She blicked the baby?": Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science, 20,* 619 - 626.

Yonas, A., & Granrud, C. E. (1984). The development of sensitivity to kinetic, binocular, and pictorial depth information in human infants. In D. Engle, D. Lee, & M. Jeannerod (Eds.), *Brain mechanisms and spatial vision* (pp. 113-145). Dordrecht: Martinus Nijhoff.

Footnotes

1. To give another example of how EBL contributes to the formation of event categories, consider how infants identify the category of *support* events, at about 4 to 5 months of age (e.g., J. Li, Baillargeon, & Needham, 2011). To start, infants notice that when a first inert object is released in contact with a second inert object, the first inert object sometimes remains stable and sometimes falls. Infants cannot predict or interpret these contrastive outcomes: the same physical representation ("first inert object released in contact with second inert object") leads to different outcomes ("first inert object remains stable", "first inert object falls"), suggesting that important information is missing from the representations. Infants then search for the conditions associated with these contrastive outcomes and eventually detect that inert objects remain stable when released *on top of* other inert objects, but not *against* or *under* them. Next, infants engage in a causal analysis of this regularity: their knowledge of gravity (e.g., Needham & Baillargeon, 1993) and solidity (e.g., Baillargeon, Spelke, & Wasserman, 1985) suggests that when object-A is placed on top of object-B, object-B can block the free fall of object-A; however, when object-A is placed against or under object-B, there is then nothing to block the free fall of object-A. This explanation leads to the creation of a new *support* category, a kind of causal interaction in which objects play the role of '*support'* and '*supportee'*. From this point on, when infants see an event in which an object is released in contact with another object, the PR system represents more precisely the spatial arrangement of the objects (i.e., on top of, against, or under); when necessary conditions are met, the PR system categorizes the event as a support event and assigns the roles of support and supportee to the appropriate objects in the event.

2. What happens when the spatiotemporal and the categorical information in the structural layer of an event's physical representation point to inconsistent numbers of objects in the event? In particular, what happens if an object disappears behind one screen and then reappears from behind another screen, without appearing in the gap between them? In this situation, the spatiotemporal information suggests that *two objects* are present (since no single object could follow such a discontinuous path), whereas the categorical information suggests that a *single object* is present (since the same object is seen to disappear and reappear). Beginning at about 100 days of age, infants resolve such conflicts by assuming that two identical objects are involved in the event (e.g., Aguiar & Baillargeon, 2002; Spelke, Kestenbaum, Simons, & Wein, 1995a). Similarly, if a cover is lowered over one object, and a second

cover is then lifted to reveal the object, infants aged 100 days and over assume that an identical object was already hidden under the second cover (e.g., Wu, Li, & Baillargeon, 2011).

3. How does the PR system gather information about a variable? Researchers have suggested that when infants attend to objects, such as the three objects at the start of our event in Figure 2A, (1) the object-tracking (OT) system assigns an *index* to each object (e.g., Leslie, Xu, Tremoulet, & Scholl, 1998; Pylyshyn, 1989), and (2) the object-representation (OR) system opens a temporary *file* for each object, listing both individual (e.g., color) and relational (e.g., relative height) features (e.g., Huttenlocher, Duffy, & Levine, 2002; Kahneman, Treisman, & Gibbs, 1992; Rose, Gottfried, Melloy-Carminar, & Bridger, 1982). When the duck and block move behind the screen, the PR system also becomes involved: the objects are now engaged in a physical interaction, and the PR system's main purpose is that of predicting how such interactions will unfold. One possibility currently being investigated is that, when the PR system requires information about a variable (e.g., color), it taps the OR system for this information (e.g., Wang & Baillargeon, 2008b; Wang & Mitroff, 2009). In the first year of life, object representations in the PR system often contain only a small subset of the featural information in the object files of the OR system; as mentioned earlier, PR representations are initially sparse and become gradually richer as infants identify relevant variables (for further discussion of the links between the OR and PR systems, see Baillargeon et al., 2011; Wang & Baillargeon, 2008b).

4. There appear to be significant developments during the first year of life in infants' ability to bind object information; several factors appear to affect this ability, including the number of objects in an event and the number of events shown side by side (e.g., Kibbe & Leslie, in press; Oakes, Ross-Sheehy, & Luck, 2006; Rose, Feldman, & Jankowski, 2001; Wilcox & Schweinle, 2002). For example, Wilcox and Schweinle (2002) found that in a very simple object-individuation task with two distinct objects, A and B, 7.5-month-olds expected to see A and B when the screen was removed and were surprised to see only A, A and A, or A and C, suggesting that they had formed precise representations of the two objects. In contrast, 5.5-month-olds were surprised to see only A, but they were not surprised to see either A and B or A and A, suggesting that they correctly expected two objects, but lacked precise expectations about their featural properties. Due to space limitations, we do not discuss the development of binding in this article.

Figure Captions

Figure 1: A. Occlusion vector representing some of the variables infants identify to predict whether the occludee will be hidden or visible behind the occluder; B. Occlusion vector representing some of the variables infants identify to predict whether the occludee that reappears from behind the occluder is the same occludee that disappeared or a different occludee.

Figure 2: A. Schematic drawing of an event in which a toy duck and a block move behind a screen; B. Schematic model of how the PR system represents the event and more specifically of what information is included in the structural and variable layers of the event's physical representation.

Figure 3: Schematic model of what object information is carried over to the next physical representation, when the event shown in Figure 2A ends with the screen's removal.

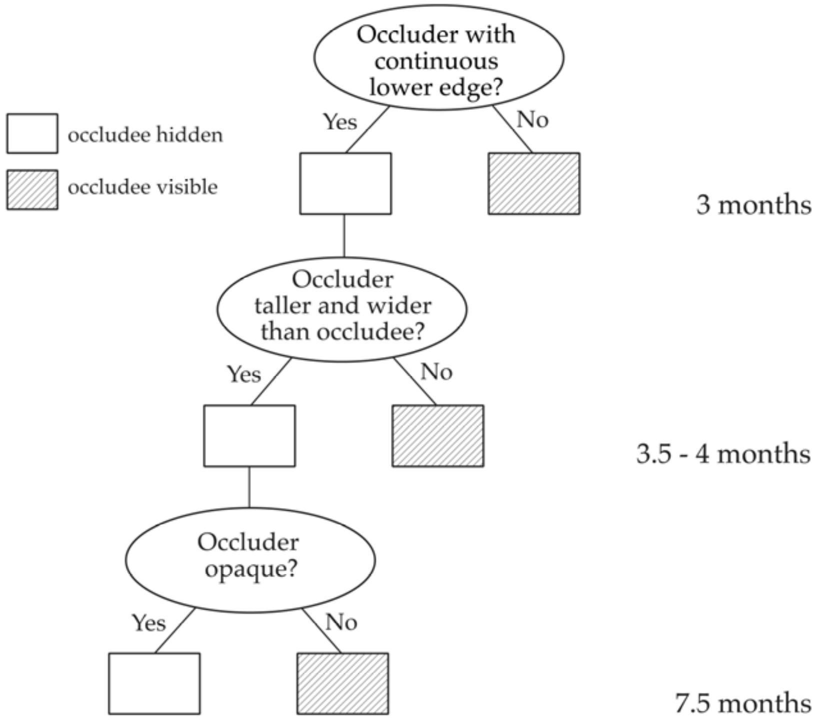
Figure 4: Schematic model of the object information included in the structural and variable layers of the occlusion event's physical representation, in three conditions conducted with 10-month-olds by Xu and Carey (1996): A. Spatiotemporal Condition; B. Single-Object Condition; and C. Categorical/Featural Condition.

Figure 5: Schematic model of the object information included in the structural and variable layers of the occlusion event's physical representation, in three conditions conducted with 12-month-olds by Xu and her colleagues: A. Categorical/Featural Condition: Different Object Categories (Xu & Carey, 1996); B. Categorical/Featural Condition: Same Object Category (Xu et al., 2004); and C. Spatiotemporal Condition: Same Object Category (Xu et al., 2004).

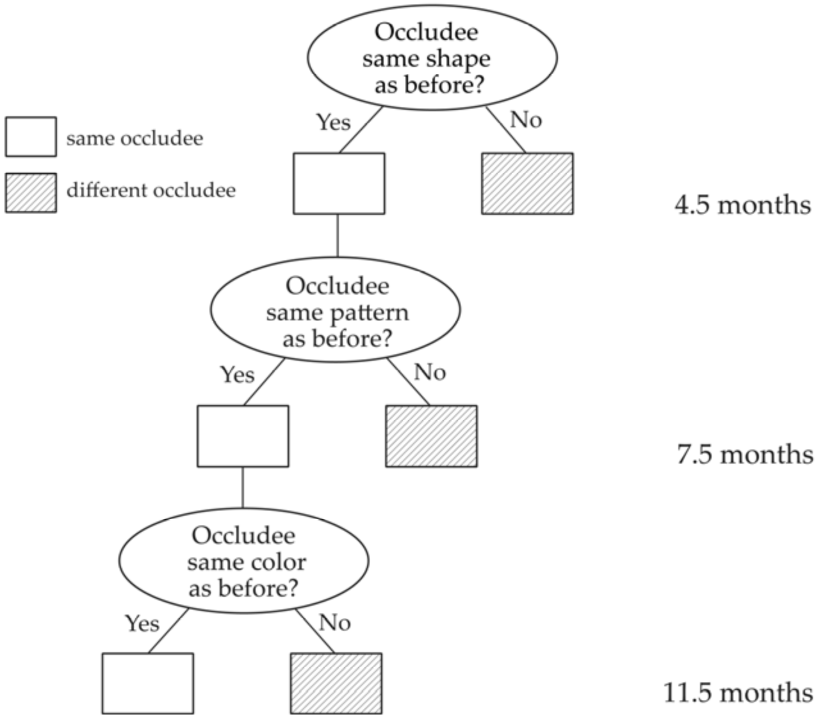
Figure 6: Schematic comparison of three occlusion event sequences, arranged from easiest to hardest in terms of the age at which infants succeed: A. infants succeed at 5.5 months (e.g., Wilcox & Baillargeon, 1998a; Wilcox & Schweinle, 2002); B. at 9 months, males succeed but females fail (Wilcox et al., in press); C. at 9.5 months, males and females both fail; at 10.5 months, male succeed but females fail (e.g., Wilcox, 2007; Wilcox & Baillargeon, 1998a). Infants succeed earlier with the occlusion sequence shown in C. if they are first given two piecemeal trials that isolate each object's trajectory. In the first piecemeal trial, the box emerges once to the left of the screen and returns behind it; in the second piecemeal trial, the ball emerges once to the right of the screen and returns behind it. Following these trials, 10.5-month-old females now succeed with the sequence in C., as do 7.5-month-old males (but not females; Wilcox, 2003, 2007).

Figure 7: Summary of object-individuation tasks where young infants correctly expect two objects when the screen is removed. **I. Spatiotemporal Information**: **Location Information**: a. the two objects are simultaneously visible at the start of the trial (Xu & Carey, 1996); **Path Information**: b. an object disappears behind one screen and reappears from behind another screen without appearing in the gap between them (Spelke et al., 1995a); **Speed Information**: c. an object disappears at one end of a screen and immediately reappears at the other end, too soon to have traveled the distance behind the screen, given its speed when in view (Wilcox & Schweinle, 2003). **II. Categorical Information: Ontological Categories**: d. infants see an inert object and a self-propelled agentive object (Surian & Caldi, 2010); e. infants see a humanlike and a non-humanlike object (Bonatti et al., 2002); **Functional Categories**: f. infants hear a rattling and a jingling object (Wilcox et al., 2006); g. infants see a hand rotate an object's dial to play a melody or pull an object's lever to flash lights (Futó et al., 2010); **Taxonomic Categories**: h. as the objects come into view, they are labeled "A duck!" and "A ball!" (Xu, 2002); i. infants receive two static priming trials involving three different cups and three different shoes, and then see one of the cups and one of the shoes emerge in alternation from behind a large screen (X. Li et al., 2011).

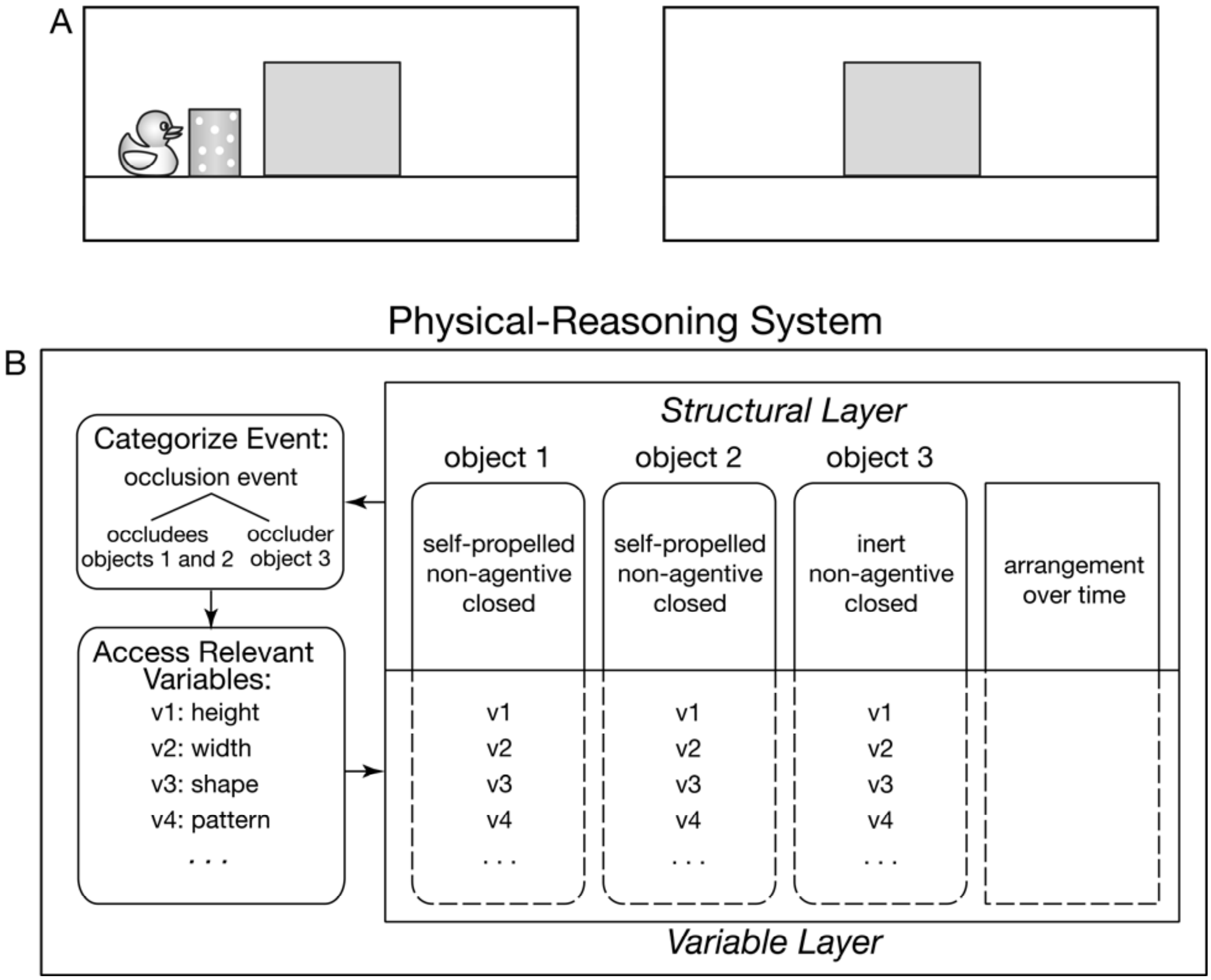
Figure 8: Examples of two-screen object-individuation tasks. **A.** Schematic drawing of the test events shown in Wilcox and Chapa (2002); copyright 2002 by Elsevier, adapted with permission of the authors; **B.** Schematic drawing of the test events shown in Hespos and Baillargeon (in Needham & Baillargeon, 2000); the shorter opaque screen used in the final phase differed from the taller opaque screen in color and pattern, to help draw infants' attention to the change.

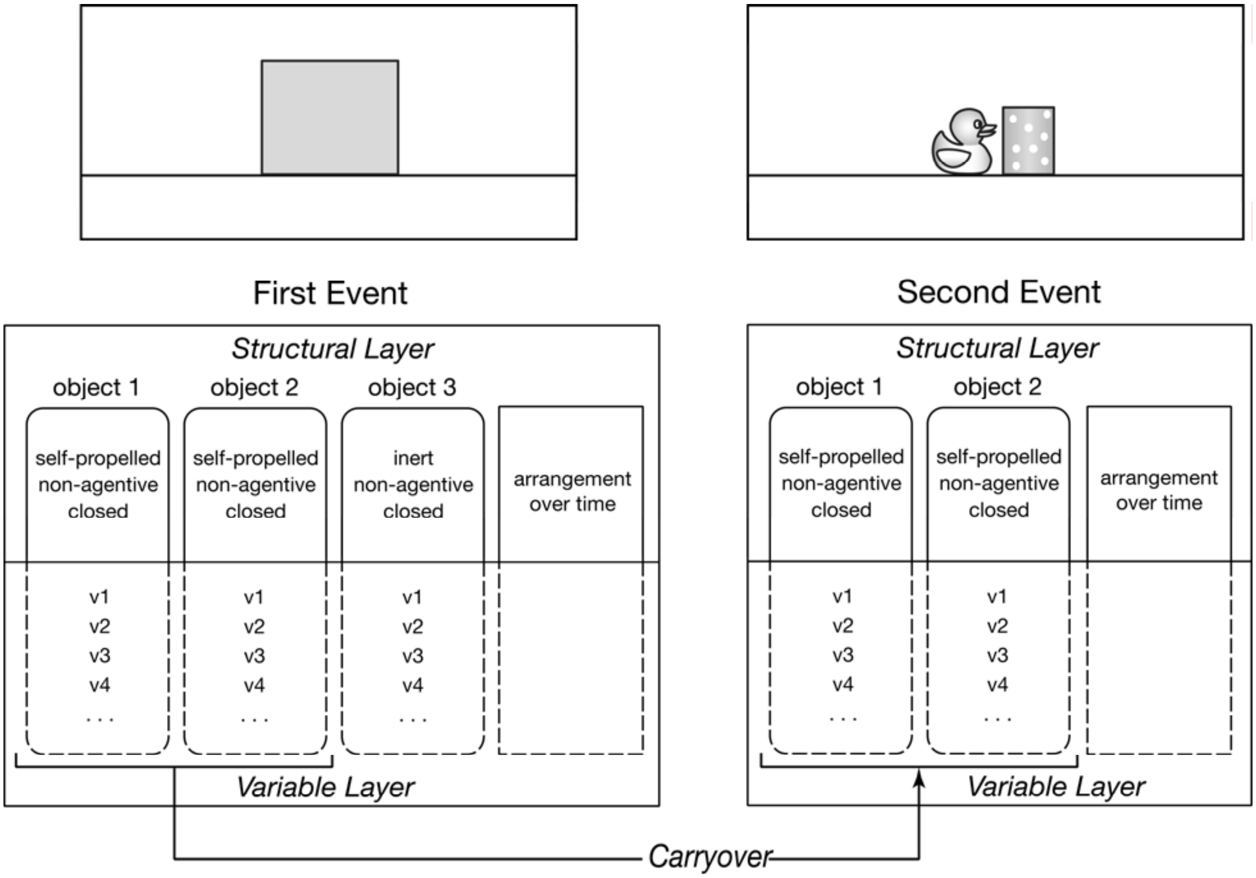


A. When is the occludee fully and continuously hidden behind the occluder?



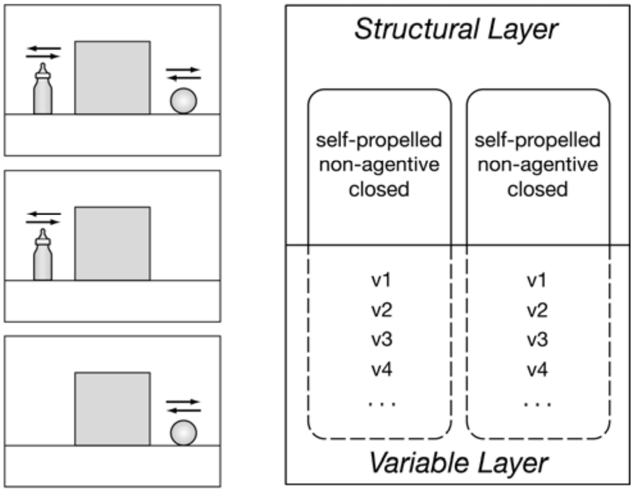
B. When is the occludee that reappears from behind the occluder the same occludee that disappeared?


A
B
Physical-Reasoning System
Categorize Event:
occlusion event
occludees
objects 1 and 2
occluder
object 3
Access Relevant Variables:
v1: height
v2: width
v3: shape
v4: pattern
. . .
Structural Layer
object 1
object 2
object 3
self-propelled
non-agentive
closed
self-propelled
non-agentive
closed
inert
non-agentive
closed
arrangement over time
v1
v2
v3
v4
. . .
Variable Layer

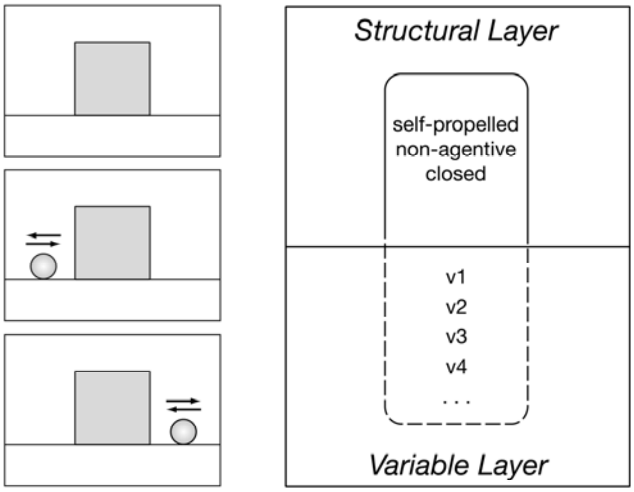

First Event
Second Event
Structural Layer
object 1
object 2
object 3
self-propelled
non-agentive
closed
self-propelled
non-agentive
closed
inert
non-agentive
closed
arrangement
over time
v1
v2
v3
v4
. . .
Variable Layer
Structural Layer
object 1
object 2
self-propelled
non-agentive
closed
self-propelled
non-agentive
closed
arrangement
over time
v1
v2
v3
v4
. . .
Variable Layer
Carryover

10-Month-Olds

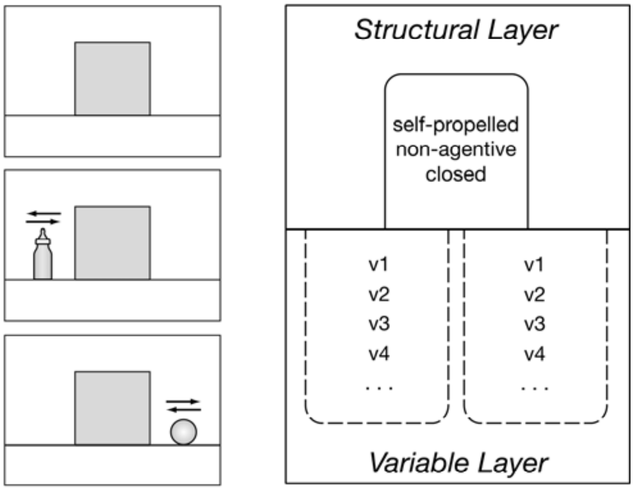
A. Spatiotemporal Condition



B. Single-Object Condition

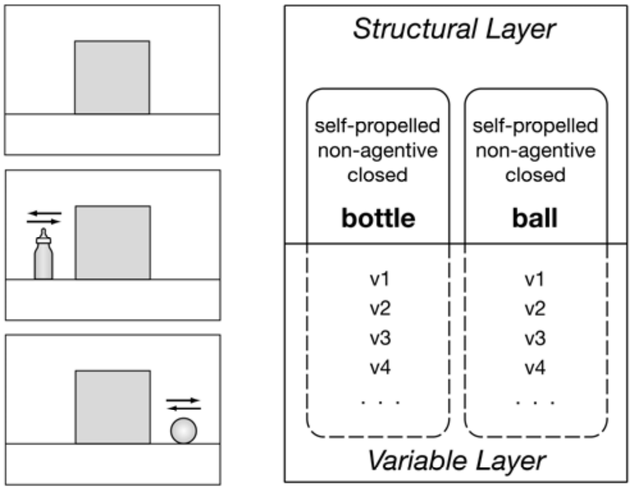


C. Categorical/Featural Condition

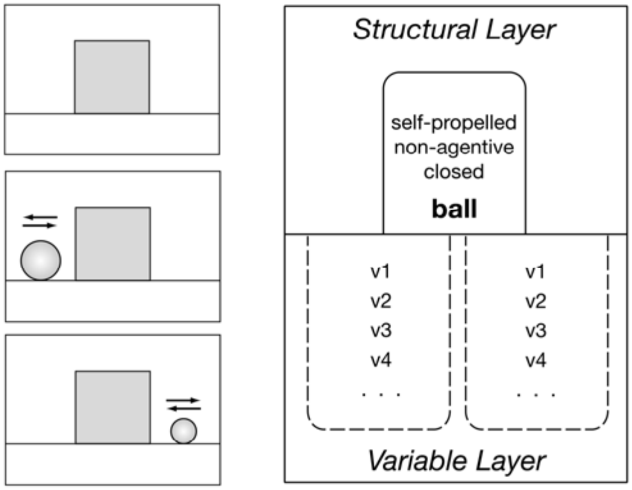


12-Month-Olds

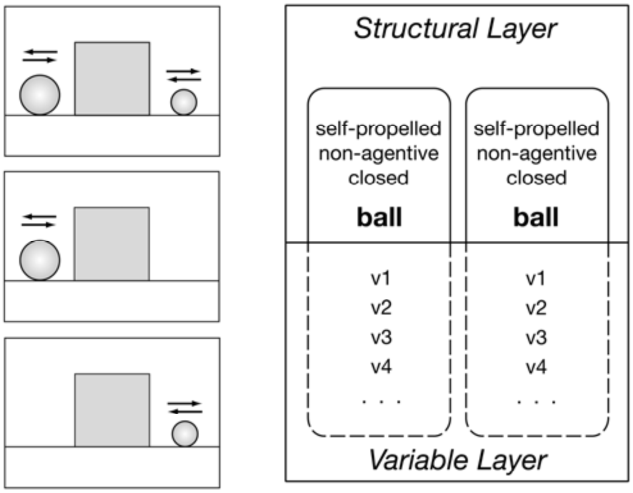
A. Categorical/Featural Condition: Different Object Categories

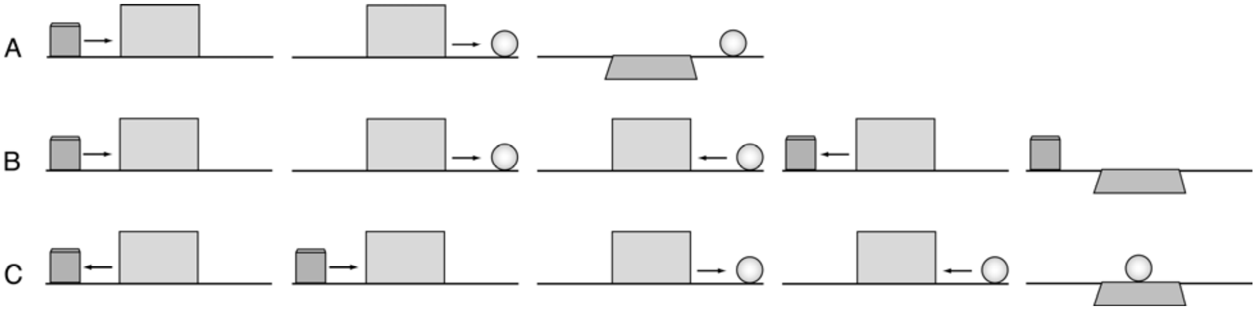


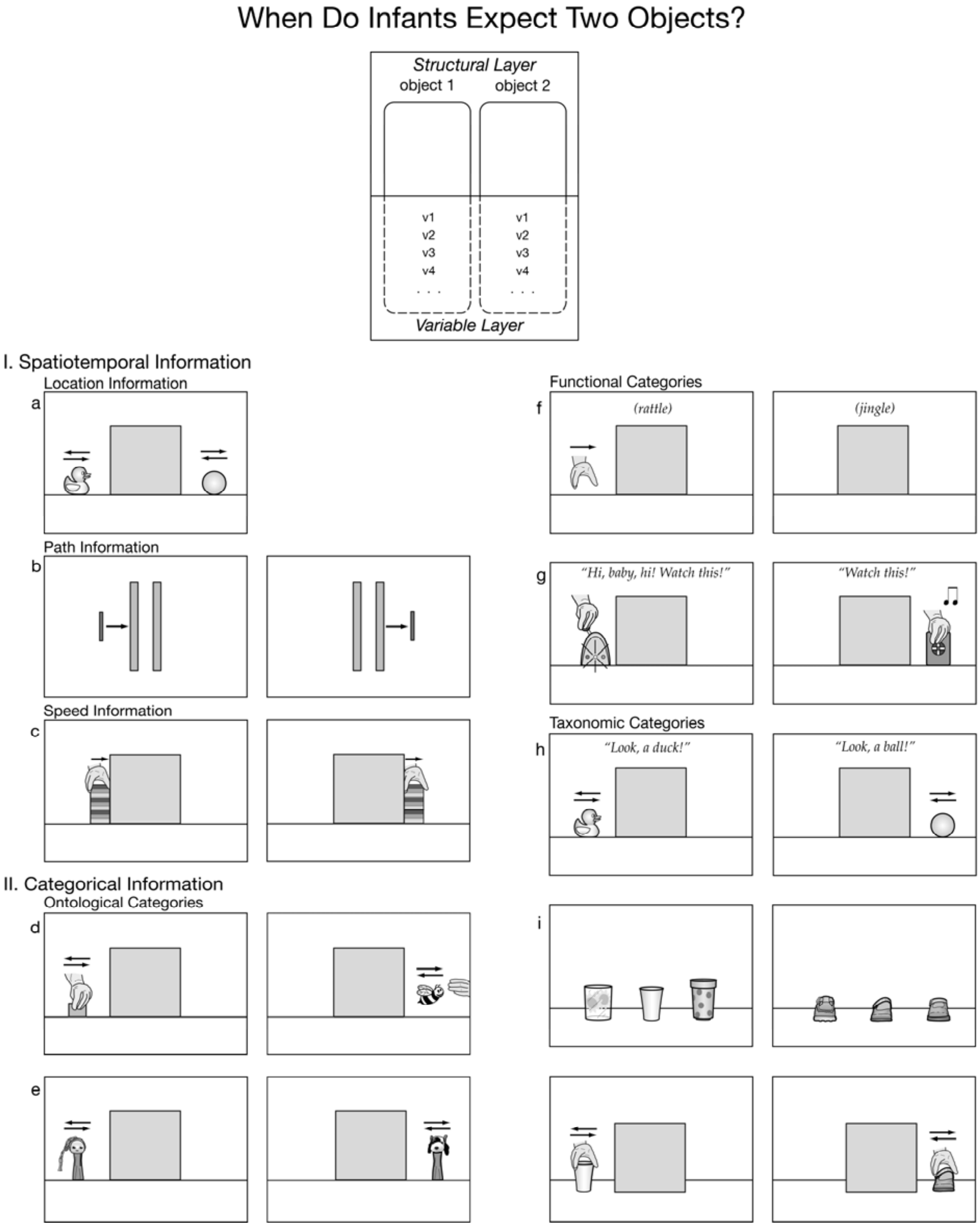
B. Categorical/Featural Condition: Same Object Category

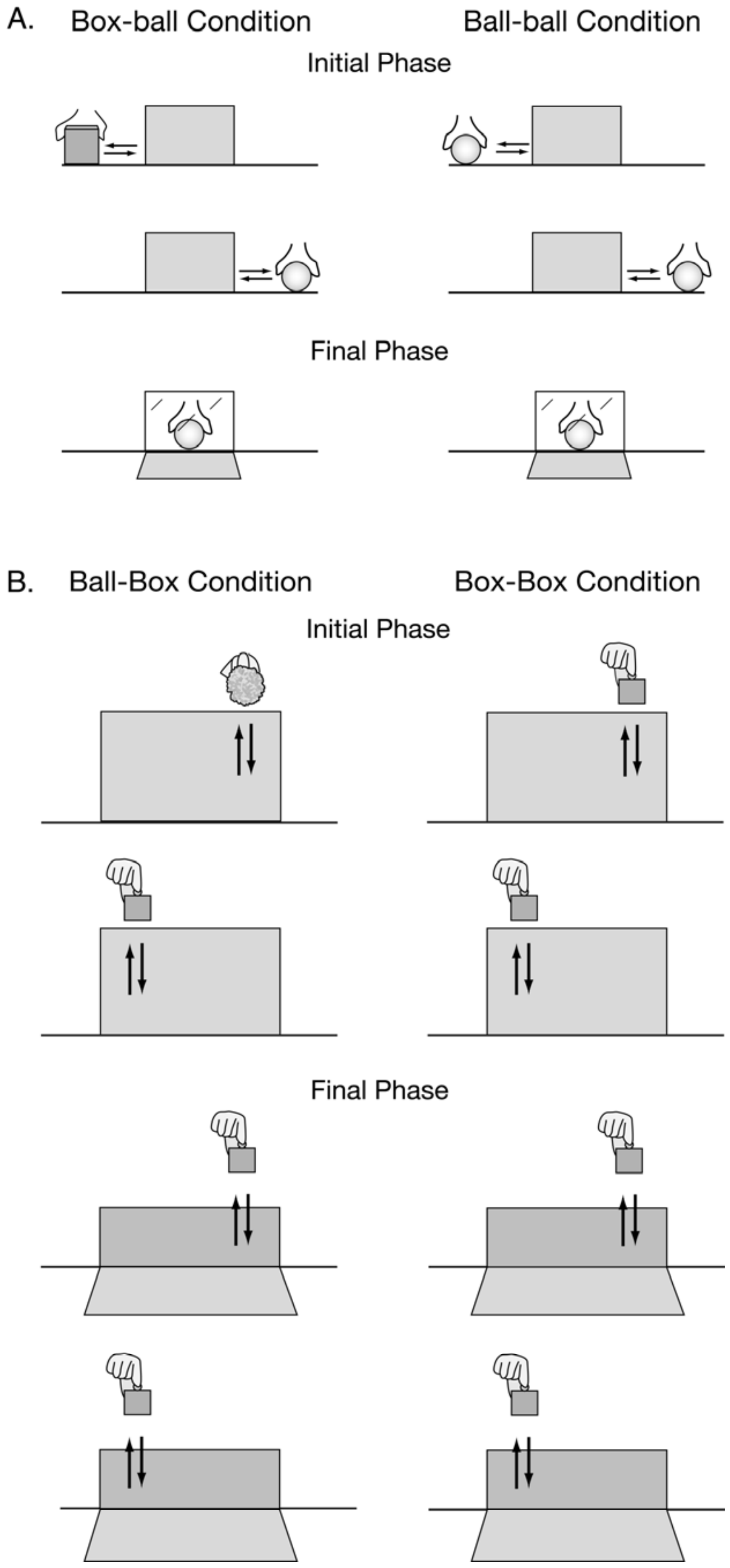


C. Spatiotemporal Condition: Same Object Category




A
B
C


When Do Infants Expect Two Objects?
Structural Layer
object 1
object 2
v1
v2
v3
v4
. . .
v1
v2
v3
v4
. . .
Variable Layer
I. Spatiotemporal Information
Location Information
a
Path Information
b
Speed Information
c
II. Categorical Information
Ontological Categories
d
e
Functional Categories
f
(rattle)
(jingle)
g
"Hi, baby, hi! Watch this!"
"Watch this!"
Taxonomic Categories
h
"Look, a duck!"
"Look, a ball!"
i


A.
Box-ball Condition
Ball-ball Condition
Initial Phase
Final Phase
B.
Ball-Box Condition
Box-Box Condition
Initial Phase
Final Phase