Evolution and Human Behavior 47 (2026) 106950

Contents lists available at ScienceDirect

Evolution and Human Behavior

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

# Minimal object representations in the infant mind: Structure, format, and inceptual role[☆]

Melissa M. Kibbe[a,*], Alan M. Leslie[b]

[a] Department of Psychological & Brain Sciences, Boston University, 64 Cummington Mall, Boston, MA 02215, United States of America
[b] Department of Psychology, Center for Cognitive Science, Rutgers University, 152 Frelinghuysen Rd., Piscataway, NJ, United States of America

ARTICLE INFO

*Keywords:*
Object representation
Object concept
Language of thought
Infancy
Face perception

ABSTRACT

What is the minimum amount of information needed to represent in the mind an object's existence? We describe studies with prelinguistic infants that support a basic format for object representations. We show that, while infants can represent an object's' identity, doing so is not a requirement for representing that object's existence. We argue that the evidence suggests that a *minimal* representation of an object – the *index* - is sufficient for this purpose. It is minimal in that it can contain no descriptive information about an object whatsoever. We suggest that richer content, including physical properties characteristic of a "Spelke object" such as solidity or cohesion, are not necessarily contained in an object representation. Instead, the minimal object representation, together with the operational characteristics of the indexing mechanism, form an *inceptual* kernel within a language of thought from which such richer object representations will develop.

As an organism moves through its environment, it encounters a world made up of other agents, substances, landscapes, and of course, objects. For humans, objects often are the tools we use, the things we talk about, and the targets of our actions, beliefs, and desires. Object representations allow us to navigate a three-dimensional environment, to think about things that are absent, to talk to each other about objects, to learn new object skills, and to take actions that allow us to complete our goals. In other words, we need to be able to represent and think about objects so that we can interact with the physical world and with each other in that physical world.

What do we represent about objects? Even the simplest objects yield a huge range of potentially "representable" information. Take, for example, the object in Fig. 1. Some of the information that we could potentially represent about this object is directly observable by perception, including the object's spatial location or its surface features like color, metric shape, texture, orientation, luminance properties, shading, or spatial frequencies. Other potentially representable information about this object is *not* directly observable by perception. For example, one could represent the category of the object at multiple levels (e.g., it is a token of the type OBJECT; it is a token of the type ROCK; it is a token of the type METAMORPHIC ROCK) and/or its corresponding verbal label (e.g., "rock"). One could represent the object's history with varying levels of detail (e.g., a rock from a beach; a rock that was picked up eight years ago on a beach in Plymouth, Massachusetts). One could represent the object's intended function (e.g., just a rock; a paperweight; a pretend racecar). The list goes on.

Just as voluminous are the cognitive processes that might use this information: it might be stored in short- or long-term memory, it might enter into numerous distinct lines of inference, action plans aplenty may be formed around it, it might be spoken of in any number of ways, become the subject of a poem "To a Rock", and so on.

Our point is that such lightly constrained maximalizing is unlikely to yield an advance for the theoretician. Instead, we ask: What is the *minimal* information about an object we must represent to maintain that object in mind? That is, if information were lost piece by piece from an object representation in memory, what's the last piece that must go such that the object would *then* have vanished from the mind?

A good way to determine the minimal format of an object representation is to tax a memory system and then observe what ends up "shaking out" in the representation. That is, when a resource-limited system (such as a human observer) faces a situation in which representational resources are spread thin (such as when the observer is tasked with representing several occluded objects simultaneously), the resulting representations are likely to be sparser than when

☆ This article is part of a Special issue entitled: 'Cognitive Development' published in Evolution and Human Behavior.

* Corresponding author.
*E-mail address:* kibbe@bu.edu (M.M. Kibbe).

https://doi.org/10.1016/j.evolhumbehav.2026.106950
Received 22 December 2025; Received in revised form 22 June 2026; Accepted 16 July 2026

**Fig. 1.** An object in the wild.

representational resources are less taxed. One could then probe these sparser representations to observe what information is prioritized over other information.

But doing so is difficult in adult human populations. Adults, when taxed, use a variety of strategies that help them overcome representational limits by "reformatting" representations into ones that are easier to maintain, especially verbal descriptions or formats that rely on adults' sophisticated conceptual knowledge or long-term memory (Baddeley & Hitch, 2019; Collins & Olson, 2014; Norris et al., 2018).

Preverbal infants, on the other hand, are an ideal population to bring to bear on the question of the minimal structure of an object representation. Like adults, infants' perceptual systems are fairly well developed, so they *perceive* the world of objects just fine. By around age 6 months, infants can perceive color, depth, motion, and other visual features like adults do (albeit nearsightedly; Dobson & Teller, 1978; Kavšek et al., 2012, Nawrot et al., 2009; Skelton et al., 2022; Yonas & Granrud, 2006). Also, around age 6 months, infants (like adults) have expectations about how objects should interact with each other mechanically (Leslie, 1982a; Leslie & Keeble, 1987). By at least 3.5 months, infants (like adults) expect an occluded object to continue to exist after it is obscured from view (Baillargeon, 1987; Baillargeon et al., 1985; Baillargeon & DeVos, 1991), suggesting they can represent an object in their minds and maintain that representation in the absence of continuous visual access to an object.

However, young infants are much more limited than adults in several important ways. Young infants have no expressive vocabulary, and have an extremely limited receptive vocabulary (Bergelson & Swingley, 2012). While young infants have some basic category knowledge (Mandler, 1992; Mandler & McDonough, 1998; Rakison & Yermolayeva, 2010), they certainly lack the vast sophisticated conceptual repertoire of adults with decades of experience with the world. They also have more limited general cognitive capacities that could support encoding and storing information about objects that are no longer in view, including attention (Colombo, 2001; Reynolds, 2015) and working memory (Fitch et al., 2016; Kibbe, 2015; Reynolds & Romano, 2016), and limited resources for reformatting incoming information using long-term memory (Behm et al., 2025; Rovee-Collier, 1993).

Infants, therefore, have much to tell us about the minimal structure of a representation of an object in the mind. We argue that the evidence suggests a format for a *minimal* representation of an object in the mind, one that is operational in infancy and beyond. We first sketch out what we mean by "minimal" object representations and distinguish these representations from the substantially richer "Spelke-object" representations typically posited to be the primitives of infants' object representational capacities. We suggest that our conception of a minimal object representation forms an inceptual kernel from which a richer object concept can develop.

## 1. Minimal representations of objects

The earliest representation of an object is an indexing operation that picks out an entity by pointing at its location and "sticking' to that entity as it moves through locations, including occluded locations (Leslie et al., 1998; Scholl & Leslie, 1999; Scholl, 2001; cf., Pylyshyn, 1994). A limited number of such indexes are available at any one time — available evidence for infants suggests up to about three (see Wang & Kibbe, 2024). These indexes operate like bare demonstratives within infants' language of thought.

The object index is a *minimal* representation, by which we mean two things. First, identifying information about the object (including its surface features, category, etc.) may optionally be bound to an index. Potentially then an object index may be enriched with all sorts of property information. This implies that in principle such bound information could be lost piece by piece and still the object be represented, just less richly. Second, should the index itself be lost, the object's existence is no longer represented. It has indeed, cognitively speaking, vanished. Thus, a bare index must be represented at minimum in order to represent an object in the mind – no other information need be represented. As we show below, evidence from infants suggests just such an architecture for object representation.

Crucially, our conception of the index is that it contains no discursive information. Representations in language of thought theory typically contain descriptive information: for example, representations of sensory features, such as size, shape, color; spatiotemporal information, such as distance, motion, duration; categorical information such as is-a-DOG, is-a-HUMAN, is-a-TRIANGLE; and so on. Sometimes these are called discursive representations (Fodor, 1975; Leslie, 1982b). The raw index, however, does not function discursively. Instead, it fulfills at least some of the (more primitive) functions of referring expressions in natural language like pure indexicals or bare demonstratives, for example, "*that!*" or indeed a 12-month-old's intentional pointing.

Importantly, this means that our characterization of a minimal object representation diverges from views that posit a richer initial representation for objects in infancy. Spelke (1990, 1994, 2022) argues for an innate format for object representations which include physical information about objects: that they are *solid* (i.e., not liquid or plasma), that they are *rigid* and will not deform without external pressure, that they are *cohesive* and will not come apart without external pressure, that they will *fall when unsupported*, and that they are *bounded* in space such that they cannot occupy the same space as another object, *cannot occupy multiple spaces at once*, and *cannot vanish and reappear*. Spelke suggests that these physical constraints are part of infants' core knowledge about the domain of objects, and that therefore infants have an innate object concept that includes such knowledge as (innate) conceptual constituents.

However, indexes need not contain *any* such physical information. We argue instead that the bare index plays the *inceptual* role in the infants' developing object concept – it serves to get the development of richer object concepts off the ground. Here we note that our argument is *not* that infants lack some or all of the above listed knowledge; to be very clear, we are open minded about the innateness of each of these items. Our point here is that the concept SPELKE-OBJECT is complex and far from atomic. For example, howsoever the infant comes by an understanding of *solidity* or *cohesion,* to take just these two properties, it is likely that the kinds of entities that attract object indexes in a typical environment will be both solid and cohesive, and moreover, probably be rigid, bounded and so forth; that is, they will likely turn out to be Spelke objects. To repeat, even if each of the listed constituent concepts is known innately, their assembly into the concept SPELKE-OBJECT may yet require learning.

To support our claim of a minimal object representation, in the next sections, we review work from our own labs and others' that leverage methods with prelinguistic infants to reveal the structure of object representations in the mind. We show that infants represent objects' surface features and the categorical identities of objects and show that the conditions under which infants represent such identifying information about objects reveal what kinds of information is prioritized in object representations. However, we also show that representing an object's identity is not a requirement for representing an object - infants can represent an object without representing any identifying information about what the object is, and indeed, without necessarily representing the object's physical properties (e.g., its solidity or rigidity). In light of this empirical evidence, we then detail the full object index model and its role in developing representational capacities in ontogeny and phylogeny.

## 2. Infants' representations of objects' surface features

An object's surface features are the properties of the object, such as color, texture, pattern, topology, metric shape, etc., that make up its visual appearance. Infants can detect featural distinctions between objects and will use some such features to individuate objects (Tremoulet et al., 2000; Wilcox, 1999; Wilcox & Chapa, 2004).

We have studied how infants use features to (re)identify individual objects by encoding such information into the object representation, which our model assumes is a process of binding to an index. Fig. 2 illustrates the general method we use. Infants first observe sets of multiple objects (two or more) with distinct featural identities (e.g., a disk and a triangle), which are then hidden sequentially, each behind its own occluder (e.g., first the disk is hidden, then the triangle is hidden), such that the objects are no longer visible to the infant. We then remove one of the occluders to reveal either the object that was hidden in that location originally (*control* outcome) or an object with a different identity than the one that was hidden (*change* outcome). We measure infants' gaze duration to these different outcomes, and compare infants' looking times between outcomes. If infants represent the featural identity of the hidden object, they should notice when the object's features have unexpectedly changed with no possible explanation, and will spend more time looking at the display (see Margoni et al., 2024; Stahl & Kibbe, 2022, for more on the violation-of-expectation method for assessing infant cognition).

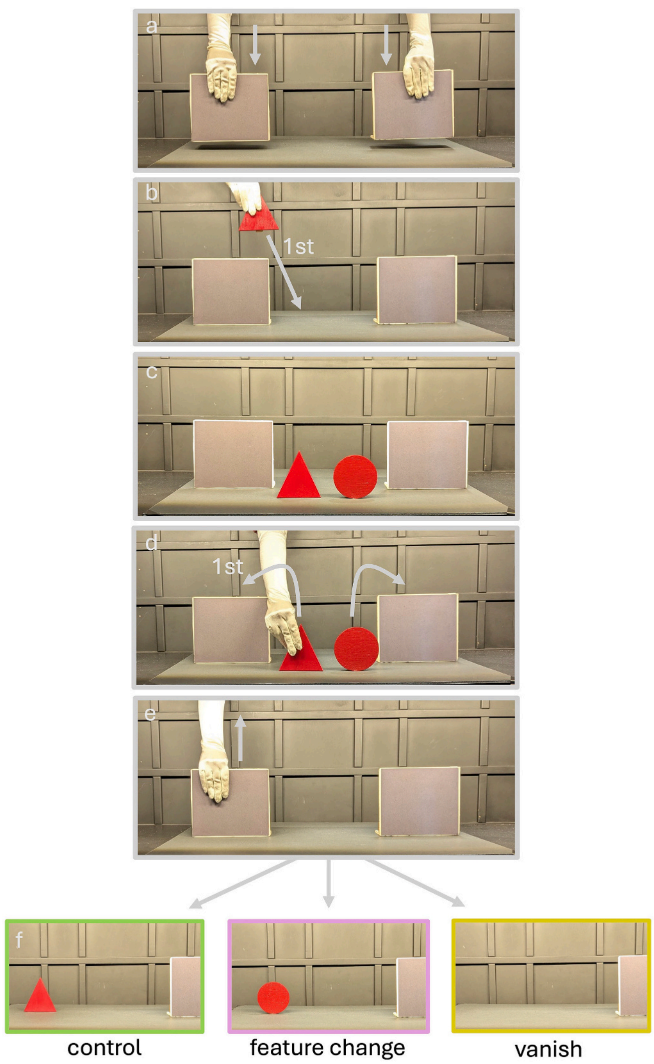


**Fig. 2.** An example sequence of events for a test trial in a study using the sequential hiding violation-of-expectation method. In this example, an experimenter places two occluders on the otherwise empty stage (panel a), then places two objects with different identities sequentially on the stage in front of the occluders (panel b). The infant is allowed to view the two objects for a few seconds (panel c), and then the experimenter places the objects behind the occluders one at a time (panel d). Finally, the experimenter removes one of the occluders (in this example, the occluder hiding the first-hidden object; panel e) to reveal one of several possible outcomes (panel f). Infants are seated in their caregiver's lap or in an infant chair facing the stage and observe these events live. The number of objects hidden, and which object is revealed, can be manipulated.

This method has several advantages for examining infants' object representations. Since each object is hidden in its own location, we can reveal each object separately, allowing us to more precisely measure infants' representations of each object without interference from other objects in the display. We also can manipulate the representational "load" for infants in two ways. First, we can manipulate the number of objects in the display, and examine how increasing the number of hidden objects impacts infants' representations of the featural identities of the objects. Second, we can manipulate which object we probe in the temporal hiding sequence. This allows us to compare infants' representations of the last-hidden object to their representations of earlier-hidden objects, the representations of which must be maintained while infants observe additional objects being hidden, allowing us to precisely examine how infants' representations of each object are impacted by representational load.

Using this method, we have observed that infants include surface features in their representations of objects. Six-month-olds who saw two featurally distinct objects hidden sequentially behind separate occluders looked longer when the last-hidden of the two objects was revealed to have changed its featural identity (e.g., when a triangle was hidden but a disk was revealed) compared to the control outcome (Káldy & Leslie, 2005). However, 6-month-olds failed to notice when the first-hidden of the two objects changed features (Káldy & Leslie, 2005; Kibbe & Leslie, 2011, 2016). Further, infants' failure to notice the change to the first-hidden object could not be attributed to the relatively longer duration between hiding and revealing the first-hidden, as compared to the last-hidden, object, because infants successfully noticed a feature change to a single object that was hidden for the same extended duration but with no intervening second-hidden object (Káldy & Leslie, 2005). Instead, the fact that infants' attention was drawn to the intervening hiding of the second, last-hidden object between the hiding and revealing of the first object appears to have driven infants' failure to include or maintain surface features in their representation of the first-hidden object.

Converging evidence for this comes from 9-month-olds. By 9 months of age, infants notice when *either* the last-hidden or the first-hidden of two objects has changed features, suggesting a developmental increase in infants' capacity to maintain featural information in their object representations (Káldy & Leslie, 2003; Kibbe & Leslie, 2013). However, 9-month-olds also show similar tradeoffs between the number of objects hidden and the number of objects for which they can retain featural information. When three featurally-distinct objects were hidden (e.g. first a square, then a triangle, last a disk), 9-month-olds failed to notice a change to the features of the second-to-last-hidden object, contrasting with their success at noticing a featural change to this object when only two objects are hidden in the display (Kibbe & Leslie, 2013). That is, when infants' attention was spread across more objects, infants encoded fewer featural details about the objects.

Infants' capacity to maintain surface featural information in their object representations continues to increase across the first three years of life, but this pattern of tradeoff between the number of hidden objects and the extent to which features are encoded for each object remains consistent (Kibbe & Applin, 2022; Kibbe & Leslie, 2013; Zosh & Feigenson, 2012) and can be influenced by the social context in which the objects are presented (for review, see Kibbe & Stahl, 2023a). This tradeoff also can be observed even into adulthood (e.g., Cowan et al., 2013; Wheeler & Treisman, 2002), in which studies show that successful representation of object features is contingent on there being sufficient attentional resources available to encode and/or maintain such information. Further, infants' pattern of behavior in our sequential hiding task suggests that when infants' representational resources are spread too thin, they default to representing the features of the *last* object that they observed, rather than representing nothing at all. This is consistent with adult work showing that, under attentional load, adults are reduced to "focal attention", able to dedicate attentional resources to only a single object in the display (Scholl, 2009).

## 3. Infants' representations of objects' categories

While infants' categorical knowledge is overall much more limited than adults', the ability to recognize and individuate conspecifics from non-conspecifics is early developing. Infants are sensitive to the distinction between human (or agent) and non-human (or non-agent) (Quinn & Eimas, 1998; Southgate et al., 2008; Spriet et al., 2022). By at least 10 months, infants use the categories HUMAN and NON-HUMAN to individuate objects. For example, infants who observed two different objects emerge one at a time behind an occluder (such that they never saw the objects together) expected there to be two objects behind the occluder when one of the objects had a human face and one did not, suggesting they used the distinctions in the identities of the objects to ascertain how many distinct individuals were present (Bonatti et al., 2002). However, when each object was from the same category but had different surface features (i.e., each object had a different face) infants failed to individuate the two objects (Bonatti et al., 2005), suggesting that infants were using the categorical, rather than the featural, identities of the objects to individuate them. Infants also failed to individuate an inverted face from a featurally-distinct non-human object (Bonatti et al., 2002), suggesting that infants do not categorize inverted faces as HUMAN (see also Otsuka et al., 2007; Turati et al., 2004).

We investigated whether infants encode such evolutionarily relevant categorical information into their representations of objects (Kibbe & Leslie, 2019). Using our sequential hiding method (described above), we showed 6-month-old infants two objects, a human face and a ball, that were distinguishable both by surface features and by category. We then hid the objects sequentially behind separate occluders. We were particularly interested in 6-month-old infants' representations of the *first-hidden* object, since our previous work showed that these infants failed to notice when the surface features of this object changed (e.g., infants failed to notice when a triangle was hidden and a disk revealed; Káldy & Leslie, 2005; Kibbe & Leslie, 2011). We therefore removed the occluder hiding the first-hidden object to reveal either the original object (e.g., the ball was hidden and the ball was revealed; or the face was hidden and the face was revealed) or that the object had changed (e.g., the ball was hidden, but the face was revealed; or the face was hidden and the ball was revealed) (see Fig. 3, left panel). We found that 6-month-old infants looked longer at the change outcome, suggesting that they had encoded the identity of the first-hidden object in their representation, and were able to maintain this information even as their attention was drawn to the second object being hidden. This was the first time we observed 6-month-old infants respond to an identity change to the first-hidden of two objects.

We conducted a second study to examine whether it was indeed the objects' distinct categories, rather than the objects' many surface featural differences, that impacted infants' ability to encode the objects' identities. A new group of infants observed the same objects as infants in the first experiment, except that the face was *inverted* (Fig. 3, middle panel). This allowed us to maintain the surface featural distinctions between the objects, while disrupting the categorical distinctions between the objects. We found that 6-month-olds' looking behavior was similar when the first-hidden object changed (e.g., a ball was hidden but an inverted face was revealed) and when the unchanged original object was revealed, consistent with our previous work that showed that 6-month-olds fail to represent surface features of the first-hidden of two objects (Káldy & Leslie, 2005; Kibbe & Leslie, 2011). That is, when the objects

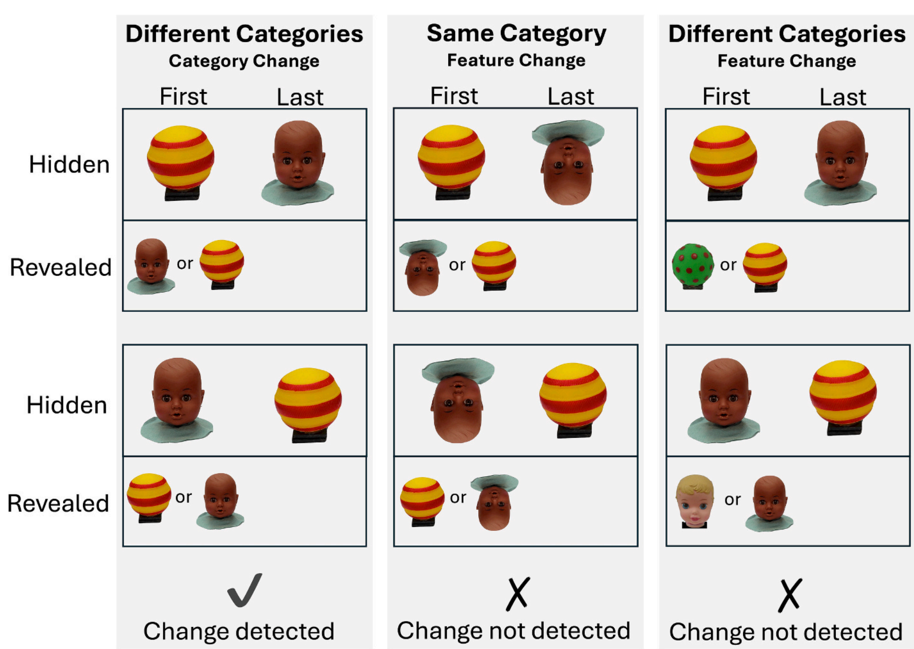


**Fig. 3.** Summary of the experimental conditions and results from Kibbe and Leslie (2019). In each experiment, two objects were hidden sequentially, and the occluder obscuring the first-hidden of the two objects was removed to reveal either the original hidden object or that the object had changed. Across experiments, we varied whether the two objects that were hidden were from different categories (face vs. object; left and right panels) or from the same category (inverted face vs. object; middle panel), and we varied whether the object was revealed to have changed categories (left panel) or to have changed features with no category change (middle and right panels). Within each experiment, we varied which object from the set was hidden first (top and bottom rows show the different hiding orders).

looked different, but were no longer categorically distinct, infants failed to represent the identity of the first-hidden object.

Finally, we asked whether the categorical differences between the objects, when available, may have supported infants' ability to represent the surface features of the objects. That is, when infants could recognize that one object was *human* and one was *non-human*, were they then able to better encode or maintain what the objects looked like? To answer this question, we again showed infants two objects from two distinct categories - the ball and the face. This time, we revealed that the first-hidden object had changed to a different object from the same category (e.g., a yellow and red striped ball was hidden and a green and blue polka-dotted ball was revealed; see Fig. 3, right panel). We again observed no differences in infants' looking behavior between the change and control outcomes.

These experiments show that, by age 6 months, infants can encode categorical information into their object representations without necessarily encoding specific surface features. Further, they suggest that infants' capacity to encode categorical identity into their object representations exceeds their capacity to encode surface featural information.

As infants gain more experience with the world, they begin to use newly acquired categories (e.g. about everyday objects, such as bottles and shoes) to individuate objects (see Baillargeon et al., 2012; Croteau et al., 2024; Xu, 2007). Recent work using EEG suggests that infants also may encode such categories into their representations of objects. Pomiechowska and Gliga (2021) found that 12-month-olds noticed between-category changes (e.g., a shoe to a brush), but not within-category changes (e.g., a shoe to a different shoe), to a single hidden object when the object was from a category that was familiar to them. Furthermore, infants' EEG patterns suggested that infants encoded familiar objects differently than unfamiliar objects even *before* the objects were occluded, suggesting that categorical identity supplanted featural identities at encoding rather than during maintenance or retrieval of the object representation from memory.

As infants develop, they become more flexible in the extent to which categorical identities supplant featural identities in their representations of objects. Kibbe and Stahl (2023b) found that 17-month-old infants could encode the surface features of objects from familiar categories into their representations, but only when those features were shown to be relevant to the task.

Together, these studies suggest that an object's categorical identity may be represented even when the objects' surface features are not and suggest a prioritized role for categorical identity in object representation that is already evident in prelinguistic infants.

## 4. Infants' representations of objects without their identities

Thus far, we have shown that infants can represent an object's surface features and/or an object's categorical identity, but sometimes they don't. When infants fail to represent an object's identity, what happens to their representation of that object? When infants lose the last piece of property information from an object representation that would allow them to re-identify that object, do they fail to represent the object entirely? Or, does an inkling of the object's existence remain, absent any identifying properties?

We addressed this question using our sequential hiding method (Kibbe & Leslie, 2011; Fig. 2). We showed 6-month-olds two differently shaped objects which we then hid sequentially behind separate occluders. Recall that 6-month-olds had previously been shown to fail to notice when the first-hidden of two objects changed featural identities (e.g. Káldy & Leslie, 2005). We therefore used this failure as an opportunity to examine whether infants represented the *existence* of the object even if they failed to represent what it looks like. To do so, we removed the occluder obscuring the first-hidden object and revealed either the expected originally hidden object, the unexpected other object, or that the object had vanished completely, and measured infants' gaze duration to these different outcomes. We found infants looked longer at the empty spot which should have been occupied by an object compared to both the control and feature change outcomes. That is, infants expected an object, even if they didn't know what the object should look like. Infants represented the object without its identifying features.

## 5. Do infants' object representations contain rich physical information?

The evidence we outline above suggests that infants can represent objects without the objects' specific identifying properties. In the "core knowledge" view of Spelke (1990, 2022), infants' object representations *necessarily* contain rich physical information about objects, including information about objects' cohesiveness, rigidity, solidity, and boundedness (for a similar argument for rich object representations grounded in perception rather than conception, see Bai et al., 2025). However, the evidence from infants suggests that these properties are *not* always represented whenever an object is. While infants do show sensitivity to violations of these properties in some contexts, infants also appear to tolerate such violations in other contexts.

To take one example, Hespos and Baillargeon (2001) showed 4.5-month-old infants events in which a tall, solid-looking cylinder was lowered behind a short curved occluder. In one condition, the object stopped when it reached the floor and therefore the top of the object remained visible above the occluder, consistent with the physical principles of solidity and rigidity (i.e., a solid object can't pass through another solid object nor undergo unexplainable deformation). In another condition, the object became fully occluded, a physical impossibility given its height and apparent solidity. They found that infants looked longer at the impossible outcome, suggesting that they expected the tall solid cylinder to remain visible once its bottom touched the stage floor. However, another group of infants observed the same events, except that the occluder was replaced with a container into which the object was lowered. This time, infants looked equally at the two outcomes. Importantly, the *only* difference between the conditions was whether the tall cylinder was lowered behind an occluder or into a container. Yet, 4.5-month-old infants appeared to expect the object to exhibit solidity and rigidity in one context and not in another. By 7.5 months of age, infants found the violation surprising in both contexts.

Baillargeon's group has identified many cases such as these, in which infants appear to detect a physical violation in some contexts, but tolerate such violations in contexts that are only slightly different (for reviews and theoretical accounts, see Baillargeon & Wang, 2002; Baillargeon, 2008; Lin et al., 2022; see also Wang, 2011; Goldman & Wang, 2019). That is, infants appear to include some physical properties in their representations of objects, *but they don't always do so* – contra the rich representation view, these physical properties are not necessarily represented anytime an object is.

Baillargeon and colleagues argue that these "decalages" in the kinds of information infants represent about objects in physical events (such as occlusion or containment) suggest protracted development of infants' physical reasoning abilities (Baillargeon, 2008; Hespos & Baillargeon, 2006, 2008). They argue that, through observation of their environments, infants learn how physical properties of objects apply to interactions between objects. We suggest that the minimal index representation plays a crucial role in this development by picking out the objects about which infants must learn and providing a means by which such properties, when identified, may be integrated into the object representation. We detail the indexing mechanism in the next section.

## 6. The format of object representations: the index

The empirical data we described above suggest a format for object representation that is best explained by the index model (Leslie et al., 1998), which we update and expand below (see Fig. 4). An "index" is a mental pointer to an object in a location. Like an arrow or a finger,

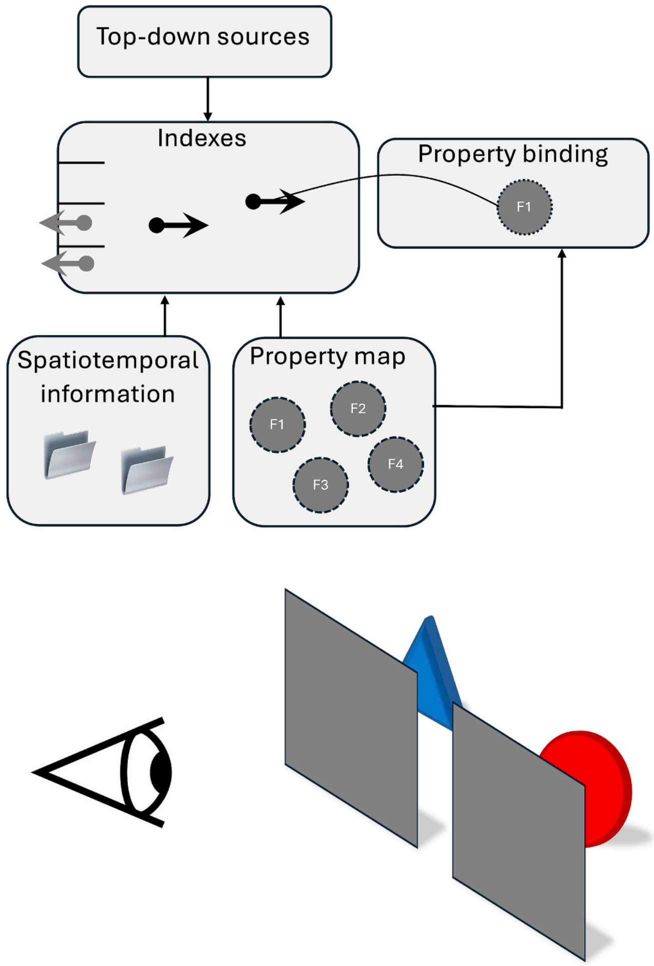


**Fig. 4.** The mature index model. Here we show three potential routes to the deployment of index representations: the spatiotemporal route (in which objects with unique spatiotemporal locations are indexed), the property map route (when spatiotemporal information is uninformative or ambiguous, the indexing system can use the property map to determine whether the an object has the same or different features as an object(s) encountered already; if different, the object is indexed), or the top-down route, in which an object may be inferred to be in a location, rather than directly observed. Once deployed, indexes can optionally have one or more properties bound to them. There are a fixed number of available indexes in the mature object indexing system. The diagram shows two indexes deployed via the spatiotemporal route, one of which has an identifying property bound to it, and two indexes still available.

indexes do not contain information about the object that they point to, but indicate the current location of the object in physical space. An index "sticks" to the object as it moves through space, much as a finger can continue to point to an object as it moves. That is, the index points to an object with a spatiotemporal location but does not point to the location itself. There are a fixed number of indexes available to represent objects (in adults, about 4; in infants and toddlers, about 3; see Wang & Kibbe, 2024, for review and meta-analysis).

An index is triggered by an encounter with an object. One possible route to an index happens when one encounters an object that has a distinct spatiotemporal location. The object can grab visual attention (triggering attention-based tracking via, e.g., Kahneman et al.'s (1992) object files or Pylyshyn's (1989) Fingers of INStantiation (FINSTs) – mechanisms by which objects with unique spatiotemporal locations can be picked out and tracked). These spatiotemporally defined representations can then elicit indexes, which can continue to represent an object even if the object moves into occlusion, representing an object's location as "somewhere behind the occluder".

The index model also accounts for situations in which an object cannot be individuated by its spatiotemporal location alone. For example, imagine that a red triangle vanishes behind an occluder and then a green disk emerges from the other side of the occluder. In this example, the available spatiotemporal information is consistent with a single object moving on a single trajectory, so spatiotemporal information alone is not sufficient to elicit a new index. Instead, the detection of new identifying properties may trigger a new index. To do so, the identifying properties of observed objects in the scene can be maintained in a "property map", which the indexing system can compare to incoming visual information and deploy a new index if new properties are detected. Importantly, this means that individuating an object by property need not involve binding property information to an index (which accounts for evidence showing that ~6-month-old infants can individuate at least two objects by shape, but only notice a change to the shape of a single object; Káldy & Leslie, 2005; Kibbe & Leslie, 2011; Wilcox, 1999).

Indexes also can be deployed to objects even if an image of the objects was *never* on the retinae. For example, when an object is inferred to be in a location rather than directly observed (e.g., the infant knows that there is an object hidden inside a container even if they did not every directly observe the object itself; Cheng & Kibbe, 2025, Cheng & Kibbe, 2026; Kibbe & Feigenson, 2014), infants use indexes to represent those objects. That is, indexes can be deployed anytime objects must be represented and their spatiotemporal locations be tracked, even if the route to these representations is not a purely "perceptual" one.

Once deployed, an index can have property information "bound" to it, but this binding occurs only after the index is assigned. That is, while *unbound* identifying properties of an object can be represented (e.g., in early visual cortex or in areas supporting semantic memory), binding these properties to a particular index is a second step that requires sustained attention (Bouvier & Treisman, 2010; Treisman, 1996). Importantly, binding of property information is not required in order to represent an object. However, it *is* required in order to later re-identify the object (e.g., by its surface features, category, or other identifying properties) or to make predictions about how such properties may dictate how an object interacts with another object (e.g., that a tall object should remain visible when lowered behind a short occluder).

The index is a "minimal" representation in that it represents nothing about the object it represents. Instead, indexes function as bare demonstratives in a language of thought, akin to "that!" in natural language. In natural language, the word "that" picks out an entity to which the speaker refers without containing any details about the referent itself. Much the same way, an index picks out an object without containing any details about the object itself. One might interject that in pointing to an object in a location a raw index must discursively represent *x,y,z* spatial co-ordinates. In Kahneman and Treisman's view just such discursive information is contained in their object file model (only it is written on the "tab" of the folder rather than contained within the file – that is, in their view, it defines the object file, while featural information is bound within the file itself). By contrast, in Pylyshyn's "fingers of instantiation" (FINST) model, coordinate information is unnecessary as long as a simple winner-takes-all network can find its way back to the retinotopic location of the indexed entity (Pylyshyn, 2000). Taking this position on the nature of the raw index, we can say it is a minimal representation because it can contain *zero* descriptive information.

While most of the things that are indexed in infants' environment will be three-dimensional physical objects, indexes are promiscuous – they pick out entities that are not solid, cohesive, bounded, rigid objects in the Spelke sense. Two-dimensional moving patches in a display can attract indexes (see Pylyshyn, 2001; Scholl & Pylyshyn, 1999), as can illusory objects (e.g. Wutz et al., 2016), objects that do not conform to Newtonian motion (Kominsky et al., 2017; Michotte, 1963), or even *sets* of individual objects if those objects appear to be grouped in some way (e.g., they move together along the same spatiotemporal trajectory, share identities, or tend to statistically co-occur; Feigenson & Halberda, 2004, 2008; Kibbe & Feigenson, 2016; Leslie & Chen, 2007; Zosh et al., 2011). However, if an object in the world loses cohesion, the object will no longer be indexed (Cheries et al., 2008; Scholl & Pylyshyn, 1999). This does not mean that an index *represents* the property of cohesiveness. Rather, just like a finger no longer clearly points to a specific entity if that entity breaks into pieces, the index will no longer indicate the spatiotemporal location of the object it represents, and it must be reassigned (perhaps to one of the pieces).

## 7. Evolution of developmental neurobiology of objecthood: a glimpse

The indexing mechanism implements the principle of spatiotemporal continuity of objecthood (Spelke, 1990, Spelke, 2022) operationally, without explicitly representing that principle; indeed it may be one of that principle's first functional embodiments. The divorce of spatiotemporal continuity from other physical principles, namely, contact mechanics, put forward in Leslie (1994) and developed in Scholl and Leslie (1999) was demonstrated by Santos (2004) in rhesus macaques (*Macaca mulatta*). Santos found that macaques are strikingly similar to human infants (and adults) on object tracking but were unable to pass similar tasks where contact-mechanical knowledge was required. She suggests that the former are modular and the latter not. These important architectural features of human cognition likely have an evolutionary history stretching at least from old world monkeys through the great apes. Before we conclude, however, that contact mechanics is an achievement only of our more recent ancestral species, we should note that according to Leslie and Keeble (1987) and Leslie (1994) the inception of contact mechanics is likely the Michotte module. So, whereas Santos (2004) advocated replacing the notion of Core Knowledge with Core Knowledges*,* we prefer Core Architecture (Leslie, 1994) a term that better encompasses the complexities of cognition in even young human infants and at the same time keeps open the many unexamined questions.

## 8. The inceptual role of minimal representations in the developing “object concept”

We argue that the index functions as a bare demonstrative in infants' (and adults') language of thought. The index can therefore serve as input into early emerging propositional sentences in a language of thought, allowing infants to think about objects, their relations, and their social significance.

Our view of a minimal structure for an object representation contrasts with views that posit richer content as the starting state for object representation (e.g., Spelke, 2022; Bai et al., 2025), in which objects are represented along with properties such as *solidity, cohesion, boundedness, rigidity*, and so forth. While we are open to the idea that infants may have concepts for these properties (perhaps innately), we suggest that the concept SPELKE OBJECT may yet be learned. The operational characteristics of the indexing architecture allow infants to pick out and represent objects (even when those objects are out of view), and to optionally bind relevant properties to those representations. While indexes can and do pick out non-Newtonian objects, most of the entities that are indexed by the young infant – those in the infants' immediate environment – will in fact be Spelke objects and exhibit the above italicized properties, and infants can learn that the properties stochastically predict each other. That is, the Spelke object is not atomic and primitive but an assembled complex concept, the result of a learning process which the indexing mechanism supports. The index, operating within infants' language of thought, plays the inceptual role in getting a more sophisticated object concept off the ground.

## CRediT authorship contribution statement

**Melissa M. Kibbe:** Writing – review & editing, Writing – original draft, Conceptualization. **Alan M. Leslie:** Writing – review & editing, Writing – original draft, Conceptualization.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

Baddeley, A. D., & Hitch, G. J. (2019). The phonological loop as a buffer store: An update. *Cortex, 112*, 91–106.
Bai, D., Hafri, A., Izard, V., Firestone, C., & Strickland, B. (2025). Core perception: Re-imagining precocious reasoning as sophisticated perceiving. *Behavioral and Brain Sciences*, 1–75.
Baillargeon, R. (1987). Object permanence in 3½-and 4½-month-old infants. *Developmental Psychology, 23*(5), 655.
Baillargeon, R. (2008). Innate ideas revisited: For a principle of persistence in infants' physical reasoning. *Perspectives on Psychological Science, 3*(1), 2–13.
Baillargeon, R., & DeVos, J. (1991). Object permanence in young infants: Further evidence. *Child Development, 62*(6), 1227–1246.
Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition, 20*(3), 191–208.
Baillargeon, R., Stavans, M., Wu, D., Gertner, Y., Setoh, P., Kittredge, A. K., & Bernard, A. (2012). Object individuation and physical reasoning in infancy: An integrative account. *Language Learning and Development, 8*(1), 4–46.
Baillargeon, R., & Wang, S. H. (2002). Event categorization in infancy. *Trends in Cognitive Sciences, 6*(2), 85–93.
Behm, L., Turk-Browne, N. B., & Kibbe, M. M. (2025). The ubiquity of episodic-like memory during infancy. *Trends in Cognitive Sciences, 29*(11), 1034–1047.
Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences, 109*(9), 3253–3258.
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*(4), 388–426.
Bonatti, L. L., Frot, E., & Mehler, J. (2005). What face inversion does to infants' counting abilities. *Psychological Science, 16*(7), 506–510.
Bouvier, S., & Treisman, A. (2010). Visual feature binding requires reentry. *Psychological science, 21*(2), 200–204.
Cheng, C., & Kibbe, M. (2025). *Human infants compute the successor function*. https://doi.org/10.21203/rs.3.rs-8253462/v1
Cheries, E. W., Mitroff, S. R., Wynn, K., & Scholl, B. J. (2008). Cohesion as a constraint on object persistence in infancy. *Developmental Science, 11*(3), 427–432.
Cheng, C., & Kibbe, M. M. (2026). *Infants represent objects they have never seen*. Manuscript under review.
Collins, J. A., & Olson, I. R. (2014). Knowledge is power: How conceptual knowledge transforms visual cognition. *Psychonomic Bulletin & Review, 21*(4), 843–860.
Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology, 52*(1), 337–367.
Cowan, N., Blume, C. L., & Saults, J. S. (2013). Attention to attributes and objects in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(3), 731.
Croteau, J., Cheries, E., & Xu, F. (2024). The development of kind concepts: Insights from object individuation. *Psychological Review, 133*(1), 134–155.
Dobson, V., & Teller, D. Y. (1978). Visual acuity in human infants: A review and comparison of behavioral and electrophysiological studies. *Vision Research, 18*(11), 1469–1483.
Feigenson, L., & Halberda, J. (2004). Infants chunk object arrays into sets of individuals. *Cognition, 91*(2), 173–190.
Feigenson, L., & Halberda, J. (2008). Conceptual knowledge increases infants' memory capacity. *Proceedings of the National Academy of Sciences, 105*(29), 9926–9930.
Fitch, A., Smith, H., Guillory, S. B., & Káldy, Z. (2016). Off to a good start: The early development of the neural substrates underlying visual working memory. *Frontiers in Systems Neuroscience, 10*, 68.
Fodor, J. A. (1975). *The language of thought*. New York: Crowell.
Goldman, E. J., & Wang, S. H. (2019). Comparison facilitates the use of height information by 5-month-olds in containment events. *Developmental Psychology, 55*(12), 2475.
Hespos, S. J., & Baillargeon, R. (2001). Infants' knowledge about occlusion and containment events: A surprising discrepancy. *Psychological Science, 12*(2), 141–147.
Hespos, S. J., & Baillargeon, R. (2006). Décalage in infants' knowledge about occlusion and containment events: Converging evidence from action tasks. *Cognition, 99*(2), 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*(1), 304–316.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*, 175–219.
Káldy, Z., & Leslie, A. M. (2003). Identification of objects in 9-month-old infants: Integrating 'what'and 'where'information. *Developmental Science, 6*(3), 360–373.
Káldy, Z., & Leslie, A. M. (2005). A memory span of one? Object identification in 6.5-month-old infants. *Cognition, 97*(2), 153–177.
Kavšek, M., Yonas, A., & Granrud, C. E. (2012). Infants' sensitivity to pictorial depth cues: A review and meta-analysis of looking studies. *Infant Behavior and Development, 35*(1), 109–128.
Kibbe, M. M. (2015). Varieties of visual working memory representation in infancy and beyond. *Current Directions in Psychological Science, 24*(6), 433–439.
Kibbe, M. M., & Applin, J. B. (2022). Tracking what went where across toddlerhood: Feature-location bound object representations in 2-to 3-year-olds' working memory. *Child Development, 93*(6), 1713–1726.
Kibbe, M. M., & Feigenson, L. (2014). Developmental origins of recoding and decoding in memory. *Cognitive Psychology, 75*, 55–79.
Kibbe, M. M., & Feigenson, L. (2016). Infants use temporal regularities to chunk objects in memory. *Cognition, 146*, 251–263.
Kibbe, M. M., & Leslie, A. M. (2011). What do infants remember when they forget? Location and identity in 6-month-olds' memory for objects. *Psychological Science, 22*(12), 1500–1505.
Kibbe, M. M., & Leslie, A. M. (2013). What's the object of object working memory in infancy? Unraveling 'what' and 'how many'. *Cognitive Psychology, 66*(4), 380–404.
Kibbe, M. M., & Leslie, A. M. (2016). The ring that does not bind: Topological class in infants' working memory for objects. *Cognitive Development, 38*, 1–9.
Kibbe, M. M., & Leslie, A. M. (2019). Conceptually rich, perceptually sparse: Object representations in 6-month-old infants' working memory. *Psychological Science, 30*(3), 362–375.
Kibbe, M. M., & Stahl, A. E. (2023a). Objects in a social world: Infants' object representational capacity limits are shaped by objects' social relevance. *Advances in Child Development and Behavior, 65*, 69–97.
Kibbe, M. M., & Stahl, A. E. (2023b). An object's categorizability impacts whether infants encode surface features into their object representations. *Infancy, 28*(5), 958–972.
Kominsky, J. F., Strickland, B., Wertz, A. E., Elsner, C., Wynn, K., & Keil, F. C. (2017). Categories and constraints in causal perception. *Psychological Science, 28*(11), 1649–1662.
Leslie, A. M. (1982a). The perception of causality in infants. *Perception, 11*(2), 173–186.
Leslie, A. M. (1982b). Discursive representation in infancy. In B. de Gelder (Ed.), *Knowledge and representation* (pp. 80–93). London: Routledge & Kegan Paul.
Leslie, A. M. (1994). ToMM, ToBy, and agency: Core architecture and domain specificity. In L. A. Hirschfeld, & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 119–148). Cambridge University Press.
Leslie, A. M., & Chen, M. L. (2007). Individuation of pairs of objects in infancy. *Developmental Science, 10*(4), 423–430.
Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition, 25*, 265–288.
Leslie, A. M., Xu, F., Tremoulet, P. D., & Scholl, B. J. (1998). Indexing and the object concept: Developing 'what' and 'where' systems. *Trends in Cognitive Sciences, 2*, 10–18.
Lin, Y., Stavans, M., & Baillargeon, R. (2022). Infants' physical reasoning and the cognitive architecture that supports it. *Cambridge Handbook of Cognitive Development*, 168–194.
Mandler, J. M. (1992). How to build a baby: II. Conceptual primitives. *Psychological Review, 99*(4), 587.
Mandler, J. M., & McDonough, L. (1998). On developing a knowledge base in infancy. *Developmental Psychology, 34*(6), 1274.
Margoni, F., Surian, L., & Baillargeon, R. (2024). The violation-of-expectation paradigm: A conceptual overview. *Psychological Review, 131*(3), 716.
Michotte, A. (1963). *The perception of causality*. Routledge (2017).
Nawrot, E., Mayo, S. L., & Nawrot, M. (2009). The development of depth perception from motion parallax in infancy. *Perception & Psychophysics, 71*(1), 194–199.
Norris, D., Butterfield, S., Hall, J., & Page, M. P. (2018). Phonological recoding under articulatory suppression. *Memory & Cognition, 46*(2), 173–180.
Otsuka, Y., Nakato, E., Kanazawa, S., Yamaguchi, M. K., Watanabe, S., & Kakigi, R. (2007). Neural activation to upright and inverted faces in infants measured by near infrared spectroscopy. *Neuroimage, 34*(1), 399–406.
Pomiechowska, B., & Gliga, T. (2021). Nonverbal category knowledge limits the amount of information encoded in object representations: EEG evidence from 12-month-old infants. *Royal Society Open Science, 8*(3), Article 200782.
Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition, 32*(1), 65–97.
Pylyshyn, Z. (1994). Some primitive mechanisms of spatial attention. *Cognition, 50*(1–3), 363–384.
Pylyshyn, Z. W. (2000). Situating vision in the world. *Trends in Cognitive Sciences, 4*, 197–207.
Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition, 80*(1–2), 127–158.
Quinn, P. C., & Eimas, P. D. (1998). Evidence for a global categorical representation of humans by young infants. *Journal of Experimental Child Psychology, 69*(3), 151–174.
Rakison, D. H., & Yermolayeva, Y. (2010). Infant categorization. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(6), 894–905.
Reynolds, G. D. (2015). Infant visual attention and object recognition. *Behavioural Brain Research, 285*, 34–43.
Reynolds, G. D., & Romano, A. C. (2016). The development of attention systems and working memory in infancy. *Frontiers in Systems Neuroscience, 10*, 15.
Rovee-Collier, C. (1993). The capacity for long-term memory in infancy. *Current Directions in Psychological Science, 2*(4), 130–135.
Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition, 80*(1–2), 1–46.
Scholl, B. J. (2009). What have we learned about attention from multiple object tracking (and vice versa). *Computation, Cognition, and Pylyshyn*, 49–78.
Santos, L. R. (2004). Core knowledges: A dissociation between spatiotemporal knowledge and contact-mechanics in a non-human primate? *Developmental Science, 7*(2), 167–174.
Scholl, B. J., & Leslie, A. M. (1999). Explaining the infant's object concept: Beyond the perception/cognition dichotomy. In E. Lepore, & Z. Pylyshyn (Eds.), *What is cognitive science?* (pp. 26–73). Oxford: Blackwell.
Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology, 38*(2), 259–290.
Skelton, A. E., Maule, J., & Franklin, A. (2022). Infant color perception: Insight into perceptual development. *Child Development Perspectives, 16*(2), 90–95.
Southgate, V., Csibra, G., Kaufman, J., & Johnson, M. H. (2008). Distinct processing of objects and faces in the infant brain. *Journal of Cognitive Neuroscience, 20*(4), 741–749.
Spelke, E. (2022). *What babies know: Core knowledge and composition volume 1. 1*. Oxford University Press.
Spelke, E. S. (1990). Principles of object perception. *Cognitive Science, 14*(1), 29–56.
Spriet, C., Abassi, E., Hochmann, J. R., & Papeo, L. (2022). Visual object categorization in infancy. *Proceedings of the National Academy of Sciences, 119*(8), Article e2105866119.
Stahl, A. E., & Kibbe, M. M. (2022). Great expectations: The construct validity of the violation-of-expectation method for studying infant cognition. *Infant and Child Development, 31*(6), Article e2359.
Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology, 6*(2), 171–178.
Tremoulet, P. D., Leslie, A. M., & Hall, D. G. (2000). Infant individuation and identification of objects. *Cognitive Development, 15*(4), 499–522.
Turati, C., Sangrigoli, S., Ruely, J., & de Schonen, S. (2004). Evidence of the face inversion effect in 4-month-old infants. *Infancy, 6*(2), 275–297.
Wang, J., & Kibbe, M. M. (2024). "Catastrophic" set size limits on infants' capacity to represent objects: A systematic review and Bayesian meta-analysis. *Developmental Science, 27*(4), Article e13488.
Wang, S. H. (2011). Priming 4.5-month-old infants to use height information by enhancing retrieval. *Developmental Psychology, 47*(1), 26.
Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General, 131*(1), 48.
Wilcox, T. (1999). Object individuation: Infants' use of shape, size, pattern, and color. *Cognition, 72*(2), 125–166.
Wilcox, T., & Chapa, C. (2004). Priming infants to attend to color and pattern information in an individuation task. *Cognition, 90*(3), 265–302.
Wutz, A., Drewes, J., & Melcher, D. (2016). Nonretinotopic perception of orientation: Temporal integration of basic features operates in object-based coordinates. *Journal of Vision, 16*(10), 3.
Xu, F. (2007). Sortal concepts, object individuation, and language. *Trends in Cognitive Sciences, 11*(9), 400–406.
Yonas, A., & Granrud, C. E. (2006). Infants' perception of depth from cast shadows. *Perception & Psychophysics, 68*(1), 154–160.
Zosh, J. M., & Feigenson, L. (2012). Memory load affects object individuation in 18-month-old infants. *Journal of Experimental Child Psychology, 113*(3), 322–336.
Zosh, J. M., Halberda, J., & Feigenson, L. (2011). Memory for multiple visual ensembles in infancy. *Journal of Experimental Psychology: General, 140*(2), 141.