This is an Accepted Manuscript for ***Behavioral and Brain Sciences***. This Target Article has been accepted for publication but not yet copyedited or typeset, and so may be subject to change during the production process. The article is considered published and may be cited using its DOI 10.1017/S0140525X25102756

"Core Perception": Re-imagining Precocious Reasoning as Sophisticated Perceiving

Dawei Bai[1,2]*, Alon Hafri[3], Véronique Izard[4], Chaz Firestone[5], Brent Strickland[1,6]*

[1] École Normale Supérieure, Département d'Études Cognitives, PSL Research University, Institut Jean Nicod (ENS, EHESS, CNRS), Paris, France

[2] Yale University, Department of Psychology, New Haven, CT 06520-8047, USA

[3] Department of Linguistics & Cognitive Science, University of Delaware, Newark, DE 19716, USA

[4] Université Paris Cité, CNRS, Integrative Neuroscience and Cognition Center, F-75006 Paris, France

[5] Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD 21218, USA

[6] Africa Business School and The School of Collective Intelligence. UM6P, Rabat, Morocco

***Corresponding authors' email:** dawei.bai@hotmail.com and brent.strickland@ens.fr

Keywords: cognitive architecture, core knowledge, infant cognition, perception, visual cognition

Short Abstract

"Core knowledge", which comprises early-emerging representations about the physical and social world, is canonically considered to be *conceptual* in nature, reflecting infants' *reasoning* abilities. Here we present an alternative view — "core perception" — positing that many core representations are inherently *perceptual*. This view is supported by evidence showing that core-knowledge representations of physics, geometry, number, and social agents also operate in adult vision, displaying empirical signatures of genuine perceptual processes. We argue that the best explanation of such overlap is that many infant results appealing to precocious reasoning actually reflect sophisticated perception, reconceptualizing a fundamental aspect of the origins of human knowledge.

Long Abstract

"Core knowledge" refers to a set of cognitive systems that underwrite early representations of the physical and social world, appear universally across cultures, and likely result from our genetic endowment. Although this framework is canonically considered as a hypothesis about early-emerging *conception* — how we think and reason about the world — here we present an alternative view: that many such representations are inherently *perceptual* in nature. This "core perception" view explains an intriguing (and otherwise mysterious) aspect of core-knowledge processes and representations: that they also operate in adults, where they display key empirical signatures of perceptual processing. We first illustrate this overlap using recent work on "core physics", the domain of core knowledge concerned with physical objects, representing properties such as persistence through time, cohesion, solidity, and causal interactions. We review evidence that adult vision incorporates exactly these representations of core physics, while also displaying empirical signatures of genuinely perceptual mechanisms, such as rapid and automatic operation on the basis of specific sensory inputs, informational encapsulation, and interaction with other perceptual processes. We further argue that the same pattern holds for other areas of core knowledge, including geometrical, numerical, and social domains. In light of this evidence, we conclude that many infant results appealing to precocious reasoning abilities are better explained by sophisticated perceptual mechanisms shared by infants and adults. Our core-perception view elevates the status of perception in accounting for the origins of conceptual knowledge, and generates a range of ready-to-test hypotheses in developmental psychology, vision science, and more.

1. Introduction

Certain aspects of human experience are universal. Physical objects, numerical quantities, geometrical layouts, and agents with mental states are all part of the reality that humans everywhere encounter. These entities exist and behave following certain regularities. For instance, physical objects move continuously through space and time (rather than teleporting between locations), agents have goals that they act to achieve (rather than behaving erratically), quantities of objects combine and divide in ways that respect mathematical principles, and so on. In accordance with such environmental regularities, the mind appears to be equipped with a set of "core knowledge" systems, thanks to which people are sensitive to physical, social, numerical, and geometrical features of the world, even in the absence of extensive exposure (Carey, 2011; Spelke, 2022; Spelke & Kinzler, 2007). Thus, preverbal infants react when hidden objects are shown to have spontaneously disappeared (Baillargeon, 1987; Baillargeon et al., 1985), when the number of objects revealed behind an occluder does not match the number of objects that initially went behind it (McCrink & Wynn, 2004), or when agents act irrationally relative to an obvious goal (Gergely et al., 1995; Onishi & Baillargeon, 2005). Such core knowledge is thought to be conserved by natural selection and to be part of species' typical phenotypic outcomes. These systems serve as a foundation for early learning, guiding preferences and choices from an early age.

From the mountains of empirical data collected on young infants over the last 40 years, the existence of rich and early-emerging knowledge is increasingly accepted by the field of psychology. But what is the nature of this knowledge, and where is its place in the architecture of the mind? On a canonical and highly influential interpretation of these data, core knowledge is inherently *cognitive*, such that infants' sensitivity to physical principles, the social world, number, and geometry results from precocious abilities for *reasoning* about these entities. According to this view, the types of processes underlying such representations are similar to those recruited when adults reason about sophisticated notions such as the value of an item in a marketplace or the intricacies of its material composition (an analogy given by Spelke, 1988).

In contrast to this view, here we argue that, rather than constituting a form of precocious *reasoning*, much of core knowledge is actually an intrinsic part of human *perception*. So, rather than accounting for how infants represent the behavior of objects by analogy with how adults might reason about, for example, a complex investment decision, we claim that infants represent physical objects and other "core" entities in ways that are similar to how perception (in infants and adults alike) represents the color, shape, or size of objects. On our alternative view, then, much of core knowledge consists in the computational principles that govern how perceptual processes operate and constrain the kinds of outputs they automatically generate. Put succinctly, instead of thinking of core knowledge as a primitive form of reasoning, we should think of much of it as a sophisticated form of perceiving.

Why think this? One reason is a conspicuous and intriguing suite of findings that has emerged over the course of the core-knowledge research program but that has managed to go relatively unexplained: strikingly overlapping patterns of performance and error in both infant "cognition" and adult perception. For example, the same cohesion violations (i.e., objects no longer maintaining a single bounded contour) that cause increased looking times in infants also cause adults to lose track of objects they are attending to; the same causal launching events that infants are sensitive to (such as two-object collisions) also drive adults' detection performance in vision tasks; and so on for many similar cases discussed below. Our alternative view provides an explanation for this conspicuous overlap between infants and adults: Because perceptual processes remain active throughout the lifespan, signatures of core knowledge should indeed persist into adulthood. We furthermore predict that core-knowledge processes should display behavioral signatures associated with other unambiguously perceptual mechanisms. That is, they should be fast, automatic, input-sensitive, informationally encapsulated, and interact with other perceptual processes (Fodor, 1983; Hafri & Firestone, 2021; Scholl & Gao, 2013) — properties that are not typically associated with reasoning processes (especially when these signatures are considered as a group).

To start, we provide a historical overview of theories that account for early forms of knowledge by appealing to precocious reasoning abilities (Section 2). We then lay out our alternative view, and argue that it disrupts, reshapes, and ultimately enriches a long-standing perspective from developmental psychology about the nature of core abilities (Section 3). To illustrate how our view is supported by empirical data, we focus initially on physical objects and their relations, or "core physics" (Section 4). We next show how this framework can be extended to the other core domains of geometry, number, and social agents (Section 5). We then address theoretical objections against our framework that arise separately from the positive empirical evidence (Section 6). We close by discussing how this reframing of early-emerging knowledge and capacities carves a path for exciting new lines of research across cognitive science (Section 7).

2. The core-knowledge hypothesis

2.1 Historical overview and core knowledge

Historically, views on the origins of human knowledge have been dominated by a handful of competing theories. On one hand, theories associated with the empiricist tradition see the infant mind starting as a "blooming, buzzing confusion" (James, 1890/1981), with richer representations developing with accumulating perceptual experience. Relatedly, according to Piaget's constructivist theory, the infant mind initially lacks sophisticated cognitive capacities and representations, leaving experience as the primary driver in creating such representations (Piaget, 1954). Today, however, many conclusions from this early work are rejected on the grounds that the tasks used to support these claims require capacities beyond those the researchers had intended to study (Lin et al., 2022). For example, Piaget claimed that infants lack representations of objects that disappear from view, based on the observation that infants younger than 8–12 months systematically fail to reach for desirable objects once they become occluded. However, infants may very well have been capable of representing occluded objects, but not yet capable of integrating those representations with the motor plans required to perform a reaching action.

In contrast to the empiricist tradition, the core-knowledge view postulates that humans are born with a rich but constrained set of representational systems: core-knowledge systems for representing objects, geometry, number, and aspects of the social world (Carey, 2009; Spelke et al., 1992; Spelke & Kinzler, 2007). These systems are thought to have developed through evolutionary processes, to be universal across humankind, and to be shared by various other species. This view has been supported in large part by empirical studies employing violation-of-expectation (VOE) and habituation methods, making elegant use of looking-time data (Baillargeon et al., 1985; Margoni et al., 2024). These methods are based on the principle that infants look longer at unexpected events — events that violate their expectations about the world, or that differ from a prior habituation phase. Since these methods rely primarily on eye or head movements, they are less demanding than Piagetian tasks in terms of executive resources and motor planning. Studies employing these innovative methods have revealed, for instance, that 2.5-to-5-month-olds look longer at events violating fundamental physical laws, such as two solid objects passing through one another or occluded objects ceasing to exist (Baillargeon, 1987; Baillargeon et al., 1985), contradicting the earlier Piagetian picture. In the decades that followed, these methods likewise have revealed that young infants possess a set of sophisticated mental representations not only about physics, but also about geometrical, numerical, and social aspects of the world around them (Spelke & Kinzler, 2007; although for leading alternative views, see Blumberg & Adolph, 2023; Karmiloff-Smith, 2009).

2.2 Core knowledge is postulated to be *cognitive*

What is the nature of these core-knowledge representations? One foundational distinction in characterizing mental processes and representations is that between *perception* and *cognition*. Accounts of this distinction vary in their details (Clarke & Beck, 2023), but most tend to rest on differences in inputs, outputs, and procedures for generating outputs based on inputs (see especially Fodor, 1983). Perception is traditionally conceived as a process that takes a narrow range of inputs directly from sensory apparati (or other perceptual processing stages), automatically applies stereotyped and inflexible "inference" rules over those inputs, and generates outputs which are fed into other perceptual modules or downstream cognitive systems. At the other end of the spectrum, cognition is conceived of quite differently: Purely cognitive processes (e.g., that might underlie planning a vacation) typically operate over a large range of input types, apply rules or heuristics which can be flexibly altered by background knowledge, expectations, or desires, and generate outputs that are globally available for action, inference, linguistic expression, or subsequent chains of thought. We expound upon these criteria in Section 3.2, but the above discussion may be sufficient for an initial grasp of the distinction.

Core-knowledge theorists' views about the nature of core systems have evolved over time. In the early days of core-knowledge theory, many researchers posited that core knowledge is exclusively conceptual.[1] For instance, Spelke (1988) held that "objects are *conceived*: Humans come to know about an object's unity, boundaries, and persistence in ways like those by which we come to know about its material composition or its market value." Spelke articulated this view in opposition to the alternative, that "objects are *perceived*: that humans come to know about an object's unity, boundaries, and persistence in ways like those by which we come to know about its brightness, color, or distance." Terms used by other researchers in this early literature bore similar themes: Infants "understand" that objects continue to exist when hidden (Baillargeon, 1986; Baillargeon et al., 1985; Hood & Willatts, 1986), "judge" that an object cannot pass through a screen (Baillargeon, 1987), and so on.

The 1990s brought further refinement to the core-knowledge view. Many discussions allowed that some primitive perceptual processing must trigger or enable the engagement of core-knowledge processes (Carey, 2009; Carey & Spelke, 1996; Carey & Xu, 2001), because "if human reasoning depends on domain-specific knowledge systems, then reasoners face a crucial task: They must single out the entities to which each system of knowledge applies" (Carey & Spelke, 1994). Early perceptual processes were thus entrusted with the function of identifying specific classes of "core" entities (objects, agents, etc.). These entities then supposedly fed into the core systems, which still were construed as cognitive processes (Carey, 2009). Spelke's own view has evolved still further: "I once proposed, wrongly, that objects are not grasped by a perceptual system but by the only alternative of which I could conceive: a system of central cognition, like our systems of explicit reasoning about objects and their mechanical interactions" (Spelke, 2024). On this refinement, "core knowledge systems [...] combine some, but not all, of the properties of mature perceptual systems and belief systems", but "differ from all perceptual systems in crucial ways" (Spelke, 2022; for a

[1] To characterize their views on core knowledge, scholars have used terms such as "conceptual", "cognitive", or "reasoning". While these terms do not refer to precisely the same mental processes, they all typically refer to processes that are *not perceptual*, which is the central distinction we are concerned with here.

similar view, see Xu, 2019). In other words, although this view now accepts that core knowledge is not exclusively cognitive/conceptual (which we agree with), it is decidedly *not perceptual* either (which we disagree with, as we clarify below). Other authors have wrestled with similar issues in recent discussions regarding the nature of core knowledge, inviting further theoretical development: "One issue I struggled with in writing *TOOC* [The Origin of Concepts] was differentiating between the senses in which core cognition representations are perceptual and the senses in which they are conceptual. [...] Here is one place where there is much work to be done" (Carey, 2011).

What is crucial about the above discussion is that the core-knowledge view has consistently committed itself to two claims across its various formulations. The first claim is that the infant mind is populated with a rich set of representations about objects, geometry, number, and social agents — a claim that, we agree, is convincingly supported by abundant empirical findings. The second claim concerns the nature of these representations. While this aspect of the view has shifted over the years (and has not settled on a consensus even today), one point is clear: All perspectives on core knowledge reviewed above agree that its representations are *not perceptual*. This is what we challenge here.

3. A new perspective: *Core perception*

3.1 Perception is more sophisticated than commonly assumed

One key reason why core knowledge is commonly considered cognitive (or at least non-perceptual) is a pervasive assumption that perceptual processing delivers only lower-level properties to the perceiver — Spelke's "brightness, color, or distance" (1988). On this view of perception, it is indeed hard to imagine how a purely perceptual process could output representations of objecthood, physics, agency, and other properties and contents represented by the infant's mind. In contrast, however, we argue — in large part on the basis of recent empirical results — that perception is far more sophisticated than traditionally assumed, trafficking in rich representations beyond low-level properties. In other words, we claim not only that perceptual representations serve as *inputs* to core-knowledge processes, but also that many core-knowledge representations are *literally perceptual* themselves. On this view, perceptual representations encompass more than merely "the spatially extended, perceptible surface layout [built] from incoming sensory information" (Spelke, 2022). Instead, perception also incorporates representations that are traditionally considered far beyond the scope of perception (including "what the sensed world consists of: what entities inhabit it, how those entities behave…"; Spelke, 2022).

The idea that a core-knowledge representation could be entirely perceptual may at first blush seem counterintuitive. Consider, for example, a study conducted by Spelke et al. (1992). In this study, infants saw an object being dropped behind a screen where a horizontal surface was present (Figure 3a), and expected the object to end its course atop the surface, as opposed to below it (thus respecting object solidity). Intuitively, it may seem that forming this expectation would require reasoning over conceptual representations, for at least two reasons. First, the dropped object and part of the surface were occluded from view, such that infants could not directly observe whether a physical violation occurred. Second, inferring the final position of the object based on physical laws sounds like a sophisticated inference — one that could be an exercise in a physics class. Accordingly, scholars have considered the fact that a representation can enter into inferences as a key criterion to judge that these representations are conceptual (Carey, 2009), or at least non-perceptual (Xu, 2019). Thus, could perception itself really make use of physical constraints to infer the final position of an object?

As appealing as they may seem, these intuitions are precisely those challenged by both classical and contemporary findings from the science of visual perception. First, and fairly straightforwardly, visual processing is not restricted to the information present in retinal stimulation. We return to this point in Section 6.1, but here are some quick and uncontroversial examples: Although our retina contains an area without photoreceptors, we do not perceive a hole in our visual field, because our visual system automatically fills in this area (i.e., the "blind spot") — and in surprisingly sophisticated ways that incorporate higher-order patterns from the surrounding input (Anstis, 2010). Perception also fills in scotomas and creates continuities over blinks and saccades — phenomena which demonstrate that percepts can be created by the visual system in the absence of corresponding retinal input (Komatsu, 2006; Teichmann et al., 2021). Second, we also know that the visual system implements complex processes that are often considered "inferential" in the sense relevant here. Consider depth perception. We can certainly *reason* that, by virtue of the principles of optics, an object growing bigger in our visual field is likely getting closer to us, or that occluding surfaces must be nearer than what they occlude. These principles might seem to involve sophisticated inferences; yet they describe precisely how the visual system automatically and effortlessly computes depth (Cutting & Vishton, 1995; Goldstein & Brockmole, 2016; Regan & Beverley, 1979): Your visual system *itself* "understands" that looming objects are likely headed your way, and that occluders must be nearer than what they occlude. Our view is that core-knowledge computations function in analogous ways to computations for the blind spot or depth: As sophisticated as these processes may seem, they are embedded in perception itself.

3.2 So what is (and is not) a perceptual process?

How do we know whether a process counts as perception?[2] As explained above, perceptual processes may traffic in information that is not immediately present in sensory stimulation, and they can also carry out surprisingly sophisticated inferences — so these aspects cannot themselves adjudicate whether or not a process is perceptual. One might also be tempted to decide whether a process is perceptual based on the kind of *content* processed: Your reasoning about the global economy would be conceptual because it implicates abstract concepts that cannot be cashed out in terms of sensory inputs, while your seeing colors would be perceptual because they can be cashed out in this way. More generally, it may be tempting to think that abstract representations are necessarily non-perceptual: "They [core-knowledge systems] are not sensory or perceptual systems, because they serve to represent abstract properties and relations" (Spelke, 2022). However, content also is not a foolproof criterion for identifying perceptual processes, for two reasons. First, we do not only see visual contents such as color and shape; we can also think and talk about them, and make judgments on their basis — such as when we consider what color to paint a bedroom, or how to assemble the pieces of a puzzle. Second, and conversely, the fact that a content is abstract (e.g., another agent's goal) is also not foolproof evidence that the process generating that content is conceptual. For instance, there is now considerable evidence that facial attributes (e.g., gender, and emotions such as fear, happiness, or sadness) are processed within perception (Block, 2023). A compelling line of evidence suggests emotional expressions show visual adaptation, such that prolonged fixation on happy faces makes neutral faces look sad (Butler et al., 2008; Webster et al., 2004), just like staring at leftward motion biases subsequent motion perception rightward. Adaptation effects have also been reported for facial cues to gender, and notably, adaptation can transfer between bodies and faces (i.e., adapt to faces and then aftereffect on bodies or vice versa; Ghuman et al., 2010; Palumbo et al., 2015; Weigelt et al., 2010), showing that it occurs at an abstract level. Thus, the visual system can extract and represent not only "red" and "leftward motion", but also abstract properties such as emotional states or gender. In summary, low-level perceptual properties can be represented within systems that are not perceptual, and abstract properties can be represented within the visual system. Therefore, content is a poor litmus test for whether a process or representation is perceptual.

---

[2] We consider it a prerequisite for a process to be perceptual that it takes sensory input as evidence and outputs representations of the distal world. Thus, the present domain of investigation excludes processes such as motor actions, processes that take sensory information as input but do not produce a representation of the distal world (e.g., knee reflex), or processes that output representations of the external world without sensory input (e.g., imagery or hallucinations).

So, how *can* we tell whether a process is perceptual? Several researchers, us among them, have pointed to key behavioral *signatures* that can serve as markers of perception (Hafri & Firestone, 2021; Scholl & Gao, 2013; see also Block, 2023). These signatures are based on empirical regularities observed across a wide array of traditionally accepted perceptual processes: Perceptual processes exhibit most or all of these signatures, whereas conceptual processes exhibit few or none of them. These signatures are: (1) Speed: Seeing properties such as color or orientation proceeds quite rapidly. For example, we can see that an object is big, blue, or bright after very brief exposure (tens of milliseconds) and very little processing time (often 100–200 milliseconds; Mack & Palmeri, 2015; Thorpe et al., 1996). By contrast, cognitive processes often take longer to unfold (e.g., planning a vacation or choosing an appropriate outfit for a social event). (2) Automaticity: As long as one encounters a visual stimulus, relevant perceptual processes will operate. Consider, for example, when an irrelevant bright light or movement in the periphery captures one's attention, having been processed obligatorily (Yantis, 1993). By contrast, many cognitive processes are "optional", such that they may or may not occur in a given situation. Consequently, as we will see in many studies, one way to test automaticity is to show that a given mechanism is activated regardless of whether the experimental task is relevant, whether the relevant input is attended to, and indeed even if processing that stimulus impairs performance on one's primary task. (3) Input sensitivity: Small pixel- or millisecond-scale changes to a visual stimulus may have dramatic effects on the resulting percept. For example, the precise alignment of contours in Kanizsa figures can alter, or even eliminate, their illusory shape (Kellman & Shipley, 1991). Cognitive processes, in contrast, are often indifferent to such subtle changes in visual input. (4) Informational encapsulation: Whereas cognitive processes may be highly sensitive to explicit knowledge and background beliefs, perception stands out for its *insensitivity* to those factors, so much so that we may *see* a stimulus one way even when we *know* it is not so. For example, in the Müller-Lyer illusion (Müller-Lyer, 1889), we perceive two lines as different in length even when we know they are identical (e.g., because we have measured them) or even if we are motivated to see them otherwise (e.g., because we would be rewarded for doing so). (5) Influence on other perceptual processes: Different perceptual inputs may jointly determine and constrain the outputs of perception. In practice, the most compelling evidence for this signature arises when an *unambiguously* perceptual process is affected. For example, manipulating an object's perceived distance can alter its perceived size (e.g., in the Ponzo illusion; Ponzo, 1910), and perceptual properties (e.g., motion onset) may facilitate visual search or target detection (Abrams & Christ, 2003). The influence of perceptual processes may also carry forward in *time,* a paradigmatic example of which is perceptual adaptation. Adaptation occurs when observing a property for a prolonged duration inextricably changes how one sees a subsequent stimulus (as when observing leftward motion biases one to see a subsequent static stimulus as moving rightward). Adaptation for a property, especially when retinotopic, is strong evidence for the existence of dedicated visual mechanisms for that property (Webster, 2015).

To illustrate how our signatures distinguish perceptual processes from conceptual ones, we provide one example of each. First, consider the computation of depth, an unambiguously perceptual process. This process is fast: You can extract stereoscopic depth as quickly as 140 milliseconds after stimulus onset (Mononen et al., 2025). It is triggered automatically: You compute the various depths in a scene as long as you look at it. It is encapsulated: Even in contexts where you know that the perceived depth is illusory, you still cannot will yourself to see otherwise, as shown in classic illusions like the Ames room illusion (in which members of your own family may seem implausibly large or small, or appear at the same distance when you know this property must be changing; Ittelson, 1952). It is sensitive to low-level details of the visual input: The depth cue of binocular disparity itself consists in computing extremely slight differences (often mere minutes of arc) between the images that the two eyes receive (Wheatstone, 1838). Finally, it influences other perceptual processes: Depth can influence the perception of size, as shown in the Ponzo illusion (Ponzo, 1910). Of course, depending on the specific process or representation under consideration, the amount of empirical evidence for each signature may vary; for example, even for phenomena that are clearly visual in nature (such as the effects observed in the Ames room), there still is no evidence (to our knowledge) on how quickly such effects arise. This observation will naturally also be true for many core knowledge representations we review below. But key to our claims (that much of core knowledge is perceptual in nature) is the observation that there is not only an abundance of evidence *for* these signatures, but also a total absence of evidence *against* them.

In contrast to perceptual processes, consider for example the *cognitive* process of deciding which outfit to wear for dinner. This process is slow: You are likely to spend at least minutes (if not much longer) considering different combinations of clothing. It is effortful and deliberative, rather than automatic: You can decide to not carry out such a mental process. It is relatively insensitive to subtleties in low-level visual input: A few arcminutes of difference in the visual angle of your sweater will not drastically influence whether you pick it. It is not encapsulated: The process can be influenced by your background knowledge of tonight's temperature, or whether formal attire is required. And finally, the process of selecting an outfit does not influence other perceptual processes: Your clothes and yourself will look exactly the same regardless of whether you engage in such mental activity. It is worth clarifying that each signature does not *on its own* effectively demarcate the two types of processes. For instance, a fashion designer with considerable expertise might evaluate whether a piece of cloth is wearable the moment they see it, and with little effort. For this reason, our behavioral signatures are intended to be used *as a whole*, instead of as individual criteria. To reiterate, a perceptual process would exhibit most or all the signatures, while a conceptual process would exhibit few or none of them.

In what follows, we build on the above list of criteria and review studies showing a striking overlap between representations in core knowledge and those in visual perception, first applying this approach to core physics and then extending our argument to other core-knowledge domains, namely geometry, number, and social agents. These studies show, on one hand, that many core representations display behavioral signatures of perception in adults. On the other hand, they also demonstrate that core knowledge in infants often exhibits the same capacities and limits as those observed in studies of adult vision. On our view, these two lines of evidence together suggest that visual-perception studies in adults and core-knowledge studies in infants have essentially been studying the same processes: perceptual processes operating from infancy to adulthood. However, to be clear, these findings do not imply that core-knowledge representations play no role in infants' cognitive life. Just like adults can *see* color but also *reason* about what color to paint a bedroom, "conceptual counterparts" may exist in infants' minds, too. Our core-perception view diverges from other core-knowledge views on the role of *perception* in this link — what kind of representations are computed within perception before cognition takes over. More specifically, based on evidence from core-knowledge studies and adult vision studies, we contend that perception outputs representations for core entities such as object interactions, agency, and other sophisticated contents.

4. Core physics

We assemble the case for our perceptual account of core knowledge by first focusing on a prominent domain of study: core physics, i.e., representations of objects and their physical interactions. Our aim is to show how perception itself incorporates surprisingly sophisticated properties and constraints of physics, and moreover that these properties and constraints map with striking consistency to those commonly attributed to infant reasoning — both in terms of what these processes *do* and *do not* represent, as we will see below.

4.1 Spatiotemporal continuity

One property of physical objects is *spatiotemporal continuity:* Objects move along a single continuous path from one moment to the next (regardless of whether they momentarily disappear from view, as in occlusion), rather than, for example, teleporting between locations. Core object representations in infancy adhere to this principle. Infants thus use spatiotemporal continuity to individuate objects, such that a stimulus is represented as one and the same visual object insofar as it traces a spatiotemporally continuous path; otherwise (e.g., if the path is discontinuous), multiple objects are postulated. A classic demonstration of this pattern showed 10-month-olds an object emerging alternatingly from behind two occluders separated by a gap (Figure 1a). If the object could be seen moving across the gap, infants later expected a single object to be revealed, whereas if no object appeared in the gap, then they expected two objects (Xu & Carey, 1996). Subsequent experiments showed that, in contrast to spatiotemporal discontinuities, changes in features are not sufficient for 10-month-olds to postulate the existence of several objects, even when these changes are clearly detectable: After seeing two different items (a ball and a duck) emerge in turn from behind a single large occluder, infants did not form the expectation that two objects were present behind the occluder (Figure 1b).

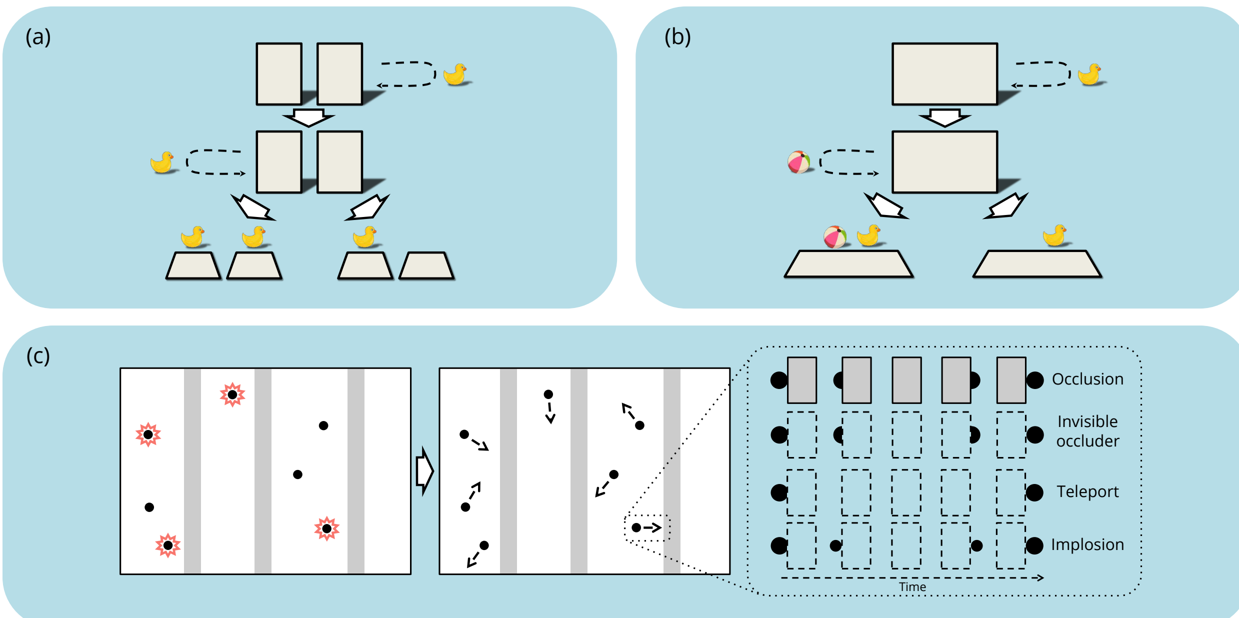

Figure 1 Spatiotemporal continuity. (a) Infants aged 10 months watched an object (e.g., a duck) moving behind and out from two screens separated by a gap. If the duck did not appear in the gap, infants expected two objects to be revealed. If the object appeared in the gap (not shown) as if it were moving in a spatiotemporally continuous path, then infants expected one object to be revealed when the screens came down (Xu & Carey, 1996). (b) Here, two distinct objects (e.g., a ball and a duck) appeared alternatingly from the two sides of a single screen. Infants did not show increased looking time if only one object was revealed (Xu & Carey, 1996). (c) Adult participants were asked to track a subset of moving objects on a screen. Their performance was maintained even when the objects moved behind (visible or invisible) occluders but was diminished if objects teleported across the occluders or imploded and exploded at the occluders' edges (Scholl & Pylyshyn, 1999). The dotted contours represent invisible occluders.

A strikingly similar pattern arises in adults' perception, which also prioritizes spatiotemporal continuity over superficial features. In the classic "tunnel effect" (Burke, 1952; Michotte et al., 1964), an object moves behind an occluder and emerges a moment later, sometimes looking the same and sometimes looking very different (e.g., entering as a kiwi and emerging as a lemon; Flombaum et al., 2004). As long as the emergence of the object from behind the occluder is spatiotemporally consistent with the previous trajectory, adults perceive the two distinct appearances as a single persisting object. This percept is maintained even over dramatic changes to object features, reminiscent of Xu and Carey's finding with 10-month-olds.

Further evidence that adults prioritize spatiotemporal cues when individuating objects is found in multiple object tracking (MOT). As moving objects maintain spatiotemporal continuity, approximately four of them can be simultaneously tracked (Pylyshyn & Storm, 1988). As in the tunnel effect, tracking is preserved even when the objects pass behind visible or even invisible occluders, a type of display compatible with spatiotemporal continuity (Scholl & Pylyshyn, 1999; Figure 1c). If however the object "teleports" across an invisible occluder — a violation of spatiotemporal continuity — tracking is dramatically impaired. Featural information, on the other hand, has much less influence: Tracking performance is not disrupted if the tracked targets undergo changes in shape (e.g., morphing from a triangle into a circle; vanMarle & Scholl, 2003; Figure 2b), or in color or luminance (Zhou et al., 2010). These studies, as well as the tunnel-effect studies reviewed above, demonstrate that adult vision prioritizes spatiotemporal information over featural information to individuate objects — just like infants.

Performance in MOT drops substantially in yet further conditions, where objects "implode" or "explode" (shrinking out of or growing into existence) as they approach occluders (instead of displaying accretion and deletion patterns) — both in adults (Scholl & Pylyshyn, 1999) and infants (Cheries et al., 2005). Object tracking is thus sensitive to subtle changes in parameters: The way the imploding/exploding objects disappear and reappear matches many low-level properties of genuine occlusion, such as rate of change in visible surface area — and yet imploding/exploding is processed very differently from occlusion. Moreover, these findings provide a striking example of informational encapsulation: Even though participants *knew* the objects would "survive" having imploded/exploded, they could not track the objects under such conditions. Thus, tracking objects based on spatiotemporal continuity possesses multiple signatures of perceptual processes, indicating that identifying objects from moment to moment — and relying on spatiotemporal continuity to do so — occurs within the visual system. Convergent with this claim, neuroimaging research finds increased activation in V1, V2, and V3 during occlusion in positions corresponding to the occluded locations, and this increase is disrupted if the objects instantly disappear instead of being gradually occluded (Erlikhman & Caplovitz, 2017).

Another study based on the tunnel effect further supports this picture, providing evidence for *all* of our signatures of perception. In this study, adult participants saw dynamic events in which several objects moved repeatedly behind and out from occluders (Flombaum & Scholl, 2006); the participants were tasked with detecting occasional color changes to the objects by pressing a key as soon as they saw a color change. When an object's movement and position before and after occlusion was spatiotemporally consistent, the object enjoyed processing advantages: Observers were better at detecting color changes for objects moving along a spatiotemporally continuous trajectory as compared to when the object's reappearance included a spatial or temporal gap. These findings demonstrate that object tracking's reliance on spatiotemporal continuity is automatic: Spatiotemporal continuity was irrelevant to the task assigned to participants — which was just to detect color changes on objects — yet it still impacted detection performance. Moreover, the study also shows that tracking objects on the basis of spatiotemporal continuity is fast. As participants were performing the task over several simultaneously moving objects and occluders, it is unlikely that slow and effortful reasoning processes could have been recruited to perform this task. Third, the results reveal a sensitivity to fine-grained differences in the visual input, as subtle spatial and temporal offsets had outsized effects on participants' performance. Fourth, processing the trajectories of objects influenced other perceptual processes, namely, change detection. Fifth, the process is encapsulated from background knowledge about the displays, since the effects did not generalize to displays of the imploding/exploding type (as in Figure 1c) nor to displays with moving screens and stationary objects, despite the explanations provided to participants.

In summary, existing findings establish that infants and adults prioritize the same visual properties when tracking objects through occlusion, and that these properties shape perceptual processing in adults performing challenging visual tasks. Such overlap has been noticed by several researchers (Carey & Xu, 2001; Cheries et al., 2009; Leslie et al., 1998; Scholl & Leslie, 1999), some of whom have argued that the same mechanisms of object-based attention guided by spatiotemporal continuity operate both in infancy and adulthood. In what follows, we demonstrate that such overlap applies to a host of other representations, and to perceptual mechanisms beyond just object-based attention.

4.2 Cohesion and rigidity

Objects are distinguished from substances like liquids or sand. This is true in at least two ways: Objects are cohesive, whereas substances can split or disperse; and objects are rigid, whereas substances can easily change shape. Infants are sensitive to these patterns: They can keep track of two objects lowered behind screens (expecting two objects to be revealed), but they fail when the items are substances such as sand — even when the objects and substances are matched on low-level properties like color and shape (Huntley-Fenner et al., 2002; Figure 2a). Crucially, the adult visual system shows a similar limitation: It treats objects which retain both cohesion and rigidity as appropriate candidates for visual object tracking, but does not do so for substances (which lack those qualities). In vanMarle & Scholl (2003), observers successfully tracked approximately four discrete objects in MOT displays. But strikingly, they failed when the items moved as substances would, by turning into a stream when traveling (as if pouring; Figure 2b). This held even though the objects and substances moved at the same speeds and along the same paths.

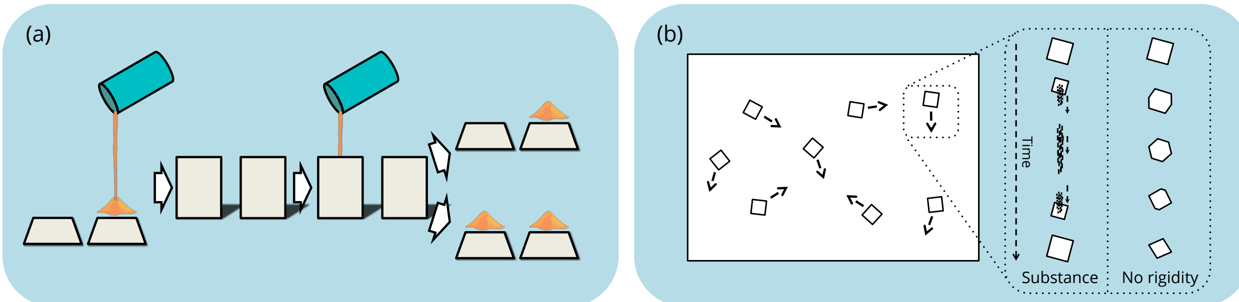

Figure 2 Cohesion and rigidity. (a) Infants were shown sand that was poured in succession behind two occluders (or similar-looking solid objects that were lowered, not shown). They expected two items to be revealed in the solid objects condition, but did not expect two piles in the sand condition (Huntley-Fenner et al., 2002). (b) Adult participants tracked a subset of target objects present on the screen. Their performance was impaired if the objects "poured" from one location to another but not if the items progressively changed shape (vanMarle & Scholl, 2003).

Further experiments have demonstrated that even simple and punctate disruptions of cohesion impair infant object representation and adult visual processing. When infants see one cracker placed in one container and two crackers placed in another container, they generally crawl toward the container with the larger quantity (Feigenson et al., 2002). Cheries et al. (2008) introduced an apparently innocuous manipulation to this paradigm, in which infants observed one small cracker placed in one container and one large cracker being broken in two before the pieces were placed in another container. After this manipulation, infants approached the two containers at similar rates, even though one container still had more crackers — and more cracker material — overall. In contrast, if infants were initially shown the two already-broken pieces without seeing the breaking event itself, they approached the container with two pieces rather than the container with just one. Thus, a simple cohesion violation disrupted infants' representations of objects. Crucially, similar effects arise in adult vision. Mitroff et al. (2004) asked what happens to the identity of a persisting visual object when it splits into two — a nearly identical cohesion violation to that in the infant study above. They did so by exploiting a marker of object maintenance through time: the object-specific preview benefit (OSPB), wherein observers are faster to respond to a previously seen target (e.g., a letter) if it appears on the same visual object as it did moments before, compared to when it appears on a different visual object (Kahneman et al., 1992). In Mitroff et al.'s task, an object visibly splitting in two caused a much reduced OSPB relative to when the object remained a single entity, suggesting that the visual system failed to treat the two items resulting from the split as persisting parts of the original entity. Moreover, the process showed evidence of automaticity, as cohesion impacted performance despite being task-irrelevant. Thus, the same cohesion violations frustrate object representation both in infants and in adults, and in adults these violations have consequences for visual processing (a pattern also noted by Cheries et al., 2009).

Compared to cohesion violations, rigidity violations appear to be less disruptive for object individuation, in both adults and infants. As reviewed above, adults' tracking performance is not typically impaired when items morph from one shape to another (vanMarle & Scholl, 2003; Figure 2b). In line with this pattern, Huntley-Fenner et al. (2002) showed that when two non-rigid objects (e.g., soft toys) were lowered behind two screens, infants expected two objects to be revealed. But if higher cognitive demand was introduced such that the soft toys were lowered behind a *single* screen, infants did not expect two objects to be revealed. In contrast, infants failed with both single and double screens in the case of sand — and succeeded in both screen setups in the case of rigid objects. Thus, unlike cohesion violations, rigidity violations do not *systematically* extinguish infants' object representation, even if such violations do make tracking more difficult.

While infants and adults can track objects through changes in surface features or metric shape, rigidity violations involving changes in *topology* appear much more disruptive in both age groups. Shapes belong to the same topological class if they are invariant under deformations like stretching or bending; by contrast, their topological class is different if such deformations cannot turn one shape into the other. For example, a ball and a cube belong to the same topological class, but a ball and a donut do not. Kibbe and Leslie (2016) found that 6-month-olds fail to remember the appearance (or even existence) of the first of two hidden objects differing in topology, suggesting that the topological contrast increased cognitive load and disrupted object individuation. Relatedly, Brun et al. (2017) found that 12-month-olds fail to infer the presence of two distinct objects when shown two topologically-distinct items emerging one at a time from behind an occluder, even though infants of the same age succeed when shown two differently-shaped objects — but sharing the same topology (e.g., a duck and a ball; Xu & Carey, 1996). Remarkably, topological changes cause similar disruptions in adults' MOT (Zhou et al., 2010): Performance drops significantly when targets undergo topological changes (e.g., when an "S" figure turns into a "9" figure), but not if targets undergo topology-preserving transformations such as color or shape changes. These results represent yet another instance of convergence in the visual properties driving object individuation in both infants and adults.

4.3 Solidity

So far, we have focused on processes involved in tracking individual objects. However, infants' core physics also encompasses principles describing interactions between objects. In particular, the physical objects we encounter follow the constraint of solidity: They cannot pass through one another. Representations of solidity have been extensively studied by developmental psychologists. As previously mentioned, Spelke et al. (1992) showed that after viewing an object dropped behind an occluder, 3.5-month-olds expected the object to reappear atop a horizontal surface instead of below it, where the latter would have required the ball to pass through the surface (Figure 3a). Similarly, 2.5-month-olds detected a violation after viewing a ball rolling horizontally behind an occluder if the ball reappeared beyond a barrier (implying that the ball had passed through it; Spelke et al., 1992), and 3.5-month-olds showed surprise when they saw a screen rotate through a space occupied by a solid object (Baillargeon, 1987). These results demonstrate that preverbal infants expect physical objects to behave according to the solidity constraint.

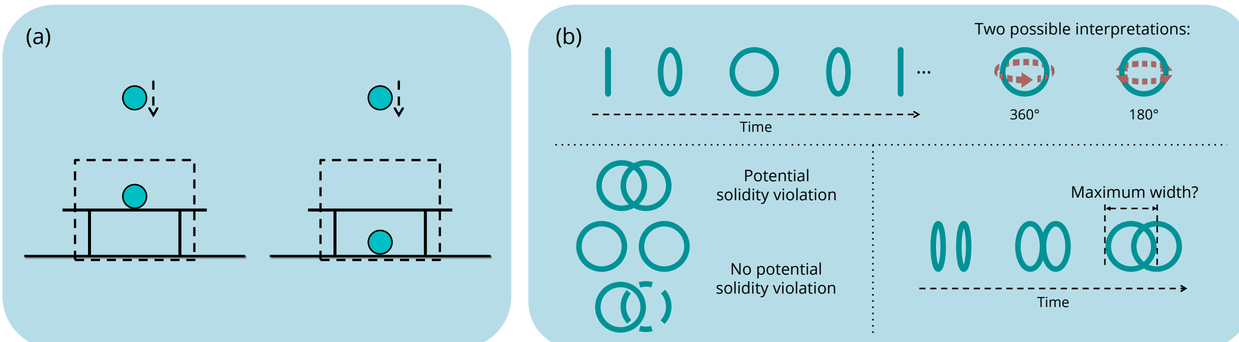

Figure 3 Solidity. (a) After viewing a ball being dropped behind an occluder (dotted lines), infants aged 3.5 months looked longer when the ball was revealed below a solid surface (right), than when it was resting *on* the surface (left; Spelke et al., 1992). (b) The Double Ring Illusion: When viewing a pair of ambiguously rotating rings that can be perceived as moving in 360° or 180° rotations, adult observers predominantly perceived 180° co-rotation if the rings were overlapping, instead of 360° co-rotations which would require the rings to pass through one another. This preference was disrupted if the rings were separated or gapped. A follow-up experiment showed that participants underestimated the maximum width of an overlapping ring more than that of a separated ring, in accordance with the fact that overlapping rings should bounce back before reaching the fronto-parallel plane due to their volume (Bai & Strickland, 2025).

The role of solidity representation in vision has long been a topic of theoretical discussion. Indeed, it has been suggested that solidity may not operate in vision (Leslie, 1988, 1994; Scholl & Leslie, 1999), due to the existence of visual illusions occurring despite obvious solidity violations. For instance, in the Pulfrich double pendulum illusion (Wilson & Robinson, 1986; see also the Pulfrich solidity illusion, Bai & Strickland, 2023), two swinging pendulums appear to pass through each other when viewed with a darkening filter over one eye. The pendulums' illusory motion is caused by the fact that darker images are processed slower in the visual system, and the resultant interocular processing delay displaces stereoscopic depth computations (i.e., the Pulfrich effect; Burge et al., 2019; Lit, 1949; Pulfrich, 1922). However, these illusions do not demonstrate that solidity is not incorporated in the visual system; instead, they only show that the solidity representation can be overridden by stereoscopic depth cues.

Indeed, a recent study demonstrated that representations of solidity actually *do* constrain aspects of visual perception in adults, in the "Double Ring Illusion" (Bai & Strickland, 2025). Adults were shown ambiguously rotating rings that are perceived as multistable — either moving in 360° rotations continuously towards one direction, or 180° rotations flipping back and forth (Figure 3b, top row). Crucially, if another ring was added such that the two rings partially overlapped, participants predominantly reported perceiving them as moving in 180° co-rotations — disfavoring the 360° co-rotation interpretation that would require the rings to pass through each other (thus violating the solidity constraint). Notably, this preference for 180° co-rotations was disrupted when the rings were separated or gapped such that no potential solidity violation could occur, suggesting that solidity is factored into computations of object motion. In a follow-up experiment, participants estimated the maximum width of a rotating ring by adjusting the width of a line (Figure 3b, bottom right). The rationale was that, due to their volume, the overlapping rings should bounce back before reaching the fronto-parallel plane (where they appear widest), leading to a potential underestimation of the maximum (apparent) width of the rings. In contrast, if the rings do not bounce back (i.e., moving in 360° co-rotations), this effect should be reduced or eliminated. Indeed, participants underestimated the maximum width of overlapping rings more than that of separated rings. This finding suggests that the influence of solidity on computations of object motion is automatic, since solidity constraints were irrelevant to the simple task of estimating width. The effect was also sensitive to small parameter changes such as relative distance and the presence of gaps. And lastly, it interacted with another perceptual process, namely size perception. Together, this evidence suggests that solidity constraints are genuinely embedded in perception, displaying key signatures of perceptual processes.

4.4 Causality

Inspired by Michotte (1946/1963), infant cognition researchers have asked whether and to what extent infants appreciate causality in collision-like “launching” events. In such events, an object A moves to contact a stationary object B, after which A stops and B begins moving forward along the same trajectory as A (Figure 4a, left, top row). These events are considered causal, as such motion patterns are spontaneously interpreted as A causing the motion of B. Early work in this area showed that 4.5-month-olds who had habituated to causal launching events dishabituated more to non-causal displays than did infants who had habituated to displays with a spatial or temporal gap at the moment of contact (Leslie, 1982; Figure 4a). Similarly, after habituating to launching events where the collision occurred behind an occluder, infants dishabituated more to displays with a spatial gap than to causal launches (Ball, 1973; Spelke et al., 1995). Subsequent work showed that 6-month-olds who had habituated to launching events in which A caused B’s motion dishabituated when the objects’ roles reversed (B launched A), but not when a temporal gap eliminated the causal interpretation of such events (Leslie & Keeble, 1987; Saxe & Carey, 2006). These findings demonstrate that infants are sensitive to causal events and treat them differently than non-causal events (which are otherwise similar in their motion properties).

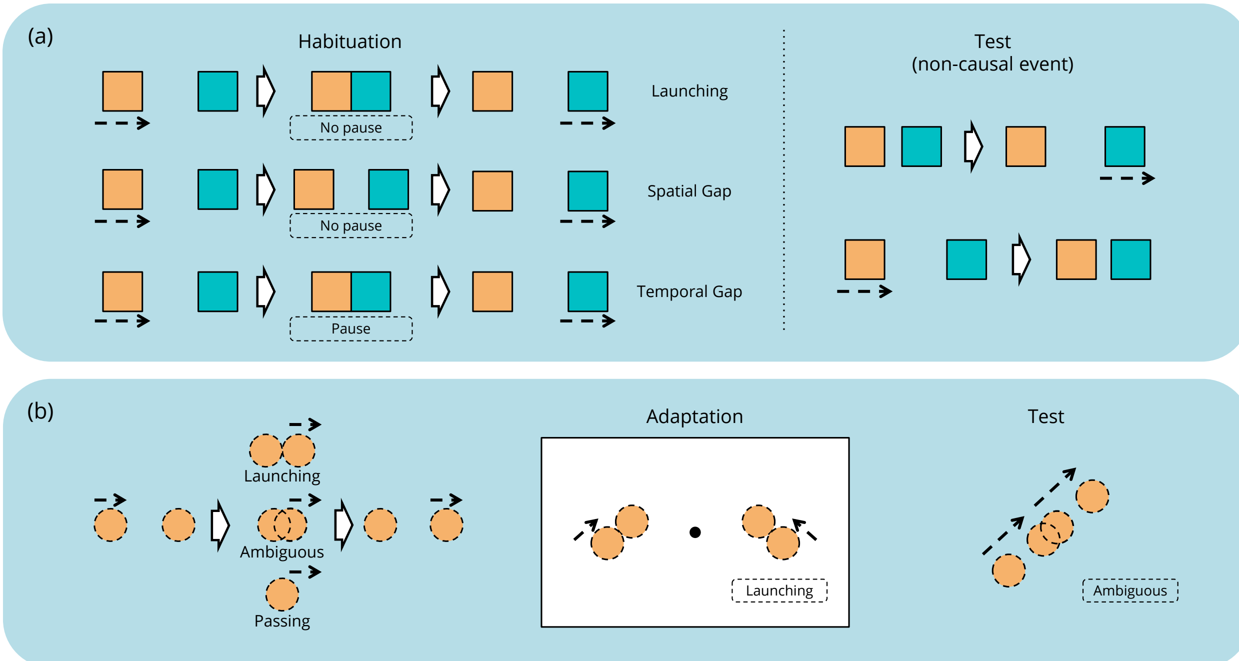

Figure 4 Causality. (a) Each infant is habituated to one of three displays. In "launching", a moving object stops when it touches another object, which immediately moves off. With a spatial gap, the first object stops and the second starts moving without contact. With a temporal gap, the second object starts moving a short while after contact. Only the group habituated with the launching display showed longer looking times when tested with a non-causal event (where one object moves away from another, or one object approaches another; Leslie, 1982). (b) Following adaptation to causal launching displays, adult participants are more likely to report seeing the first disc pass over the second disc when shown an ambiguous display, which can either be interpreted as launching or passing (Rolfs et al., 2013). Dotted lines bordering circles are for illustration purposes only and were not present in the stimulus.

The last 15 years of vision science research have provided converging evidence that causality in events like those described above is extracted in adult visual processing. Moors et al. (2017) used continuous flash suppression to assess the saliency of causal versus closely matched non-causal events like "passing" (e.g., one disc moving behind or in front of another; Figure 4b). Participants received different inputs to their two eyes: a faint target stimulus in one eye and a strong textured mask in the other. Target stimuli were initially suppressed by the mask, but causal events broke into awareness faster than non-causal events, indicating that they are privileged by visual processing. Importantly, participants were not asked about causality — they were simply asked to respond as soon as they perceived one or more discs — and yet causality automatically influenced another visual process, namely interocular suppression. The effect was also sensitive to input settings, as the subtle degree of overlap led to strikingly different responses. Likewise, Kominsky et al. (2017) showed that subtle changes in object speed can have dramatic effects on adults' ability to find "oddballs" among different causal launching events: Adults better detected events in which the speed of the second object was faster than what is physically possible (assuming no internally generated force), compared to a control condition where the two objects' speeds were swapped (i.e., a fast object causing the second object to move slowly, which does not violate Newtonian physics). Notably, this advantage vanished if a temporal or spatial gap was introduced at the moment of contact. Lastly, perhaps the most striking evidence that the visual system extracts causal events is that causal launches display retinotopic adaptation (Rolfs et al., 2013; see also Kominsky & Scholl, 2020). After viewing causal launching events in an adaptation phase, observers had to classify ambiguous stimuli (in which A partially overlaps B before B moves) as causal launches or non-causal passes (Figure 4b). After adaptation, observers classified ambiguous stimuli more often as passes than launches. Remarkably, this only occurred in the retinotopic position where stimuli were presented during adaptation. This finding provides strong evidence that perception computes causal launching events (see discussion in Hafri & Firestone, 2021, and Phillips & Firestone, 2023).

4.5 Containment and occlusion

Young infants sort dynamic information into event categories such as containment, occlusion, covering, tube events, and others (Baillargeon, 2008; Baillargeon & Wang, 2002; Casasola, 2008). These representations then orient attention to certain object features in an event-specific manner. For example, when viewing a containment event — such as an object being lowered into a vertically oriented tube — 4.5-month-olds look longer at physical violations (compared to non-violations) involving object width (e.g., an object somehow fitting inside a container that should be too narrow for it). However, they only show looking-time increases for object *height* violations (e.g., an object fitting inside a container that should be too *short*) at 7.5 months of age (see Figure 5a). By contrast, for visually similar occlusion events (where an object is lowered *behind* a tube), infants reliably encode both the height and width of an object by around 4.5 months of age (Hespos & Baillargeon, 2001; Mou & Luo, 2017; Wang et al., 2004; Wang & Onishi, 2017). Put differently, the infant's mind prioritizes object width over object height for containment events, but for occlusion events these are given equal priority.

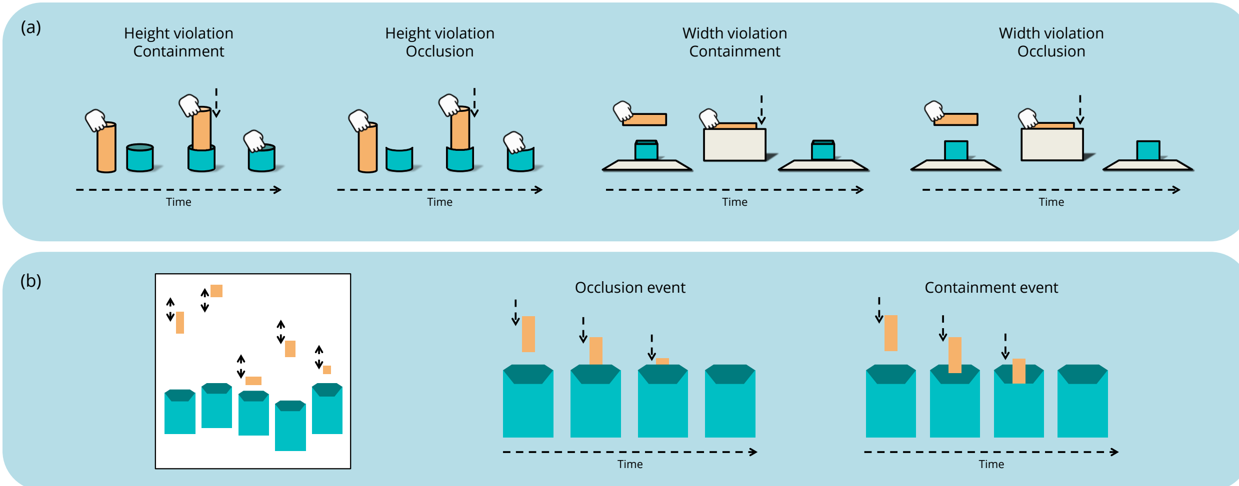

Figure 5 Containment and occlusion. (a) 4.5-month-olds were shown a cylinder lowered into a container or behind an occluder, which were either taller or shorter than the cylinder. Infants looked longer at the short occluder than the tall occluder, but no such difference was found for containers (two left-hand figures; Hespos & Baillargeon, 2001). Experiments involving width were similar and are depicted in the two right-hand figures (note that the gray occluder was necessary here, because otherwise the orange objects would visibly extend outside the blue container/occluder during the event). Here, 4-month-olds detected the violation for both containment and occlusion events (Wang et al., 2004). (b) Adult participants viewed five rectangles (in orange) repeatedly moving up-and-down between the top of the screen and inside or behind the containers (in blue). As a rectangle re-emerged from inside/behind a container, its height or width changed. Participants detected width changes better than height changes for containment events, but this difference was not found for occlusion events (Strickland & Scholl, 2015).

A study with adults found strikingly similar prioritizations of object features in visual processing (Strickland & Scholl, 2015). Participants were shown several rectangles repeatedly moving up and down above vertically oriented containers, which the rectangles either fell *into* or *behind* before re-emerging moments later (Figure 5b). While objects were inside or behind the container, their height or width could change, and participants were tasked with detecting such changes. Participants detected width changes better than height changes for containment events, while no such difference was found for occlusion events — a pattern strikingly similar to that observed in infant studies. Remarkably, this effect arose even for participants who reported not noticing the existence of the two distinct event types in the experiment, suggesting that the underlying mechanism is automatic and encapsulated from beliefs about the displays. Such a process is also sensitive to subtle stimulus parameters, as occlusion and containment events differed from each other only by the edges at which the rectangles disappeared, which were separated by only about 1° of visual angle.

4.6 Summary

In summary, the wealth of evidence demonstrating infants' sensitivity and insensitivity to physical properties of objects and events is mirrored by strikingly similar patterns of sensitivity and insensitivity in adult visual processing of those same properties of objects and events. Infants expect objects to trace a spatiotemporally continuous path (even through occlusion), but are less concerned with objects' featural changes for purposes of tracking and individuation; similarly, adults can visually track objects moving in a spatiotemporally continuous manner through occlusion, even if those objects change features. Infants can represent cohesive objects that are occluded, but struggle with substances; adults can track multiple moving discs, but not pouring substances. Infants' object individuation is sensitive to topological categories; adult visual attention is similarly frustrated by topological changes. Infants expect objects to be solid and not pass through one another; adult perception makes the same assumption when processing object motion. Infants interpret certain events as causal; adult vision shows retinotopic adaptation to the same events. Infants prioritize certain physical dimensions in specific events (e.g., containment); adult vision is better able to detect changes in the same dimensions and event types. And still other patterns (reviewed above) show similar alignment.

These observations thus constitute a powerful "case study" of our central claim: The reason these patterns — many of which are counterintuitive and not easily predictable in advance — arise so consistently in both infant cognition and adult perception is that the *same mechanisms* are operative in both cases. Within the adult visual system are embedded physical principles that create the patterns of behavior reviewed above — and that visual system is, of course, present in the infant's mind as well. Our view is thus that what had initially appeared to be cases of reasoning in infants in fact reflects the operation of their perceptual systems, which turn out to be more sophisticated than our research traditions have assumed. That, in essence, is the core-perception view.

5. Expanding to other core knowledge domains

Beyond objects and their physical interactions, our view is that a much wider range of core representations and processes that have traditionally been considered as purely conceptual should *also* be re-interpreted as reflecting sophisticated perceptual mechanisms. We review a wealth of recent evidence suggesting a perceptual basis for core representations of geometry, number, and even aspects of the social world. Across these cases, the amount of existing evidence may vary due to the recency of this theoretical view. But we argue that, just like representations of core physics, the emerging picture is that such properties are genuinely perceived, rather than only conceived.

5.1 Geometry

Geometry has traditionally been presented as one domain of core knowledge, based on findings demonstrating that toddlers, children, adults, and non-human animals are sensitive to the geometry of their surrounding environment (e.g., Spelke & Kinzler, 2007). Hence, when disoriented, 18-month-olds use the shape of the room to reorient (Hermer & Spelke, 1996), prioritizing geometric properties of the room's layout over non-spatial features such as color (Hermer & Spelke, 1994; Hermer-Vazquez et al., 2001; Lee & Spelke, 2008). Young children's navigation is particularly sensitive to certain boundary cues (Lee, 2017) such as a 2-cm bump from the ground or objects protruding from walls, but not texture-defined boundaries, markings on the floor, or objects detached from walls (Lee & Spelke, 2010, 2011). Similar abilities and limitations have been observed in non-human animals (Cheng, 1986; Cheng & Gallistel, 1984; Lee et al., 2012, 2020), including in animals reared in controlled, circular environments (Brown et al., 2007; Chiandetti & Vallortigara, 2008). These findings have suggested an innate and evolutionarily conserved origin of these geometric abilities — hence their description as a domain of core knowledge.

There is now considerable evidence that a system exhibiting the above characteristics also functions in adults, and that it displays key signatures of perceptual processing. When asked to reorient in a room while verbally shadowing a pre-recorded textual passage (a dual task with verbal interference), adults — similarly to 18-month-olds — use the geometry of the room, but not featural cues such as wall colors (Hermer-Vazquez et al., 1999). Moreover, adults' visual brain areas are sensitive to the same boundary-defining cues that impact performance in children. The scene-selective occipital place area (OPA) plays a causal role in representing distance to visual boundaries (Julian et al., 2016), and does so in ways that are tolerant to textural differences — corresponding closely to children's sensitivity to geometry and not features. Similarly, the parahippocampal place area (PPA) displays a categorical difference between 2D texture-defined boundaries and very slight 3D boundaries (Ferrara & Park, 2016). Importantly, the processes encoding geometry for navigation exhibit several key signatures of perception in adults. First, they display a sharp sensitivity to visual input — as demonstrated by the impact of subtle differences between 2D and 3D boundaries. Second, they operate rapidly: Adults can extract navigationally relevant properties like the spatial openness or closedness of a natural scene within 100 ms, sometimes even before they can detect objects in the scene (Greene & Oliva, 2009). Third, global scene properties such as openness and depth demonstrate adaptation aftereffects (Greene & Oliva, 2010). Lastly, scene representations are activated even when observers are engaged in simple orthogonal visual tasks (e.g., detecting a dot; Bonner & Epstein, 2017), providing evidence of automaticity.

Besides their use of geometry in navigation, infants are also sensitive to the geometrical properties of figures and objects — leading to the proposal that they possess a second system of core geometry, dedicated to small-scale shapes (Dillon et al., 2013; Spelke, 2022; Spelke et al., 2010). Sensitivity to shape features traces back to the very first days of life. For example, few-day-old newborns can discriminate between a cylinder and a triangular prism, and can visually recognize these shapes after holding them (Sann & Streri, 2007). They can also discriminate between objects of different sizes (even when retinal size is controlled for; Slater et al., 1990); between lines presented at different orientations (Slater et al., 1991); and between coarsely contrasted visual shapes (e.g., cross vs. square; Antell & Caron, 1985; Turati et al., 2003). By and large, while the parameters of shape perception continue to be refined throughout childhood and adolescence (Jüttner et al., 2014; Pereira & Smith, 2009; Smith et al., 2014), infants' and children's shape perception still largely overlaps with adults'. For example, 4-month-olds can discriminate vertically symmetrical shapes from other patterns (Fisher et al., 1981), a finding reminiscent of adults' ability to rapidly and effortlessly detect symmetric figures or patterns, especially vertical symmetry (Baylis & Driver, 1994; Palmer & Hemenway, 1978; Wagemans, 1997). Perhaps more impressively, adults and infants alike represent the global shape of objects by extracting their skeletal structure, based on their medial axis (Ayzenberg & Lourenco, 2019, 2022; Firestone & Scholl, 2014; Lowet et al., 2018). Importantly, in adults, sensitivity to the medial axis displays multiple signatures of perceptual processes: It is automatic — skeletal sensitivity arises even while observers display no awareness of it (Firestone & Scholl, 2014) — and it can influence other typically perceptual processes, like orientation detection and visual search (Kovács & Julesz, 1994; Sun & Firestone, 2021).

Together, these studies suggest that infants are exquisitely sensitive to geometrical properties at both small and large scales, and that adults represent these same properties in ways that are genuinely perceptual. It is also worth noting that sensitivity to many geometrical aspects has not been systematically investigated in both infancy and adulthood, leaving many avenues for further research. For example, when they have been disoriented, young children rely on the distances between the walls to reorient, but ignore other geometric cues such as the walls' lengths or relative angles (Lee et al., 2012). Our core-perception view predicts that adult vision should display similar sensitivities and insensitivities. Conversely, studies on adult vision have unearthed a variety of cues collectively determining shape perception (e.g., Morgenstern et al., 2021); we suspect that these same cues may drive infants' perception as well.

5.2 Number

Infants can represent the number of items in a collection in their first year of life (Feigenson et al., 2004; Mou & vanMarle, 2014). At 6 months, they detect number changes in visual arrays (Xu & Spelke, 2000) or auditory streams (Lipton & Spelke, 2003), even when various non-numerical variables are controlled for (e.g., cumulative surface area or spatial density). Infants can also compare numbers across sensory modalities: When presented simultaneously with visual arrays and sound sequences, neonates look longer when the number of objects they see matches the number of sounds they hear (Coubart et al., 2014; Izard et al., 2009; McCrink et al., 2020). Importantly, infants' number representations are approximate, and their ability to discriminate between two quantities is ratio-dependent. Thus, for instance, 6-month-olds can discriminate between 8 and 16 items (a 1:2 ratio), but not between 8 and 12 or between 16 and 24 (a 2:3 ratio; Libertus & Brannon, 2010; Lipton & Spelke, 2003; Xu & Spelke, 2000). Likewise, in children and adults, number discrimination tasks yield psychometric curves governed by ratio (Halberda & Feigenson, 2008; Piazza et al., 2004; Van Oeffelen & Vos, 1982), and number estimations display scalar variability (i.e., the variability of estimates scales with the size of these estimates; Izard & Dehaene, 2008; Sullivan & Barner, 2014). Such ratio-dependence is also present in non-human animals (Nieder, 2019) and has been taken as evidence for the existence of a core system of number, the "Approximate Number System" (ANS).

As revealed over the last 15 years, numerosity processing displays key behavioral signatures of perception. First, adults extract the numerosity of a visual array after very brief exposure — 16 ms (Inglis & Gilmore, 2013) — as brief as even the most basic visual properties. Furthermore, number can be decoded from neural responses in early visual areas 80 ms after stimulus onset (Fornaciai et al., 2017; Park et al., 2016). Second, number exhibits visual adaptation: After staring at an array containing a large number of dots, observers underestimate number in subsequent arrays presented at the same retinotopic location (Burr et al., 2018; but see Yousif et al., 2024 and Burr et al., 2025 for debates). Interestingly, observers seem to experience this effect even when it violates their explicit knowledge, suggesting information encapsulation: In a demo released by Burr and Ross (2008), one cannot help but experience two dot arrays as different in number after adaptation, all while seeing that the two arrays are made of the same dots — and thus must logically be equal in number. Third, numerosity estimation is sensitive to subtle stimulus details. For example, people's estimates are smaller when the dots form regular patterns (Zhao & Yu, 2016), or are grouped in small clusters (Im et al., 2016), when some of the dots are connected by a line (Franconeri et al., 2009; He et al., 2009), or when the elements are Kanizsa-like figures manipulated to face towards each other (Guan et al., 2020). Lastly, numerosity perception is intertwined with perception of other magnitudes such as duration and spatial extent, as well as with spatial positions (de Hevia, 2021). For instance, both adults and infants are faster to detect a target located on the left following a small numerosity prime, and faster to detect a target located on their right following a large numerosity prime (Bulf et al., 2014, 2016). In summary, computations of numerosity thus exhibit several signatures of genuinely perceptual processes, namely speed, informational encapsulation, input sensitivity, and influence on other perceptual processes. Some researchers even consider number as a primary feature processed by the adult perceptual system (Anobile et al., 2016).

Beyond mere quantity estimation or comparison, adults can intuitively perform various tasks *on* numerosities: They can order numerosities, perform addition and subtraction, and compute proportions (e.g., Barth et al., 2006; Pica et al., 2004). Similar abilities have been identified in children (Barth et al., 2005; Szkudlarek & Brannon, 2021), and even in infants (Brannon, 2002; McCrink & Wynn, 2004, 2007). Whether these arithmetic inferences are performed only by higher-level cognitive systems or also by sophisticated perceptual processes is currently an open question.

5.3 Agency and social interaction

A fourth domain of core-knowledge is concerned with social cognition: Preverbal infants possess knowledge about agents' goals and beliefs, and how agents interact with each other and the world. By 6 months, infants assume that social agents (but not inanimate objects) are goal-oriented, such that agents who reach for an object are expected to reach toward that same object again rather than merely the same location (Woodward, 1998, 1999). By one year, infants are likewise able to evaluate the rationality of agents' actions, expecting agents to take efficient and rational paths toward goals (Csibra et al., 2003; Gergely et al., 1995; Liu et al., 2019; Skerry et al., 2013). Furthermore, infants can discriminate dyads based on whether the agents seemingly engage in social interactions (e.g., two face-to-face bodies versus two back-to-back ones; Goupil et al., 2022). Recent work has further shown that infants can differentiate "agents" from "patients" in interacting dyads (Papeo et al., 2024), similarly to non-human primates (Brocard et al., 2024; Wilson et al., 2024). Infants also differentiate roles in other social events such as chasing (Yin & Csibra, 2015) and resource-transfer actions like giving and taking (Tatone et al., 2015, 2021).

Perhaps more so than other core-knowledge domains, agency and social interactions may seem to require deliberate reasoning, as the properties infants are sensitive to (e.g., goals and roles) may seem particularly high-level and abstract. Yet, as we have argued in Section 3.2, representations with abstract content need not be generated by purely conceptual processes. Indeed, recent work in adults has revealed that processing agents and social interactions also demonstrates signatures of perception (Adams et al., 2010; McMahon & Isik, 2023; Papeo, 2020) — part of a broader trend of "vision going social" (Nakayama, 2011).

One line of adult vision research focuses on the perception of social interaction and agency in moving shapes. A canonical example is chasing, in which contingently moving shapes create the vivid impression that they possess mental states and are even pursuing one another (Heider & Simmel, 1944). The extraction of such social information is informationally encapsulated: While the stimuli clearly consist of simple geometric shapes devoid of minds, this knowledge fails to remove the impression of agency. Moreover, perception of such interactions appears to be automatic. When participants performed a "foraging" task where they controlled a shape to collect dots scattered around the scene, they were slower if the shape they controlled was flanked by arrows that appeared to chase it, even though the arrows were irrelevant to the foraging task (van Buren et al., 2016). Chasing percepts are also sensitive to subtle visual parameters, as the above effect is disrupted if the chaser's motion path remains constant but its angle is altered to point away from the "chasee" (Gao et al., 2009, 2010). Similar studies have revealed that adult vision is additionally sensitive to related notions such as goal-directedness (van Buren et al., 2016) and rationality (Gao & Scholl, 2011), mirroring representations of goal-directedness and rationality in infancy.

Another line of vision research has investigated the perception of facing dyads, and found converging evidence consistent with findings in infants mentioned above: People are more sensitive to front-facing dyads — as if interacting — than non-facing ones. Thus, two bodies facing one another are automatically grouped, but not if they are facing away from each other, leading to visual search asymmetries (Papeo et al., 2019). Furthermore, recognition of bodies facing toward but not away from one another is impaired when the bodies are inverted, suggesting configural processing of facing dyads (Papeo et al., 2017). Crucially, these effects occur even when the task is orthogonal to the presence of social interactions, demonstrating the automaticity of such social processing. For example, the grouping effect is found in facing bodies when participants are tasked with finding a target individual (Papeo et al., 2019) or with simply determining whether the display contains human bodies at all (Papeo et al., 2017). Adults can also process social information very rapidly. With a visual exposure as brief as 37 ms, adults can extract social event categories between human figures (e.g., tickling and pushing) and the relational structure of such events (i.e., identifying the agent and the patient) based on postural cues (Dobel et al., 2010; Hafri et al., 2013), even when such information is task-irrelevant (Hafri et al., 2018; Vettori et al., 2025).

Just as looking-time measures demonstrate that preverbal infants differentiate giving from taking actions (Tatone et al., 2015, 2021), an adaptation study demonstrates that visual processing also encodes similar resource-transfer actions (Fedorov et al., 2018). In this study, after being repeatedly shown videos of a giving action, participants perceived an ambiguous action along the giving-throwing continuum as throwing (throwing being the action that lies at the opposite end of that continuum). Strikingly, after participants underwent the same adaptation phase (to actions of giving), subsequent testing on an ambiguous action along a different but related continuum — taking (the opposite of giving) and catching (the opposite of throwing) — caused them to predominantly report catching. In other words, even though participants never directly adapted to the taking-catching continuum, their repeated exposure to the action of giving activated the representation of a contingent action — taking — thus leading to the adaptation aftereffect of catching (i.e., the opposite of taking on that continuum). This suggests that the contingent action pair of giving-taking is embedded in visual processing, and is co-activated during visual adaptation.

Other evidence that social information is encoded in adult vision comes from studies on biological motion. It is well established that observers can extract a remarkably rich set of individual and social attributes (e.g., gender and emotion) from point-light displays, in which only a set of points representing key joints or body parts are visible (Blake & Shiffrar, 2007; Pavlova, 2012). Beyond information about single bodies, the visual system can also extract information about social interactions such as two agents fighting or dancing with each other. In Neri et al. (2006), participants saw two displays masked with noise and selected the display containing two agents instead of one. In the crucial control condition, the two agents were desynchronized; hence, even though both agents were dancing or fighting individually, their interaction was not meaningful. Results showed higher discrimination sensitivity when the two agents were synchronized compared to when they were not. This effect was found even though the temporal offset between the two conditions was subtle, thus reflecting the perceptual signature of input-sensitivity. The process is also automatic, because the task — determining the number of agents — was orthogonal to whether the agents were interacting or not. Other work using similar methods reported converging results: Participants better detected an agent if another agent was performing a communicative gesture towards the former compared to a non-communicative gesture (Manera, Becchio, et al., 2011; Manera, Del Giudice, et al., 2011). Together, these studies demonstrate that the spatiotemporal contingencies of human interaction are integrated into the visual system's processing of biological motion.

Due to the relative recency of psychophysical studies of social perception, many open questions remain. For instance, work with infants suggests that they are sensitive to social relationships (Thomas, 2024) like dominance (Mascaro & Csibra, 2012; Mascaro et al., 2023; Thomsen et al., 2011) and intimacy (Thomas et al., 2022). Future work could explore whether the adult visual system is sensitive to these same representations.

6. Skeptical positions and responses

The foregoing discussion demonstrates the striking overlap between representations and processes in infant core knowledge and adult vision, across multiple domains. Our interpretation of this overlap is that the same mechanisms are operative in both cases — sophisticated perception in both infants and adults. However, a skeptical reader might worry that, no matter the extent of the empirical evidence for this overlap, there are independent reasons to believe that core knowledge *cannot* be computed in the course of visual processing. In this section, we address these worries.

6.1 Vision cannot operate over the unseen

One worry arises from the fact that numerous core-knowledge tasks involve temporarily occluded objects (e.g., in the tunnel effect or in containment events). Could visual processes really represent stimuli that cast no light onto our eyes? Indeed, they can, and routinely do. Consider the familiar case of partial occlusion: When a surface masks part of an object, visual processes fill in the object's missing portion to represent an amodally completed whole. This effect is revealed not only by perceptual phenomenology, but also by visual search tasks using masked displays to pinpoint the precise timecourse of the completion process (Rauschenberger & Yantis, 2001). Furthermore, such perceptual completion leads to enhanced neural activity in regions of the early visual cortex corresponding to the filled-in area (Meng et al., 2005). Another example is apparent motion: When two nearby stimuli appear sequentially, visual processing generates a percept of a single stimulus traversing a continuous path between the two locations, accompanied by V1 activation at the unstimulated regions along the inferred path (Muckli et al., 2005). Other examples of perceptual filling-in include completion at the blind spot, at a scotoma, and during blinking (discussed in Section 3.1), neon color spreading, illusory contours, and so on. Such filling-in can operate over a range of visual properties such as color, edge, surface, object, and more (Komatsu, 2006; Teichmann et al., 2021). These phenomena effectively demonstrate that vivid percepts can be created by the visual system in the absence of corresponding retinal input.

6.2 Vision can only operate over a (very) short time scale

While some core-knowledge theorists now accept that "perceptual systems [...] include processes for extrapolating visible surfaces behind occluders", they may still contend that there remains a disanalogy: The long "time scales over which object representations are maintained, processed, and integrated" (Spelke, 2022) would support a "non-perceptual" account of core knowledge (for a similar assumption, see Mandler, 1988). Indeed, infant studies typically involve occluding objects for many seconds — which may seem like a lifetime for perceptual processes. However, perceptual representations can also be maintained for many seconds. For instance, it is known that what we presently perceive is systematically biased by our visual experience in the past several seconds (as in the phenomenon of serial dependence). Studies have demonstrated serial dependence for orientation (Fischer & Whitney, 2014), color (Barbosa & Compte, 2020), shape (Collins, 2022), and more (Manassi et al., 2023). Some researchers estimate that serial dependence effects are robust up to 10 seconds after stimulus presentation (Manassi et al., 2023). In other words, perceptual processes are *fast*, but they can also be far-*reaching* and create representations that are *temporally extended*. Thus, the fact that object representations are maintained for several seconds when the object is not visible (as when they are occluded) does not rule out such representations being perceptual in nature.

6.3 Perceptual processes cannot be cross-modal

Another reason to insist that core-knowledge representations must be cognitive is that some early-emerging representations appear to transfer across different perceptual modalities (an argument advanced by Spelke, 2022, and supported by Xu, 2019). For example, newborn infants associate numerosities presented in auditory sequences to those presented visually (Izard et al., 2009). Is this not evidence that infants' numerosity representations must be computed by a (central) cognitive system? We think not: Cross-modal transfer of this sort is actually a routine occurrence in perceptual systems, without any need to implicate cognitive processes. For example, both experimental and modeling work suggests that visual, haptic, and auditory information about objects' size, shape, and position are integrated in automatic and input-driven ways: They conform to the relative precision of each sensory modality (Ernst & Banks, 2002), and they are contingent on the precise degree of disparity between multisensory cues (Körding et al., 2007). Collaboration across senses can even give rise to illusory percepts, as in the McGurk effect, where we receive conflicting visual and auditory information: Hearing "ba" while observing a mouth producing "ga" results in a percept of "da" (McGurk & MacDonald, 1976). This effect is impervious to higher-level beliefs or knowledge (e.g., knowledge that only "ba" is playing), fast (within 100 ms of signal onset; van Wassenhove et al., 2005), and sensitive to fine-grained details of the sensory input (as the percept depends on the precise temporal characteristics of the stimuli, shape of the speaker's mouth, etc.; Munhall et al., 1996; van Wassenhove et al., 2007). The existence of such cross-modal effects within perception is not surprising, as we must often process objects' properties or speech in an "online" manner and in noisy environments — and fast, automatic perceptual mechanisms that operate autonomously and integrate information from different senses are well-suited to such situations. Integration mechanisms of a similar sort are likely in play when newborns match numerosities across vision and audition. Thus, demonstrations of cross-modal effects for core representations do not dispel the possibility that such representations are perceptual in nature.

6.4 Core knowledge must be conceptual because it supports conceptual processes

Another worry appears at the interface between core-knowledge representations and higher-level cognition: How could core knowledge be perceptual when such representations have direct effects on unambiguously conceptual processes, such as hypothesis-testing behaviors and word learning? Indeed, this argument is given explicitly by core-knowledge theorists: "A final property of the core object system distinguishes it from perceptual systems: It supports infants' efforts to explain the events that they perceive." (Spelke, 2022). For example, infants will diligently explore an object after it seemingly violates the solidity principle (Stahl & Feigenson, 2015). Similarly, children preferentially associate novel words to objects sharing the same shape rather than the same texture (the "shape bias"; Landau et al., 1988). For representations to serve exploration or language acquisition, this objection might go, mustn't they themselves be conceptual (or at least non-perceptual) in nature?

In our view, the crucial point is the following: You can think, talk about, and explore all sorts of entities whose representations originate as outputs of perceptual processing. This is true both in straightforward or everyday scenarios (e.g., one can think about the colors one sees when mulling over the best color for a wall; see Section 3.2), and even in situations that might be characterized as hypothesis-testing and exploration. For example, under the yellowish moonlight of a dark night, a glass of water and a glass of apple juice might look similar; visual perception outputs an ambiguous color percept consistent with either. This ambiguity can naturally lead to hypothesis-testing behavior such as sniffing the beverage to determine which one it is. The decision to engage in such exploratory behavior is surely cognitive, but it is nevertheless clearly informed by perception. Similarly, when adults are shown physical violations of the sort used in infant studies, they may choose to further interrogate the violations detected by their visual system and draw hypotheses to explain how the events were generated (e.g., surmising that the researchers used video editing software). Likewise, the fact that infants display exploratory behavior following, for instance, viewing solidity violations need not mean that solidity constraints are not initially computed within perception. Instead, in both infants and adults, inconsistencies (e.g., physical violations) can be detected and flagged at the perceptual stage, and then fed into post-perceptual reasoning processes.

While core representations may eventually be used by higher cognitive processes for purposes such as exploration or reasoning, the empirical evidence we have marshaled in this article nevertheless supports the hypothesis that many core representations are output by genuine perceptual processes.

6.5 An unfalsifiable view?

It is important here to be clear about the scope of the view we are presenting. Our central claim is that many core abilities assumed to reside in cognitive processes instead arise from sophisticated perceptual processing. However, this does not entail that *all* such processes and representations are perceptual in nature. For example, 4-month-olds expect adults to comfort a crying baby (Jin et al., 2018), but such expectations could very well reside wholly in higher cognition, since there is currently no evidence that this representation — or more generally morality itself — is computed by perception (for another example, see

Firestone & Scholl, 2015; though see also evidence that rapidly extracted visual representations of causality and harm may automatically *inform* moral judgments; De Freitas & Alvarez, 2018; De Freitas & Hafri, 2024). Ultimately, for any given representation or process, it is an empirical question whether perception can process it or not.

In light of this, one might raise the concern that our view is unfalsifiable, because it is open to the possibilities that (1) some core representations are exclusively cognitive, and (2) some core representations operate both in cognition and perception (as discussed in Section 6.4). This worry is misplaced. Our discussion has showcased a precise set of core-knowledge representations (in Sections 4 and 5) that we argue are part of perception. If future evidence reveals that some of these very representations do *not* operate in infants' perceptual systems, then our view is falsified in those cases. Specifically, if studies show that infants' representation of spatiotemporal continuity, for instance, does *not* display any key signature of perception (so that the processes involved are always slow, optional, and so on), then that strongly undermines our view. This consideration also points to the importance of investigating key signatures of perception in infants' core-knowledge processes, a point which we discuss in Section 7.1.

6.6 Top-down effects of cognition on perception?

A further objection might also serve as a simple way to explain the overlap between representations in adult perception and infant cognition: positing that cognition exerts top-down effects on perception. On this view, core representations are thoroughly cognitive, but they then influence how objects, physical interactions, geometry, numerical properties, and agents are *perceived*. However, there are independent reasons to be skeptical of such effects. Firestone and Scholl (2016) systematically evaluated the evidence for top-down effects reported in hundreds of studies and concluded that all these attempts have suffered from a set of pitfalls that undermine such claims.

For instance, one classic study of this kind reported that wearing a heavy backpack makes hills look steeper (Bhalla & Proffitt, 1999), such that spatial perception reflects the organism's abilities and intentions. However, instead of truly *perceiving* hills as steeper, participants might have *guessed* the purpose of the experiment and responded cooperatively. Indeed, later studies found that if participants were given a cover story that prevented them from guessing the purpose of the study, the effect disappeared (Durgin et al., 2009). Another issue with these studies concerns the measures of perception. For example, it has been claimed that *knowing* the typical colors of objects (e.g., that bananas are yellow) biases *perception* of those objects towards their known colors: When participants adjust an image of a banana to appear gray, they actually make it slightly *blue* (yellow's opponent color), as if countering some extra yellow added by their minds (Hansen et al., 2006; also Witzel, 2016), whereas this does not occur for a disk (which would have no canonical color). However, Valenti and Firestone (2019) showed that when participants select the "odd one out" from a bluish banana, a bluish disk, and a gray disk, they choose the gray disk, suggesting they did in fact perceive the bluish banana as blue (like the bluish disk), instead of gray. Thus, what appeared to be a long-term effect of cognition on perception actually reflected certain response biases (e.g., intentionally adjusting away from yellow as a strategy). Additional issues include experimenter biases, other confounding factors, and explanations arising from other cognitive systems (Firestone & Scholl, 2015; see also Amir & Firestone, 2025). Collectively, these pitfalls account for studies allegedly demonstrating top-down effects from cognition to perception, such that arguments relying on them may be tenuous.

7. Looking ahead

We have reviewed an extensive body of evidence that perception traffics in core-knowledge representations which are commonly attributed to non-perceptual cognitive processes. Our view has two primary upshots: (1) That many core-knowledge representations are actually perceptual in nature, and (2) that perception is more sophisticated than typically assumed, generating surprisingly rich representations across diverse domains including objects, physics, geometry, number, and social agents. Such processes display key signatures associated with perception per se, including speed, automaticity, sensitivity to subtle aspects of sensory input, informational encapsulation, and interaction with other unambiguously perceptual processes. In this final section, we draw out further theoretical and practical implications from our view, and point to new research directions beyond those outlined earlier.

7.1 Fruitful dialogues between research fields

Our framework holds that developmental and perceptual psychologists are often studying the same underlying processes (i.e., sophisticated perceptual processes). Thus, an implication of our view is that one should be able to generate predictions about adult perceptual performance from infant data, and vice versa.

Much of the work reviewed above illustrates that such predictions can be successful. For example, Xu and Carey's (1996) classic finding that infants prioritize spatiotemporal continuity over featural continuity directly motivated predictions that change detection and MOT should be sensitive to the same principles — predictions that turned out to be correct (Flombaum & Scholl, 2006; Scholl & Pylyshyn, 1999). Similarly, based on findings that infants' sensitivity to object dimensions varies by event types (containment vs. occlusion; Hespos & Baillargeon, 2001; Wang et al., 2004), Strickland and Scholl (2015) correctly hypothesized that adults would show similar patterns in change detection in analogous displays.

This exchange has also proceeded in the other direction. For example, Strickland and Scholl (2015) found that when containers were turned on their sides, the prioritized dimension also rotated by 90° (i.e., height was better detected than width in containment events). Based on this finding, Goldman and Wang (2019) discovered that the same pattern holds in infants who see horizontal containment events. Along similar lines, work on causality (Kominsky et al., 2017; Leslie, 1982; Michotte, 1946/1963) and dyadic interactions (Goupil et al., 2022; Papeo et al., 2017, 2019) also benefited from such considerations. The productive back-and-forth between vision and infancy research has deepened our understanding of the underlying mechanisms — and continues to do so.

Along a similar vein, findings in adult vision have benefited research investigating core knowledge in non-human animals. For example, studies showing that adult vision groups facing dyads together (Papeo et al., 2017, 2019) directly inspired work finding that macaques prefer to look at displays of facing dyads (of conspecifics) over non-facing ones (Goupil et al., 2024). Perhaps most strikingly, newly-hatched female chicks also exhibit a visual preference for point-light displays of facing dyads (of hens) over non-facing dyads (Zanon et al., 2024). Comparative studies allow researchers to address the (often overlooked) nativist claims of the core-knowledge theory, which have been primarily justified by data from young infants, leading some developmental researchers to doubt whether core representations are truly innate (Blumberg & Adolph, 2023); work in non-human animals helps to rebut this criticism. Dialogues between these scientific fields can thus provide important insights into the origins of core representations.

Our framework additionally generates hypotheses about the neural bases of core knowledge. Under a view that considers core knowledge solely a part of higher cognition, one might expect core representations to recruit only regions associated with higher-level reasoning (e.g., prefrontal cortex; Wertheim & Ragni, 2020). By contrast, our view predicts that much of core knowledge may have a basis in functionally specific perceptual areas. Indeed, a growing literature has begun to provide just such evidence. For example, intuitive physics tasks engage a functionally specific network that includes parietal areas in the dorsal visual pathway (Fischer et al., 2016; Pramod et al., 2022). Likewise, specialized areas in high-level visual regions have been implicated in other core-knowledge domains, including number (Nieder, 2019), environmental geometry (Bonner & Epstein, 2017; Julian et al., 2016), and social interaction (Abassi & Papeo, 2020; Gandolfo et al., 2024; Isik et al., 2017). Future work could investigate the developmental trajectory of such specialized perceptual areas, as has been done for high-level visual categories such as faces, scenes, and number (Buiatti et al., 2019; Deen et al., 2017; Hyde, 2023; Kamps et al., 2020; Kosakowski et al., 2022, 2024; Nakai et al., 2023; Spriet et al., 2022).

Furthermore, our framework suggests an intriguing and — to our knowledge — heretofore unexplored line of research: developmental studies explicitly aimed at testing the signatures of perception on core knowledge. Though infant research poses substantial challenges for investigating these signatures, relevant studies have successfully demonstrated influence on other perceptual processes, input sensitivity, and more. For example, creative behavioral studies have revealed patterns of discrimination performance that are consistent with categorical perception of color (Franklin, 2013; Franklin et al. 2008), and neuroscientific and behavioral methods have been crafted to measure the speed of infant face perception (Gelskov & Kouider, 2010; Halit et al., 2004; Hochmann & Kouider, 2022; Kouider et al., 2013). Developmental research on core knowledge, however, has seen substantially less work investigating these signatures of perception (with perhaps the exception of input sensitivity, which has been explored by, for instance, studies on causality and numerosity). Our framework invites researchers to identify the other behavioral signatures in processes giving rise to core-knowledge representations. We believe this general approach can be a productive next step for uncovering the intricate interplay between core knowledge and perceptual mechanisms.

7.2 The format of core-knowledge representations in perception and cognition

The existence of sophisticated core-knowledge representations in perception has implications for the format of perceptual representations and for how perceptual and cognitive systems interact. A common assumption about perception is that it is inherently iconic or "image-like", such that parts of a representation correspond to parts of the represented scene or state of affairs, much like a photograph (Carey, 2009; Dretske, 1981; Kosslyn et al., 2006). Recent work broadens this notion, framing iconicity in terms of an analog format — where representational elements mirror graded differences in features (e.g., orientation or brightness) without requiring holistic, picture-like mappings (Beck, 2019; Burge, 2022; Block, 2023) — and Clarke (2022), for example, even suggests that certain core-knowledge representations employ this analog format.

However, although such formats are well-tailored to encode continuous, low-level visual properties like orientation or brightness, they may be ill-suited to capturing the categorical properties central to core-knowledge systems, such as causality, containment, and the like. By contrast, symbolic formats are better suited for this purpose (for a detailed treatment of the iconic-symbolic spectrum, see Greenberg 2023). Thus, our core-perception view may implicate a greater role for such symbolic formats in perception than has typically been considered. This perspective dovetails with a resurgence of interest in language-of-thought-like representational formats (Dehaene et al., 2025; Fodor, 1975; Goodman et al., 2014; Sablé-Meyer et al., 2022) — including in visual perception (Cavanagh, 2021; Green & Quilty-Dunn, 2021; Hafri et al., 2023; Quilty-Dunn et al., 2023). Such shared formats would facilitate (and even make less mysterious) the interface between perceptual and conceptual systems, with perception outputting high-level representations in a format "immediately consumable by cognition" (Quilty-Dunn, 2020).

7.3 Perceptual roots of conceptual knowledge?

In holding that much of core knowledge is perceptual in nature, our framework raises questions and hypotheses about the development of *conceptual* representations. Children and adults have the capacity to deliberate and reason about many of the same properties that form core domains; what are the origins of that capacity? Although related questions also arise on the traditional view of core knowledge (see, e.g., Carey, 2009), different answers to them may be more or less attractive under the framework of core perception.

If much of our knowledge is initially perceptual in nature, then one possibility is that these initial perceptual representations *serve as the basis* for later conceptual representations. In other words, if representations of, say, solidity or goal-directedness are initially embedded in infants' perceptual systems but are not yet available for more domain-general reasoning processes, these initial perceptual representations may help guide infants to derive corresponding conceptual ones. This hypothesis may at first glance seem reminiscent of empiricism (Locke, 1690/1959; Hume, 1740/2000), a prominent view of conceptual development postulating that our conceptual repertoire is wholly derived from perceptual experience. However, a crucial point of divergence here is that, according to empiricism, the to-be-learned conceptual representations are initially absent from the mind in any form (including perceptual), such that learning or acquiring such concepts must involve some kind of abstraction process (from low-level properties to high-level ones). In contrast, the core perception framework invites us to consider a process of conceptual development that is typically absent from empiricist views of the mind — one in which high-level representations that are already present in perception are *transferred* into conceptual systems without the need for constructing or abstracting entirely new representational structures.[3]

What mechanism might achieve such perception-to-conception transfer? One potential candidate could be attention-triggered exploration and learning (in line with what earlier theorists have posited, e.g., Mandler, 2000). This process could operate in two (non-mutually exclusive) ways. One way is for attention to select "natural units" in the environment such as physical objects and agents (Scholl, 2001), with these object-level units — rather than lower-level properties like brightness, orientations, or edges — handed off as the

[3] An alternative possibility is that perceptual and conceptual representations both exist from infancy, but that the mapping between perception and conception — i.e., semantic access to the relevant concepts via perception (Fodor, 1998) — develops over time. On either account, the perceptual representations effectively precede access to their conceptual counterparts.

primary currency of more general reasoning processes. Thus, for instance, learners may predominantly reason about objects that bounce off of each other upon collision, rather than about changes in brightness caused by a passing shadow; over time learners may derive a principle pertaining to object solidity and causality, rather than a principle related to light intensity. A second way is for attention to flag violations to the learner, inviting the learner to gather more information (see discussion in Section 6.4). Supporting such a mechanism, it has been shown that speed patterns that violate Newtonian mechanics are found faster amongst distractors than speed patterns respecting this constraint (Kominsky et al., 2017), suggesting that physical violations do spontaneously capture attention. In infants, it is well established that viewing events that display violations leads to prolonged looking time (as shown by studies using the VOE paradigm), as well as exploration (Perez & Feigenson, 2022; Stahl & Feigenson, 2015, 2019). Thus, attentional processes may be guided by underlying perceptual representations and accordingly direct learners to relevant properties in the environment. In doing so, the attention system gives learners greater opportunity to derive explicit knowledge about those properties and to develop hypotheses about which perceptual features are relevant for gaining semantic access to specific concepts.

7.4 Conclusion

Since its emergence four decades ago, the core-knowledge research program has been one of the most ambitious endeavors in cognitive science. At its heart, it addresses some of the deepest and most enduring questions in the study of the mind: What is the origin of human knowledge? How do our earliest representations come to be? How much of what we know is innate, and how much is acquired from experience?

Unlike many ambitious projects, core-knowledge research is a *success* story. It has yielded mountains of converging data, and has served as a wellspring for groundbreaking discoveries in fields as far-ranging as psychology, anthropology, linguistics, vision science, computer science, and more. This work has delivered compelling answers to many of its motivating questions, and will no doubt continue to do so.

However, bound up with the questions that animate core-knowledge research is another equally important question: What is the *nature* of this early knowledge, and where is its place in the architecture of the mind? We believe there is now substantial evidence for a heterodox answer. Having assembled and examined findings on both infants' early representations and adults' perception, the consistent alignment of these two sources of evidence — across diverse domains of core knowledge — is, in our view, too consistent and too striking to be coincidental. The most compelling and parsimonious explanation is that much of core knowledge is a part of perception itself.

Acknowledgements

We would like to thank Roberto Casati, Jérôme Dokic, Zsuzsa Kaldy, Michael Murez, François Recanati, and Brian Scholl for their valuable feedback and conversations over the years around the topics covered in this paper.

Funding statement

This work received support from:

- Université Paris Sciences & Lettres (IPFBW 2016-151).
- Agence Nationale de la Recherche (grants ANR-10-IDEX-0001-02 PSL and ANR-17-EURE-0017 FrontCog).
- The European Union (FP/2007-2013; 324115, FRONTSEM; 778077, Orisem).
- Centre National de la Recherche Scientifique (CNRS).
- National Science Foundation #2021053 awarded to C.F.

Conflicts of interest

The authors declare no conflicts of interest.

References

Abassi, E., & Papeo, L. (2020). The representation of two-body shapes in the human visual cortex. *The Journal of Neuroscience*, *40*(4), 852–863. https://doi.org/10.1523/JNEUROSCI.1378-19.2019

Abrams, R. A., & Christ, S. E. (2003). Motion onset captures attention. *Psychological Science*, *14*(5), 427–432. https://doi.org/10.1111/1467-9280.01458

Adams, R. , Ambady, N., Nakayama, K., & Shimojo, S. (Eds.). (2010). *The science of social vision*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195333176.001.0001

Amir, D., & Firestone, C. (2025). Is visual perception WEIRD? The Müller-Lyer illusion and the Cultural Byproduct Hypothesis. *Psychological Review*. Advance online publication. https://doi.org/10.1037/rev0000549

Anobile, G., Cicchini, G. M., & Burr, D. C. (2016). Number as a primary perceptual attribute: A review. *Perception*, *45*(1–2), 5–31. https://doi.org/10.1177/0301006615602599

Anstis, S. (2010). Visual filling-in. *Current Biology*, *20*(16), R664–R666. https://doi.org/10.1016/j.cub.2010.06.029

Antell, S. E., & Caron, A. J. (1985). Neonatal perception of spatial relationships. *Infant Behavior and Development*, *8*(1), 15–23. https://doi.org/10.1016/S0163-6383(85)80013-8

Ayzenberg, V., & Lourenco, S. F. (2019). Skeletal descriptions of shape provide unique perceptual information for object recognition. *Scientific Reports*, *9*(1), 9359. https://doi.org/10.1038/s41598-019-45268-y

Ayzenberg, V., & Lourenco, S. (2022). Perception of an object's global shape is best described by a model of skeletal structure in human infants. *eLife*, *11*, e74943. https://doi.org/10.7554/eLife.74943

Bai, D., & Strickland, B. (2023). The Pulfrich Solidity Illusion: A surprising demonstration of the visual system's tolerance of solidity violations. *Psychonomic Bulletin & Review*, *30*(5), 1782–1787. https://doi.org/10.3758/s13423-023-02271-9

Bai, D., & Strickland, B. (2025). The "Double Ring Illusion": The physical constraint of solidity shapes visual processing. *Journal of Experimental Psychology: General*. Advance online publication. https://doi.org/10.1037/xge0001839

Baillargeon, R. (1986). Representing the existence and the location of hidden objects: Object permanence in 6-and 8-month-old infants. *Cognition*, *23*(1), 21–41. https://doi.org/10.1016/0010-0277(86)90052-1

Baillargeon, R. (1987). Object permanence in 3½-and 4½-month-old infants. *Developmental Psychology*, *23*(5), 655–664. https://doi.org/10.1037/0012-1649.23.5.655

Baillargeon, R. (2008). Innate ideas revisited: For a principle of persistence in infants' physical reasoning. *Perspectives on Psychological Science*, *3*(1), 2–13. https://doi.org/10.1111/j.1745-6916.2008.00056.x

Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, *20*(3), 191–208. https://doi.org/10.1016/0010-0277(85)90008-3

Baillargeon, R., & Wang, S. H. (2002). Event categorization in infancy. *Trends in Cognitive Sciences*, *6(2)*, 85–93. https://doi.org/10.1016/S1364-6613(00)01836-2

Ball, W. A. (1973). The perception of causality in the infant. Paper presented at the biannual meeting of the *Society for Research in Child Development*, Philadelphia, PA.

Barbosa, J., & Compte, A. (2020). Build-up of serial dependence in color working memory. *Scientific Reports*, *10*(1), 10959. https://doi.org/10.1038/s41598-020-67861-2

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Non-symbolic arithmetic in adults and young children. *Cognition*, *98*(3), 199–222. https://doi.org/10.1016/j.cognition.2004.09.011

Barth, H., La Mont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Sciences*, *102*(39), 14116–14121. https://doi.org/10.1073/pnas.0505512102

Baylis, G. C., & Driver, J. (1994). Parallel computation of symmetry but not repetition within single visual shapes. *Visual Cognition*, *1*(4), 377–400. https://doi.org/10.1080/13506289408401715

Beck, J. (2019). Perception is Analog: The Argument from Weber's Law. *The Journal of Philosophy*, *116*(6), 319–349. https://doi.org/10.5840/jphil2019116621

Bhalla, M., & Proffitt, D. R. (1999). Visual–motor recalibration in geographical slant perception. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(4), 1076–1096. https://doi.org/10.1037/0096-1523.25.4.1076

Blake, R., & Shiffrar, M. (2007). Perception of human motion. *Annual Review of Psychology*, *58(1)*, 47–73. https://doi.org/10.1146/annurev.psych.57.102904.190152

Block, N. (2023). *The border between seeing and thinking*. Oxford University Press. https://doi.org/10.1093/oso/9780197622223.001.0001

Blumberg, M. S., & Adolph, K. E. (2023). Protracted development of motor cortex constrains rich interpretations of infant cognition. *Trends in Cognitive Sciences*, *27*(3), 233–245. https://doi.org/10.1016/j.tics.2022.12.014

Bonner, M. F., & Epstein, R. A. (2017). Coding of navigational affordances in the human visual system. *Proceedings of the National Academy of Sciences*, *114*(18), 4793–4798. https://doi.org/10.1073/pnas.1618228114

Brannon, E. M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition*, *83*(3), 223–240. https://doi.org/10.1016/S0010-0277(02)00005-7

Brocard, S., Wilson, V. A., Berton, C., Zuberbühler, K., & Bickel, B. (2024). A universal preference for animate agents in hominids. *iScience*, *27*(6). https://doi.org/10.1016/j.isci.2024.109996

Brown, A. A., Spetch, M. L., & Hurd, P. L. (2007). Growing in circles: Rearing environment alters spatial navigation in fish. *Psychological Science*, *18*(7), 569–573. https://doi.org/10.1111/j.1467-9280.2007.01941.x

Brun, M., Martin L., Strickland B., Marie J., Le Saint K., Izard V. (2017). Does topology help object individuation?, *Budapest CEU Conference for Cognitive Development*, Hungary.

Buiatti, M., Di Giorgio, E., Piazza, M., Polloni, C., Menna, G., Taddei, F., Baldo, E., & Vallortigara, G. (2019). Cortical route for facelike pattern processing in human newborns. *Proceedings of the National Academy of Sciences*, *116*(10), 4625–4630. https://doi.org/doi:10.1073/pnas.1812419116

Bulf, H., de Hevia, M. D., & Macchi Cassia, V. (2016). Small on the left, large on the right: Numbers orient visual attention onto space in preverbal infants. *Developmental Science*, *19*(3), 394–401. https://doi.org/10.1111/desc.12315

Bulf, H., Macchi Cassia, V., & de Hevia, M. D. (2014). Are numbers, size and brightness equally efficient in orienting visual attention? Evidence from an eye-tracking study. *PLOS One*, *9*(6), e99499. https://doi.org/10.1371/journal.pone.0099499

Burge, J., Rodriguez-Lopez, V., & Dorronsoro, C. (2019). Monovision and the Misperception of Motion. *Current Biology*, *29*(15), 2586–2592. https://doi.org/10.1016/j.cub.2019.06.070

Burge, T. (2022). *Perception: First form of mind*. Oxford University Press. https://doi.org/10.1093/oso/9780198871002.001.0001

Burke, L. (1952). On the tunnel effect. *Quarterly Journal of Experimental Psychology*, *4*(3), 121–138. https://doi.org/10.1080/17470215208416611

Burr, D. C., Anobile, G., & Arrighi, R. (2018). Psychophysical evidence for the number sense. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1740), 20170045. https://doi.org/10.1098/rstb.2017.0045

Burr, D., Anobile, G., & Arrighi, R. (2025). Number adaptation: reply. *Cognition*, *254*, 105870. https://doi.org/10.1016/j.cognition.2024.105870

Burr, D., & Ross, J. (2008). A visual sense of number. *Current biology*, *18*(6), 425–428. https://doi.org/10.1016/j.cub.2008.02.052

Butler, A., Oruc, I., Fox, C. J., & Barton, J. J. (2008). Factors contributing to the adaptation aftereffects of facial expression. *Brain research*, *1191*, 116–126. https://doi.org/10.1016/j.brainres.2007.10.101

Carey, S. (2009). *The Origin of Concepts*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195367638.001.0001

Carey, S. (2011). Précis of the origin of concepts. *Behavioral and Brain Sciences*, *34*(3), 113–124. https://doi.org/10.1017/S0140525X10000919

Carey, S., & Spelke, E. (1994). Domain-Specific knowledge and conceptual change. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 169–200). Cambridge University Press. https://doi.org/10.1017/CBO9780511752902.008

Carey, S., & Spelke, E. (1996). Science and core knowledge. *Philosophy of Science*, *63*(4), 515–533. https://doi.org/10.1086/289971

Carey, S., & Xu, F. (2001). Infants' knowledge of objects: Beyond object files and object tracking. *Cognition*, *80*(1–2), 179–213. https://doi.org/10.1016/S0010-0277(00)00154-2

Casasola, M. (2008). The development of infants' spatial categories. *Current Directions in Psychological Science*, *17*(1), 21–25. https://doi.org/10.1111/j.1467-8721.2008.00541.x

Cavanagh, P. (2021). The language of vision. *Perception*, *50*(3), 195–215. https://doi.org/10.1177/0301006621991491

Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition*, *23*(2), 149–178. https://doi.org/10.1016/0010-0277(86)90041-7

Cheng, K., & Gallistel, C. R. (1984). Testing the geometric power of an animal's spatial representation. In H. L. Roitblat, T. C. Bever, & H. S. Terrace (Eds.), *Animal Cognition* (pp. 409–423). Erlbaum.

Cheries, E. W., Feigenson, L., Scholl, B. J., & Carey, S. (2005). Cues to object persistence in infancy: Tracking objects through occlusion vs. implosion. *Journal of Vision*, *5*(8), 352–352. https://doi.org/10.1167/5.8.352

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. https://doi.org/10.1111/j.1467-7687.2008.00687.x

Cheries, E. W., Mitroff, S. R., Wynn, K., & Scholl, B. J. (2009). Do the same principles constrain persisting object representation in infant cognition and adult perception? The cases of continuity and cohesion. In B. Hood & L. Santos (Eds.), *The origins of object knowledge* (pp.107–134). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199216895.003.0005

Chiandetti, C., & Vallortigara, G. (2008). Is there an innate geometric module? Effects of experience with angular geometric cues on spatial re-orientation based on the shape of the environment. *Animal Cognition*, *11*(1), 139–146. https://doi.org/10.1007/s10071-007-0099-y

Clarke, S. (2022). Beyond the icon: Core cognition and the bounds of perception. *Mind & Language*, *37*(1), 94–113. https://doi.org/10.1111/mila.12315

Clarke, S., & Beck, J. (2023). Border Disputes: Recent Debates along the Perception–Cognition Border. *Philosophy Compass*, *18*(8), e12936. https://doi.org/10.1111/phc3.12936

Collins, T. (2022). Serial dependence occurs at the level of both features and integrated object representations. *Journal of Experimental Psychology. General*, *151*(8), 1821–1832. https://doi.org/10.1037/xge0001159

Cooper, L. A., & Shepard, R. N. (1973). Chronometric studies of the rotation of mental images. In W. G. Chase (Eds.), *Visual information processing* (pp. 75– 176). Academic Press. https://doi.org/10.1016/B978-0-12-170150-5.50009-3

Coubart, A., Izard, V., Spelke, E. S., Marie, J., & Streri, A. (2014). Dissociation between small and large numerosities in newborn infants. *Developmental Science*, *17*(1), 11–22. https://doi.org/10.1111/desc.12108

Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, *27*(1), 111–133. https://doi.org/10.1207/s15516709cog2701_4

Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In *Perception of space and motion* (pp. 69–117). Academic Press. https://doi.org/10.1016/B978-012240530-3/50005-5

De Freitas, J., & Alvarez, G. A. (2018). Your visual system provides all the information you need to make moral judgments about generic visual events. *Cognition*, *178*, 133–146. https://doi.org/10.1016/j.cognition.2018.05.017

De Freitas, J., & Hafri, A. (2024). Moral thin-slicing: Forming moral impressions from a brief glance. *Journal of Experimental Social Psychology, 112*, 104588. https://doi.org/10.1016/j.jesp.2023.104588

de Hevia, M. D. (2021). How the human mind grounds numerical quantities on space. *Child Development Perspectives*, *15*(1), 44–50. https://doi.org/10.1111/cdep.12398

Deen, B., Richardson, H., Dilks, D. D., Takahashi, A., Keil, B., Wald, L. L., Kanwisher, N., & Saxe, R. (2017). Organization of high-level visual cortex in human infants. *Nature Communications*, *8*(1), 13995. https://doi.org/10.1038/ncomms13995

Dehaene, S., Sablé-Meyer, M., & Ciccione, L. (2025). Origins of numbers: a shared language-of-thought for arithmetic and geometry?. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2025.03.001

Dillon, M. R., Huang, Y., & Spelke, E. S. (2013). Core foundations of abstract geometry. *Proceedings of the National Academy of Sciences*, *110*(35), 14191–14195. https://doi.org/10.1073/pnas.1312640110

Dobel, C., Glanemann, R., Kreysa, H., Zwitserlood, P., & Eisenbeiss, S. (2010). Visual encoding of coherent and non-coherent scenes. In J. Bohnemeyer & E. Pederson (Eds.), *Event representation in language: Encoding events at the language cognition interface*, 189–215. Cambridge University Press. https://doi.org/10.1017/CBO9780511782039.009

Dretske, F. I. (1981). *Knowledge and the flow of information*. MIT Press.

Durgin, F. H., Baird, J. A., Greenburg, M., Russell, R., Shaughnessy, K., & Waymouth, S. (2009). Who is being deceived? The experimental demands of wearing a backpack. *Psychonomic Bulletin & Review*, *16*(5), 964–969. https://doi.org/10.3758/PBR.16.5.964

Erlikhman, G., & Caplovitz, G. P. (2017). Decoding information about dynamically occluded objects in visual cortex. *NeuroImage*, *146*, 778–788. https://doi.org/10.1016/j.neuroimage.2016.09.024

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433. https://doi.org/10.1038/415429a

Fedorov, L. A., Chang, D. S., Giese, M. A., Bülthoff, H. H., & De la Rosa, S. (2018). Adaptation aftereffects reveal representations for encoding of contingent social actions. *Proceedings of the National Academy of Sciences*, *115*(29), 7515–7520. https://doi.org/10.1073/pnas.1801364115

Feigenson, L., Carey, S., & Hauser, M. (2002). The representations underlying infants' choice of more: Object files versus analog magnitudes. *Psychological Science*, *13*(2), 150–156. https://doi.org/10.1111/1467-9280.00427

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002

Ferrara, K., & Park, S. (2016). Neural representation of scene boundaries. *Neuropsychologia*, *89*, 180–190. https://doi.org/10.1016/j.neuropsychologia.2016.05.012

Firestone, C., & Scholl, B. J. (2014). "Please tap the shape, anywhere you like": Shape skeletons in human vision revealed by an exceedingly simple measure. *Psychological Science*, *25*(2), 377–386. https://doi.org/10.1177/0956797613507584

Firestone, C., & Scholl, B. J. (2015). Enhanced visual awareness for morality and pajamas? Perception vs. memory in 'top-down' effects. *Cognition*, *136*, 409–416. https://doi.org/10.1016/j.cognition.2014.10.014

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences*, *39*, e229. https://doi.org/10.1017/S0140525X15000965

Fischer, J., Mikhael, J. G., Tenenbaum, J. B., & Kanwisher, N. (2016). Functional neuroanatomy of intuitive physical inference. *Proceedings of the National Academy of Sciences*, *113*(34), E5072–E5081. https://doi.org/10.1073/pnas.1610344113

Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience*, *17*(5), 738–743. https://doi.org/10.1038/nn.3689

Fisher, C. B., Ferdinandsen, K., & Bornstein, M. H. (1981). The role of symmetry in infant form discrimination. *Child Development*, *52*(2), 457–462. https://doi.org/10.2307/1129162

Flombaum, J. I., Kundey, S. M., Santos, L. R., & Scholl, B. J. (2004). Dynamic object individuation in rhesus macaques: A study of the tunnel effect. *Psychological Science*, *15*(12), 795–800. https://doi.org/10.1111/j.0956-7976.2004.00758.x

Flombaum, J. I., & Scholl, B. J. (2006). A temporal same-object advantage in the tunnel effect: facilitated change detection for persisting objects. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(4), 840. https://doi.org/10.1037/0096-1523.32.4.840

Fodor, J. A. (1975). The language of thought. Harvard university press.

Fodor, J. A. (1983). *The modularity of mind*. MIT press.

Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong*. Oxford University Press. https://doi.org/10.1093/0198236360.001.0001

Fornaciai, M., Brannon, E. M., Woldorff, M. G., & Park, J. (2017). Numerosity processing in early visual cortex. *NeuroImage*, *157*, 429–438. https://doi.org/10.1016/j.neuroimage.2017.05.069

Franconeri, S. L., Bemis, D. K., & Alvarez, G. A. (2009). Number estimation relies on a set of segmented objects. *Cognition*, *113*(1), 1–13. https://doi.org/10.1016/j.cognition.2009.07.002

Franklin, A. (2013). Infant color categories. In L. Ronier (Ed.), *Encyclopedia of color science and technology*. Springer-Verlag (Online). https://doi.org/10.1007/978-3-642-27851-8

Franklin, A., Drivonikou, A., Bevis, L., Davies, I., Kay, P., & Regier, T. (2008). Categorical perception of color is lateralized to the right hemisphere in infants, but to the left hemisphere in adults. *Proceedings of the National Academy of Sciences*, *105*(9), 3221–3225. https://doi.org/10.1073/pnas.0712286105

Gandolfo, M., Abassi, E., Balgova, E., Downing, P. E., Papeo, L., & Koldewyn, K. (2024). Converging evidence that left extrastriate body area supports visual sensitivity to social interactions. *Current Biology*, *34*(2), 343–351. https://doi.org/10.1016/j.cub.2023.12.009

Gao, T., McCarthy, G., & Scholl, B. J. (2010). The wolfpack effect: Perception of animacy irresistibly influences interactive behavior. *Psychological Science*, *21*(12), 1845–1853. https://doi.org/10.1177/0956797610388814

Gao, T., Newman, G. E., & Scholl, B. J. (2009). The psychophysics of chasing: A case study in the perception of animacy. *Cognitive Psychology*, *59*(2), 154–179. https://doi.org/10.1016/j.cogpsych.2009.03.001

Gao, T., & Scholl, B. J. (2011). Chasing vs. stalking: interrupting the perception of animacy. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(3), 669. https://doi.org/10.1037/a0020735

Gelskov, S. V., & Kouider, S. (2010). Psychophysical thresholds of face visibility during infancy. *Cognition*, *114*(2), 285–292. https://doi.org/10.1016/j.cognition.2009.09.012

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*(2), 165–193. https://doi.org/10.1016/0010-0277(95)00661-H

Ghuman, A. S., McDaniel, J. R., & Martin, A. (2010). Face Adaptation without a Face. *Current Biology*, *20*(1), 32–36. https://doi.org/10.1016/j.cub.2009.10.077

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), 2475–2482. https://doi.org/10.1037/dev0000823

Goldstein, E. B., & Brockmole, J. (2016). *Sensation and Perception* (10th ed.). Cengage Learning.

Goodman, N. D., Tenenbaum, J. B., Gerstenberg, T. (2014). Concepts in a Probabilistic Language of Thought. In E. Margolis & S. Laurence (Eds.), *The conceptual mind: New directions in the study of concepts* (pp. 623–654). MIT Press.

Goupil, N., Papeo, L., & Hochmann, J. R. (2022). Visual perception grounding of social cognition in preverbal infants. *Infancy*, *27*(2), 210–231. https://doi.org/10.1111/infa.12453

Goupil, N., Rayson, H., Serraille, E., Massera, A., Ferrari, P. F., Hochmann, J. R., & Papeo, L. (2024). Visual preference for socially relevant spatial relations in humans and monkeys. *Psychological Science, 35*(6), 681–693. https://doi.org/10.1177/09567976241242995

Green, E. J., & Quilty-Dunn, J. (2021). What is an Object File? *British Journal for the Philosophy of Science*, *72*(3), 665–699. https://doi.org/10.1093/bjps/axx055

Greenberg, G. (2023). The iconic-symbolic spectrum. *Philosophical Review*, *132*(4), 579–627. https://doi.org/10.1215/00318108-10697558

Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding. *Psychological Science*, *20*(4), 464–472. https://doi.org/10.1111/j.1467-9280.2009.02316.x

Greene, M. R., & Oliva, A. (2010). High-level aftereffects to global scene properties. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1430–1442. https://doi.org/10.1037/a0019058

Guan, C., Schwitzgebel, D., Hafri, A., & Firestone, C. (2020). Possible objects count: perceived numerosity is altered by representations of possibility. *Journal of Vision*, *20*(11), 847–847. https://doi.org/10.1167/jov.20.11.847

Hafri, A., & Firestone, C. (2021). The perception of relations. *Trends in Cognitive Sciences*, *25*(6), 475–492. https://doi.org/10.1016/j.tics.2021.01.006

Hafri, A., Green, E. J., & Firestone, C. (2023). Compositionality in visual perception. *Behavioral and Brain Sciences*, *46*, e277. https://doi.org/10.1017/S0140525X23001838

Hafri, A., Papafragou, A., & Trueswell, J. C. (2013). Getting the gist of events: recognition of two-participant actions from brief displays. *Journal of Experimental Psychology: General*, *142*(3), 880. https://doi.org/10.1037/a0030045

Hafri, A., Trueswell, J. C., & Strickland, B. (2018). Encoding of event roles from visual scenes is rapid, spontaneous, and interacts with higher-level visual processing. *Cognition*, *175*, 36–52. https://doi.org/10.1016/j.cognition.2018.02.011

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "Number Sense": The Approximate Number System in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, *44*(5), 1457. https://doi.org/10.1037/a0012682

Halit, H., Csibra, G., Volein, A., & Johnson, M. H. (2004). Face-sensitive cortical processing in early infancy. *Journal of Child Psychology and Psychiatry*, *45*(7), 1228–1234. https://doi.org/10.1111/j.1469-7610.2004.00321.x

Hansen, T., Olkkonen, M., Walter, S., & Gegenfurtner, K. R. (2006). Memory modulates color appearance. *Nature Neuroscience*, *9*(11), 1367–1368. https://doi.org/10.1038/nn1794

He, L., Zhang, J., Zhou, T., & Chen, L. (2009). Connectedness affects dot numerosity judgment: Implications for configural processing. *Psychonomic Bulletin & Review*, *16*(3), 509–517. https://doi.org/10.3758/PBR.16.3.509

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, *57*(2), 243–259. https://doi.org/10.2307/1416950

Hermer, L., & Spelke, E. (1996). Modularity and development: The case of spatial reorientation. *Cognition*, *61*(3), 195–232. https://doi.org/10.1016/S0010-0277(96)00714-7

Hermer, L., & Spelke, E. S. (1994). A geometric process for spatial reorientation in young children. *Nature*, *370*(6484), 57–59. https://doi.org/10.1038/370057a0

Hermer-Vazquez, L., Moffet, A., & Munkholm, P. (2001). Language, space, and the development of cognitive flexibility in humans: The case of two spatial memory tasks. *Cognition*, *79*(3), 263–299. https://doi.org/10.1016/S0010-0277(00)00120-7

Hermer-Vazquez, L., Spelke, E. S., & Katsnelson, A. S. (1999). Sources of flexibility in human cognition: dual-task studies of space and language. *Cognitive Psychology*, *39*(1), 3–36. https://doi.org/10.1006/cogp.1998.0713

Hespos, S. J., & Baillargeon, R. (2001). Infants' knowledge about occlusion and containment events: A surprising discrepancy. *Psychological Science*, *12*(2), 140–147. https://doi.org/10.1111/1467-9280.00324

Hochmann, J. R., & Kouider, S. (2022). Acceleration of information processing en route to perceptual awareness in infancy. *Current Biology*, *32*(5), 1206–1210. https://doi.org/10.1016/j.cub.2022.01.029

Hood, B., & Willatts, P. (1986). Reaching in the dark to an object's remembered position: Evidence for object permanence in 5-month-old infants. *British Journal of Developmental Psychology*, *4*(1), 57–65. https://doi.org/10.1111/j.2044-835X.1986.tb00998.x

Hume, D. (1740/2000). *A treatise of human nature*. Oxford University Press. (Original work published 1740)

Huntley-Fenner, G., Carey, S., & Solimando, A. (2002). Objects are individuals but stuff doesn't count: Perceived rigidity and cohesiveness influence infants' representations of small groups of discrete entities. *Cognition*, *85*(3), 203–221. https://doi.org/10.1016/S0010-0277(02)00088-4

Hyde, D. C. (2023). Cognitive neuroscience: An abstract sense of number in the infant brain. *Current Biology*, *33*(10), R400–R402. https://doi.org/10.1016/j.cub.2023.03.079

Im, H. Y., Zhong, S. H., & Halberda, J. (2016). Grouping by proximity and the visual impression of approximate number in random dot arrays. *Vision Research*, *126*, 291–307. https://doi.org/10.1016/j.visres.2015.08.013

Inglis, M., & Gilmore, C. (2013). Sampling from the mental number line: How are approximate number system representations formed? *Cognition*, *129*(1), 63–69. https://doi.org/10.1016/j.cognition.2013.06.003

Isik, L., Koldewyn, K., Beeler, D., & Kanwisher, N. (2017). Perceiving social interactions in the posterior superior temporal sulcus. *Proceedings of the National Academy of Sciences*, *114*(43), E9145–E9152. https://doi.org/10.1073/pnas.1714471114

Ittelson, W. H. (1952). *The Ames demonstrations in perception; a guide to their construction and use*. Princeton University Press.

Izard, V., & Dehaene, S. (2008). Calibrating the mental number line. *Cognition*, *106*(3), 1221–1247. https://doi.org/10.1016/j.cognition.2007.06.004

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences*, *106*(25), 10382–10385. https://doi.org/10.1073/pnas.0812142106

James, W. (1890/1981). *The principles of psychology.* Harvard University Press. (Original work published 1890)

Jin, K.-s., Houston, J. L., Baillargeon, R., Groh, A. M., & Roisman, G. I. (2018). Young infants expect an unfamiliar adult to comfort a crying baby: Evidence from a standard violation-of-expectation task and a novel infant-triggered-video task. *Cognitive Psychology*, *102*, 1–20. https://doi.org/10.1016/j.cogpsych.2017.12.004

Julian, J. B., Ryan, J., Hamilton, R. H., & Epstein, R. A. (2016). The occipital place area is causally involved in representing environmental boundaries during navigation. *Current Biology*, *26*(8), 1104–1109. http://dx.doi.org/10.1016/j.cub.2016.02.066

Jüttner, M., Petters, D., Wakui, E., & Davidoff, J. (2014). Late development of metric part-relational processing in object recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1718–1734. https://doi.org/10.1037/a0037288

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*(2), 175–219. https://doi.org/10.1016/0010-0285(92)90007-O

Kamps, F. S., Hendrix, C. L., Brennan, P. A., & Dilks, D. D. (2020). Connectivity at the origins of domain specificity in the cortical face and place networks. *Proceedings of the National Academy of Sciences*, *117*(11), 6163–6169. https://doi.org/10.1073/pnas.1911359117

Karmiloff-Smith, A. (2009). Nativism versus neuroconstructivism: Rethinking the study of developmental disorders. *Developmental Psychology*, *45*(1), 56–63. https://doi.org/10.1037/a0014506

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*(2), 141–221. https://doi.org/10.1016/0010-0285(91)90009-D

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. https://doi.org/10.1016/j.cogdev.2015.12.001

Komatsu, H. (2006). The neural mechanisms of perceptual filling-in. *Nature Reviews Neuroscience*, *7*(3), 220–231. https://doi.org/10.1038/nrn1869

Kominsky, J. F., & Scholl, B. J. (2020). Retinotopic adaptation reveals distinct categories of causal perception. *Cognition*, *203*, 104339. https://doi.org/10.1016/j.cognition.2020.104339

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. https://doi.org/10.1177/0956797617719930

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. PLOS *One*, *2*(9), e943. https://doi.org/10.1371/journal.pone.0000943

Kosakowski, H. L., Cohen, M. A., Herrera, L., Nichoson, I., Kanwisher, N., & Saxe, R. (2024). Cortical face-selective responses emerge early in human infancy. *eNeuro, 11*(7). https://doi.org/10.1523/ENEURO.0117-24.2024

Kosakowski, H. L., Cohen, M. A., Takahashi, A., Keil, B., Kanwisher, N., & Saxe, R. (2022). Selective responses to faces, scenes, and bodies in the ventral visual pathway of infants. *Current Biology*, *32*(2), 265–274. https://doi.org/10.1016/j.cub.2021.10.064

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195179088.001.0001

Kouider, S., Stahlhut, C., Gelskov, S. V., Barbosa, L. S., Dutat, M., de Gardelle, V., Christophe, A., Dehaene, S., & Dehaene-Lambertz, G. (2013). A neural marker of perceptual consciousness in infants. *Science*, *340*(6130), 376–380. https://doi.org/10.1126/science.1232509

Kovács, I., & Julesz, B. (1994). Perceptual sensitivity maps within globally defined visual shapes. *Nature*, *370*(6491), 644–646. https://doi.org/10.1038/370644a0

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*(3), 299–321. https://doi.org/10.1016/0885-2014(88)90014-7

Lee, S. A. (2017). The boundary-based view of spatial cognition: a synthesis. *Current Opinion in Behavioral Sciences*, *16*, 58–65. https://doi.org/10.1016/j.cobeha.2017.03.006

Lee, S. A., Austen, J. M., Sovrano, V. A., Vallortigara, G., McGregor, A., & Lever, C. (2020). Distinct and combined responses to environmental geometry and features in a working-memory reorientation task in rats and chicks. *Scientific Reports*, *10*(1), 1–9. https://doi.org/10.1038/s41598-020-64366-w

Lee, S. A., Sovrano, V. A., & Spelke, E. S. (2012). Navigation as a source of geometric knowledge: Young children's use of length, angle, distance, and direction in a reorientation task. Cognition, *123*(1), 144–161. https://doi.org/10.1016/j.cognition.2011.12.015

Lee, S. A., & Spelke, E. S. (2008). Children's use of geometry for reorientation. *Developmental Science*, *11*(5), 743–749. https://doi.org/10.1111/j.1467-7687.2008.00724.x

Lee, S. A., & Spelke, E. S. (2010). A modular geometric mechanism for reorientation in children. *Cognitive Psychology*, *61*(2), 152–176. https://doi.org/10.1016/j.cogpsych.2010.04.002

Lee, S. A., & Spelke, E. S. (2011). Young children reorient by computing layout geometry, not by matching images of the environment. *Psychonomic Bulletin & Review*, *18*(1), 192–198. https://doi.org/10.3758/s13423-010-0035-z

Lee, S. A., Spelke, E. S., & Vallortigara, G. (2012). Chicks, like children, spontaneously reorient by three-dimensional environmental geometry, not by image matching. *Biology Letters*, *8*(4), 492–494. https://doi.org/10.1098/rsbl.2012.0067

Leslie, A. M. (1982). The perception of causality in infants. *Perception*, *11*(2), 173–186. https://doi.org/10.1068/p110173

Leslie, A. M. (1988). The necessity of illusion: Perception and thought in infancy. In L. Weiskrantz (Ed.), *Thought without language* (pp. 185–210). Clarendon Press/Oxford University Press.

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. https://doi.org/10.1017/CBO9780511752902.006

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition*, *25*(3), 265–288. https://doi.org/10.1016/S0010-0277(87)80006-9

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*(1), 10–18. https://doi.org/10.1016/S1364-6613(97)01113-3

Libertus, M. E., & Brannon, E. M. (2010). Stable individual differences in number discrimination in infancy. *Developmental Science*, *13*(6), 900–906. https://doi.org/10.1111/j.1467-7687.2009.00948.x

Lin, Y., Stavans, M., & Baillargeon, R. (2022). Infants' physical reasoning and the cognitive architecture that supports It. In O. Houdé & G. Borst (Eds.), *The Cambridge Handbook of Cognitive Development* (pp. 168–194). Cambridge University Press. https://doi.org/10.1017/9781108399838.012

Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense: Large-number discrimination in human infants. *Psychological Science*, *14*(5), 396–401. https://doi.org/10.1111/1467-9280.01453

Lit, A. (1949). The magnitude of the Pulfrich stereophenomenon as a function of binocular differences of intensity at various levels of illumination. *The American Journal of Psychology*, *62*(2), 159–181. https://doi.org/10.2307/1418457

Liu, S., Brooks, N. B., & Spelke, E. S. (2019). Origins of the concepts cause, cost, and goal in prereaching infants. *Proceedings of the National Academy of Sciences*, *116*(36), 17747–17752. https://doi.org/10.1073/pnas.1904410116

Locke, J. (1690/1959). *An essay concerning human understanding, vols. I and II*. Dover. (Original work published 1690).

Lowet, A. S., Firestone, C., & Scholl, B. J. (2018). Seeing structure: Shape skeletons modulate perceived similarity. *Attention, Perception, & Psychophysics*, *80*(5), 1278–1289. https://doi.org/10.3758/s13414-017-1457-8

Mack, M. L., & Palmeri, T. J. (2015). The dynamics of categorization: Unraveling rapid categorization. *Journal of Experimental Psychology: General*, *144*(3), 551–569. https://doi.org/10.1037/a0039184

Manassi, M., Murai, Y., & Whitney, D. (2023). Serial dependence in visual perception: A meta-analysis and review. *Journal of Vision*, *23*(8), 18–18. https://doi.org/10.1167/jov.23.8.18

Mandler, J. M. (1988). How to build a baby: On the development of an accessible representational system. *Cognitive Development*, *3*(2), 113–136. https://doi.org/10.1016/0885-2014(88)90015-9

Mandler, J. M. (2000). Perceptual and conceptual processes in infancy. *Journal of Cognition and Development*, *1*(1), 3–36. https://doi.org/10.1207/S15327647JCD0101N_2

Manera, V., Becchio, C., Schouten, B., Bara, B. G., & Verfaillie, K. (2011). Communicative interactions improve visual detection of biological motion. *PLOS One*, *6*(1), e14594. https://doi.org/10.1371/journal.pone.0014594

Manera, V., Del Giudice, M., Bara, B. G., Verfaillie, K., & Becchio, C. (2011). The second-agent effect: communicative gestures increase the likelihood of perceiving a second agent. *PLOS One*, *6*(7), e22650. https://doi.org/10.1371/journal.pone.0022650

Margoni, F., Surian, L., & Baillargeon, R. (2024). The violation-of-expectation paradigm: A conceptual overview. *Psychological Review, 131*(3), 716–748. https://doi.org/10.1037/rev0000450

Mascaro, O., & Csibra, G. (2012). Representation of stable social dominance relations by human infants. *Proceedings of the National Academy of Sciences*, *109*(18), 6862–6867. https://doi.org/10.1073/pnas.1113194109

Mascaro, O., Goupil, N., Pantecouteau, H., Depierreux, A., Van der Henst, J. B., & Claidière, N. (2023). Human and animal dominance hierarchies show a pyramidal structure guiding adult and infant social inferences. *Nature Human Behaviour*, *7*(8), 1294–1306. https://doi.org/10.1038/s41562-023-01634-5

McCrink, K., Veggiotti, L., & de Hevia, M. D. (2020). A left visual advantage for quantity processing in neonates. *Annals of the New York Academy of Sciences*, *1477*(1), 71–78. https://doi.org/10.1111/nyas.14457

McCrink, K., & Wynn, K. (2004). Large-number addition and subtraction by 9-month-old infants. *Psychological Science*, *15*(11), 776–781. https://doi.org/10.1111/j.0956-7976.2004.00755.x

McCrink, K., & Wynn, K. (2007). Ratio abstraction by 6-month-old infants. *Psychological Science*, *18*(8), 740–745. https://doi.org/10.1111/j.1467-9280.2007.01969.x

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748. https://doi.org/10.1038/264746a0

McMahon, E., & Isik, L. (2023). Seeing social interactions. *Trends in Cognitive Sciences*, *27*(12), 1165–1179. https://doi.org/10.1016/j.tics.2023.09.001

Meng, M., Remus, D. A., & Tong, F. (2005). Filling-in of visual phantoms in the human brain. *Nature Neuroscience*, *8*(9), 1248–1254. https://doi.org/10.1038/nn1518

Michotte, A. (1963). *La perception de la causalité*. Institut Supérieur de Philosophie, 1946. English translation of updated edition by T. Miles & E. Miles, *The perception of causality*, Basic Books, 1963. (Original work published 1946)

Michotte, A., Thinés, G., & Crabbé, G. (1964). *Les compléments amodaux des structures perceptives.* Publications Universitaires de Louvain.

Mitroff, S. R., Scholl, B. J., & Wynn, K. (2004). Divide and conquer: How object files adapt when a persisting object splits into two. *Psychological Science*, *15*(6), 420–425. https://doi.org/10.1111/j.0956-7976.2004.00695.x

Mononen, R., Saarela, T., Vallinoja, J., Olkkonen, M., & Henriksson, L. (2025). Cortical encoding of spatial structure and semantic content in 3D natural scenes. *Journal of Neuroscience*, *45*(9), e2157232024. https://doi.org/10.1523/JNEUROSCI.2157-23.2024

Moors, P., Wagemans, J., & de-Wit, L. (2017). Causal events enter awareness faster than non-causal events. *PeerJ*, *5*, e2932. https://doi.org/10.7717/peerj.2932

Morgenstern, Y., Hartmann, F., Schmidt, F., Tiedemann, H., Prokott, E., Maiello, G., & Fleming, R. W. (2021). An image-computable model of human visual shape similarity. *PLOS Computational Biology*, *17*(6), e1008981. https://doi.org/10.1371/journal.pcbi.1008981

Mou, Y., & Luo, Y. (2017). Is it a container? Young infants' understanding of containment events. *Infancy*, *22*(2), 256–270. https://doi.org/10.1111/infa.12148

Mou, Y., & vanMarle, K. (2014). Two core systems of numerical representation in infants. *Developmental Review*, *34*(1), 1–25. https://doi.org/10.1016/j.dr.2013.11.001

Muckli, L., Kohler, A., Kriegeskorte, N., & Singer, W. (2005). Primary visual cortex activity along the apparent-motion trace reflects illusory perception. *PLOS Biology*, *3*(8), e265. https://doi.org/10.1371/journal.pbio.0030265

Müller-Lyer, F.C. (1889). Optische Urtheilstäuschungen. *Archiv für Anatomie und Physiologie, Physiologische Abteilung, Supplement-Band*, 263–270.

Munhall, K. G., Gribble, P., Sacco, L., & Ward, M. (1996). Temporal constraints on the McGurk effect. *Perception & Psychophysics*, *58*(3), 351–362. https://doi.org/10.3758/BF03206811

Nakai, T., Girard, C., Longo, L., Chesnokova, H., & Prado, J. (2023). Cortical representations of numbers and nonsymbolic quantities expand and segregate in children from 5 to 8 years of age. *PLOS Biology*, *21*(1), e3001935. https://doi.org/10.1371/journal.pbio.3001935

Nakayama, K. (2011). Introduction: Vision going social. In R. B. Adams, N. Ambady, K. Nakayama & S. Shimojo (Eds.), *The Science of Social Vision*, (pp. xv–xxviii). Oxford University Press.

Neri, P., Luu, J. Y., & Levi, D. M. (2006). Meaningful interactions can enhance visual discrimination of human agents. *Nature Neuroscience*, *9*(9), 1186–1192. https://doi.org/10.1038/nn1759

Nieder, A. (2019). *A brain for numbers: the biology of the number instinct*. MIT press. https://doi.org/10.7551/mitpress/11565.001.0001

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*(5719), 255–258. https://doi.org/10.1126/science.1107621

Palmer, S. E., & Hemenway, K. (1978). Orientation and symmetry: effects of multiple, rotational, and near symmetries. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(4), 691. https://doi.org/10.1037/0096-1523.4.4.691

Palumbo, R., D'Ascenzo, S., & Tommasi, L. (2015). Cross-category adaptation: exposure to faces produces gender aftereffects in body perception. *Psychological Research*, *79*, 380–388. https://doi.org/10.1007/s00426-014-0576-2

Papeo, L. (2020). Twos in human visual perception. *Cortex*, *132*, 473–478. https://doi.org/10.1016/j.cortex.2020.06.005

Papeo, L., Goupil, N., & Soto-Faraco, S. (2019). Visual search for people among people. *Psychological Science*, *30*(10), 1483–1496. https://doi.org/10.1177/0956797619867295

Papeo, L., Stein, T., & Soto-Faraco, S. (2017). The two-body inversion effect. *Psychological Science*, *28*(3), 369–379. https://doi.org/10.1177/0956797616685769

Papeo, L., Vettori, S., Serraille, E., Odin, C., Rostami, F., & Hochmann, J. R. (2024). Abstract thematic roles in infants' representation of social events. *Current Biology*, *34*(18), 4294–4300. https://doi.org/10.1016/j.cub.2024.07.081

Park, J., DeWind, N. K., Woldorff, M. G., & Brannon, E. M. (2016). Rapid and direct encoding of numerosity in the visual stream. *Cerebral Cortex*, *26*(2), 748–763. https://doi.org/10.1093/cercor/bhv017

Pavlova, M. A. (2012). Biological motion processing as a hallmark of social cognition. *Cerebral Cortex*, *22*(5), 981–995. https://doi.org/10.1093/cercor/bhr156

Pereira, A. F., & Smith, L. B. (2009). Developmental changes in visual object recognition between 18 and 24 months of age. *Developmental Science*, *12*(1), 67–80. https://doi.org/10.1111/j.1467-7687.2008.00747.x

Perez, J., & Feigenson, L. (2022). Violations of expectation trigger infants to search for explanations. *Cognition*, *218*, 104942. https://doi.org/10.1016/j.cognition.2021.104942

Phillips, I., & Firestone, C. (2023). Visual adaptation and the purpose of perception. *Analysis*, *83*, 555–575. https://doi.org/10.1093/analys/anac060

Piaget, J. (1954). *The construction of reality in the child*. (M. Cook, Trans.). Basic Books. https://doi.org/10.1037/11168-000

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron*, *44*(3), 547–555. https://doi.org/10.1016/j.neuron.2004.10.014

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*(5695), 499–503. https://doi.org/10.1126/science.1102085

Ponzo, M. (1910). Intorno ad alcune illusioni nel campo delle sensazioni tattili sull'illusione di Aristotele e fenomeni analoghi [On some tactile illusions, Aristotle's illusion, and similar phenomena]. *Archive für die Gesamte Psychologie*, *16*, 307–345.

Pramod, R. T., Cohen, M. A., Tenenbaum, J. B., & Kanwisher, N. (2022). Invariant representation of physical stability in the human brain. *Elife*, *11*, e71736. https://doi.org/10.7554/eLife.71736

Pulfrich, C. (1922). Die Stereoskopie im Dienste der isochromen und heterochromen Photometrie. *Naturwissenschaften*, *10*(35), 751–761. https://doi.org/10.1007/BF01565171

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial vision*, *3*(3), 179–197. https://doi.org/10.1163/156856888x00122

Quilty-Dunn, J. (2020). Concepts and predication from perception to cognition. *Philosophical Issues*, *30*, 273–292. https://doi.org/10.1111/phis.12185

Quilty-Dunn, J., Porot, N., & Mandelbaum, E. (2023). The best game in town: The reemergence of the language-of-thought hypothesis across the cognitive sciences. *Behavioral and Brain Sciences*, *46*, e261. doi:10.1017/S0140525X22002849

Rauschenberger, R., & Yantis, S. (2001). Masking unveils pre-amodal completion representation in visual search. *Nature*, *410*(6826), 369–372. https://doi.org/10.1038/35066577

Regan, D., & Beverley, K. I. (1979). Binocular and monocular stimuli for motion in depth: Changing-disparity and changing-size feed the same motion-in-depth stage. *Vision Research*, *19*(12), 1331–1342. https://doi.org/10.1016/0042-6989(79)90205-0

Rolfs, M., Dambacher, M., & Cavanagh, P. (2013). Visual adaptation of the perception of causality. *Current Biology*, *23*(3), 250–254. https://doi.org/10.1016/j.cub.2012.12.017

Sablé-Meyer, M., Ellis, K., Tenenbaum, J. B., & Dehaene, S. (2022). A language of thought for the mental representation of geometric shapes. *Cognitive Psychology*, *139*, 101527. https://doi.org/10.1016/j.cogpsych.2022.101527

Sann, C., & Streri, A. (2007). Perception of object shape and texture in human newborns: evidence from cross-modal transfer tasks. *Developmental Science*, *10*(3), 399–410. https://doi.org/10.1111/j.1467-7687.2007.00593.x

Saxe, R., & Carey, S. (2006). The perception of causality in infancy. *Acta Psychologica*, *123*(1–2), 144–165. https://doi.org/10.1016/j.actpsy.2006.05.005

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, *80*(1–2), 1–46. https://doi.org/10.1016/S0010-0277(00)00152-9

Scholl, B. J., & Gao, T. (2013). Perceiving animacy and intentionality: Visual processing or higher-level judgment? In M. D. Rutherford & V. A. Kuhlmeier (Eds.), *Social perception: Detection and interpretation of animacy, agency, and intention* (pp. 197–229). Boston Review. https://doi.org/10.7551/mitpress/9780262019279.003.0009

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). Blackwell.

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology*, *38*(2), 259–290. https://doi.org/10.1006/cogp.1998.0698

Skerry, A. E., Carey, S. E., & Spelke, E. S. (2013). First-person action experience reveals sensitivity to action efficiency in prereaching infants. *Proceedings of the National Academy of Sciences*, *110*(46), 18728–18733. https://doi.org/10.1073/pnas.1312322110

Slater, A., Mattock, A., & Brown, E. (1990). Size constancy at birth: Newborn infants' responses to retinal and real size. *Journal of Experimental Child Psychology*, *49*(2), 314–322. https://doi.org/10.1016/0022-0965(90)90061-C

Slater, A., Mattock, A., Brown, E., & Bremner, J. G. (1991). Form perception at birth: Cohen and Younger (1984) revisited. *Journal of Experimental Child Psychology*, *51*(3), 395–406. https://doi.org/10.1016/0022-0965(91)90084-6

Smith, L. B., Street, S., Jones, S. S., & James, K. H. (2014). Using the axis of elongation to align shapes: Developmental changes between 18 and 24 months of age. *Journal of Experimental Child Psychology*, *123*, 15–35. https://doi.org/10.1016/j.jecp.2014.01.009

Spelke, E. S. (1988). The origins of physical knowledge. In L. Weiskrantz (Ed.), *Thought without language* (pp. 168–184). Clarendon Press/Oxford University Press.

Spelke, E. S. (2022). *What Babies Know: Core Knowledge and Composition*. Oxford University Press. https://doi.org/10.1093/oso/9780190618247.001.0001

Spelke, E. S. (2024). Précis of *What Babies Know*. *Behavioral and Brain Sciences*, *47*, e120. https://doi.org/10.1017/S0140525X23002443

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, *99*(4), 605–632. https://doi.org/10.1037/0033-295X.99.4.605

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*(1), 89–96. https://doi.org/10.1111/j.1467-7687.2007.00569.x

Spelke, E., Lee, S. A., & Izard, V. (2010). Beyond core knowledge: Natural geometry. *Cognitive Science*, *34*(5), 863–884. https://doi.org/10.1111/j.1551-6709.2010.01110.x

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

Spriet, C., Abassi, E., Hochmann, J. R., & Papeo, L. (2022). Visual object categorization in infancy. *Proceedings of the National Academy of Sciences*, *119*(8), e2105866119. https://doi.org/10.1073/pnas.2105866119

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science*, *348*(6230), 91–94. https://doi.org/10.1126/science.aaa3799

Stahl, A. E., & Feigenson, L. (2019). Violations of core knowledge shape early learning. *Topics in Cognitive Science*, *11*(1), 136–153. https://doi.org/10.1111/tops.12389

Strickland, B., & Scholl, B. J. (2015). Visual perception involves 'event type' representations: The case of containment vs. occlusion. *Journal of Experimental Psychology: General*, *144*, 570–580. https://doi.org/10.1037/a0037750

Sullivan, J., & Barner, D. (2014). Inference and association in children's early numerical estimation. *Child Development*, *85*(4), 1740–1755. https://doi.org/10.1111/cdev.12211

Sun, Z., & Firestone, C. (2021). Curious objects: How visual complexity guides attention and engagement. *Cognitive Science*, *45*(4), e12933. https://doi.org/10.1111/cogs.12933

Szkudlarek, E., & Brannon, E. M. (2021). First and second graders successfully reason about ratios with both dot arrays and arabic numerals. *Child Development*, *92*(3), 1011–1027. https://doi.org/10.1111/cdev.13470

Tatone, D., Geraci, A., & Csibra, G. (2015). Giving and taking: Representational building blocks of active resource-transfer events in human infants. *Cognition*, *137*, 47–62. https://doi.org/10.1016/j.cognition.2014.12.007

Tatone, D., Hernik, M., & Csibra, G. (2021). Facilitation of object encoding in infants by the observation of giving. *Scientific Reports*, *11*(1), 18305. https://doi.org/10.1038/s41598-021-97910-3

Teichmann, L., Edwards, G., & Baker, C. I. (2021). Resolving visual motion through perceptual gaps. *Trends in Cognitive Sciences*, *25*(11), 978–991. https://doi.org/10.1016/j.tics.2021.07.017

Thomas, A. J. (2024). Cognitive Representations of Social Relationships and their Developmental Origins. *Behavioral and Brain Sciences*, 1–53. https://doi.org/10.1017/S0140525X24001328

Thomas, A. J., Woo, B., Nettle, D., Spelke, E. S., & Saxe, R. (2022). Early concepts of intimacy: Young humans use saliva sharing to infer close relationships. *Science*, *375*(6578), 311–315. https://doi.org/10.1126/science.abh1054

Thomsen, L., Frankenhuis, W. E., Ingold-Smith, M., & Carey, S. (2011). Big and mighty: Preverbal infants mentally represent social dominance. *Science*, *331*(6016), 477–480. https://doi.org/10.1126/science.1199198

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522. https://doi.org/10.1038/381520a0

Turati, C., Simion, F., & Zanon, L. (2003). Newborns' perceptual categorization for closed and open geometric forms. *Infancy*, *4*(3), 309–325. https://doi.org/10.1207/S15327078IN0403_01

Valenti, J. J., & Firestone, C. (2019). Finding the "odd one out": Memory color effects and the logic of appearance. *Cognition*, *191*, 103934. https://doi.org/10.1016/j.cognition.2019.04.003

van Buren, B., Uddenberg, S., & Scholl, B. J. (2016). The automaticity of perceiving animacy: Goal-directed motion in simple shapes influences visuomotor behavior even when task-irrelevant. *Psychonomic Bulletin & Review*, *23*(3), 797–802. https://doi.org/10.3758/s13423-015-0966-5

Van Oeffelen, M. P., & Vos, P. G. (1982). A probabilistic model for the discrimination of visual number. *Perception & Psychophysics*, *32*(2), 163–170. https://doi.org/10.3758/BF03204275

Van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences*, *102*(4), 1181–1186. https://doi.org/10.1073/pnas.0408949102

Van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*, *45*(3), 598–607. https://doi.org/10.1016/j.neuropsychologia.2006.01.001

vanMarle, K., & Scholl, B. J. (2003). Attentive tracking of objects versus substances. *Psychological Science*, *14*(5), 498–504. https://doi.org/10.1111/1467-9280.03451

Vettori, S., Odin, C., Hochmann, J., & Papeo, L. (2025). A perceptual cue-based mechanism for automatic assignment of thematic agent and patient roles. *Journal of Experimental Psychology: General*, *154*(3), 787–798. https://doi.org/10.1037/xge0001657

Wagemans, J. (1997). Characteristics and models of human symmetry detection. *Trends in Cognitive Sciences*, *1*(9), 346–352. https://doi.org/10.1016/S1364-6613(97)01105-4

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. https://doi.org/10.1016/j.cognition.2003.09.012

Wang, S. H., & Onishi, K. H. (2017). Enhancing young infants' representations of physical events through improved retrieval (not encoding) of information. *Journal of Cognition and Development*, *18*, 289–308. https://doi.org/10.1080/15248372.2017.1281279

Webster, M. A. (2015). Visual adaptation. *Annual Review of Vision Science*, *1*(1), 547–567. https://doi.org/10.1146/annurev-vision-082114-035509

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature*, *428*(6982), 557–561. https://doi.org/10.1038/nature02420

Weigelt, S., Koldewyn, K., & Doehrmann, O. (2010). Cross-category adaptation reveals tight coupling of face and body perception. *Journal of Neurophysiology*, *104*(2), 581–583. https://doi.org/10.1152/jn.00288.2010

Wertheim, J., & Ragni, M. (2020). The neurocognitive correlates of human reasoning: A meta-analysis of conditional and syllogistic inferences. *Journal of Cognitive Neuroscience*, *32*(6), 1061–1078. https://doi.org/10.1162/jocn_a_01531

Wheatstone, C. (1838). On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London*, *33*, 371–394.

Wilson, J. A., & Robinson, J. O. (1986). The impossibly twisted Pulfrich pendulum. *Perception*, *15*(4), 503–504. https://doi.org/10.1068/p150503

Wilson, V. A., Sauppe, S., Brocard, S., Ringen, E., Daum, M. M., Wermelinger, S., Gu, N., Andrews, C., Isasi-Isasmendi, A., Bickel, B., & Zuberbühler, K. (2024). Humans and great apes visually track event roles in similar ways. *PLOS Biology*, *22*(11), e3002857. https://doi.org/10.1371/journal.pbio.3002857

Witzel, C. (2016). An easy way to show memory color effects. *I-Perception*, *7*(5). https://doi.org/10.1177/2041669516663751

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1–34. https://doi.org/10.1016/S0010-0277(98)00058-4

Woodward, A. L. (1999). Infants' ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behavior and Development*, *22*(2), 145–160. https://doi.org/10.1016/S0163-6383(99)00007-7

Xu, F. (2019). Towards a rational constructivist theory of cognitive development. *Psychological Review*, *126*(6), 841–864. https://doi.org/10.1037/rev0000153

Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive psychology*, *30*(2), 111–153. https://doi.org/10.1006/cogp.1996.0005

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), B1–B11. https://doi.org/10.1016/S0010-0277(99)00066-9

Yantis, S. (1993). Stimulus-driven attentional capture and attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(3), 676–681. https://doi.org/10.1037/0096-1523.19.3.676

Yin, J., & Csibra, G. (2015). Concept-based word learning in human infants. *Psychological Science*, *26*(8), 1316–1324. https://doi.org/10.1177/0956797615588753

Yousif, S., Clarke, S., & Brannon, E. (2024). Number adaptation: A critical look. *Cognition*, *249*, 105813. https://doi.org/10.1016/j.cognition.2024.105813

Zanon, M., Lemaire, B. S., Papeo, L., & Vallortigara, G. (2024). Innate sensitivity to face-to-face biological motion. *iScience*, *27*(2). https://doi.org/10.1016/j.isci.2024.108793

Zhao, J., & Yu, R. Q. (2016). Statistical regularities reduce perceived numerosity. *Cognition*, *146*, 217–222. https://doi.org/10.1016/j.cognition.2015.09.018

Zhou, K., Luo, H., Zhou, T., Zhuo, Y., & Chen, L. (2010). Topological change disturbs object continuity in attentive tracking. *Proceedings of the National Academy of Sciences*, *107*(50), 21920–21924. https://doi.org/10.1073/pnas.1010919108