



Cognitive Science 34 (2010) 1158–1184
 Copyright © 2010 Cognitive Science Society, Inc. All rights reserved.
 ISSN: 0364-0213 print / 1551-6709 online
 DOI: 10.1111/j.1551-6709.2010.01127.x

How Infants Learn About the Visual World

Scott P. Johnson

Department of Psychology, University of California, Los Angeles

Received 22 August 2008; received in revised form 30 April 2010; accepted 20 May 2010

Abstract

The visual world of adults consists of objects at various distances, partly occluding one another, substantial and stable across space and time. The visual world of young infants, in contrast, is often fragmented and unstable, consisting not of coherent objects but rather surfaces that move in unpredictable ways. Evidence from computational modeling and from experiments with human infants highlights three kinds of learning that contribute to infants' knowledge of the visual world: learning via association, learning via active assembly, and learning via visual-manual exploration. Infants acquire knowledge by observing objects move in and out of sight, forming associations of these different views. In addition, the infant's own self-produced behavior—oculomotor patterns and manual experience, in particular—is an important means by which infants discover and construct their visual world.

Keywords: Visual development; Cognitive development; Models of development; Object perception; Infants; Learning

1. Introduction

When we look around us, we encounter environments characterized by numerous objects and people at varying distances. The scene depicted in Fig. 1 is typical. It shows a garden in Southern California occupied by flora, people, and artifacts such as buildings and walkways. The scene is busy and cluttered. The objects have multiple parts and are located at various distances from the observer; nearer objects obscure farther objects—the trees hide part of the building's roof, and many of the people on the patio cannot be seen. In some instances, color, texture, size, and shape can serve as information for the unity of objects. The leaves on the trees, for example, are all roughly similar in appearance, and they are perceived as

Correspondence should be sent to Scott P. Johnson, Department of Psychology, Franz Hall, UCLA, Los Angeles, CA 90095. E-mail: scott.johnson@ucla.edu



Fig. 1. A visual scene.

grouping together. In other instances, however, we see objects as unified despite considerable discrepancies in these kinds of visual information: The girls on the stairs wear blue shorts and red shirts, yet we do not see these distinctions in color as denoting four “parts” of girls, but instead we see them as belonging to objects in common. In real-world scenes, motion of observers and of objects provides additional information to determine the contents of our surroundings. We can move through the environment, obtain new perspectives, and see parts of objects invisible from previous vantage points. And as objects move, we can track them across periods of temporary invisibility, often predicting their reappearance.

These facts about the visual world are at once ordinary and remarkable. They are ordinary because virtually every sighted observer experiences the environment as composed of separate, distinct objects of varying complexity and appearance. Yet they are nonetheless remarkable because of the intricate cortical machinery needed to produce them (Zeki, 1993): Several dozen areas of the brain, each responsible for processing a distinct aspect of the visual scene or coordinating the outputs of other areas, working in concert to yield a more-or-less seamless and coherent experience. Action systems are likewise elaborate (Gibson, 1950): eyes, head, and body, each with distinct control systems, working in concert to explore the visual environment.

In this article, I consider theory and research on the development of infants’ perception of the visual environment, in particular object perception. The garden example illustrates some of the issues faced by researchers who wish to better understand these processes. Visual perception has a “bottom-up” foundation built upon coding and integrating distinctive kinds of visual information (variations in color, luminance, distance, texture, orientation, motion, and so forth). Visual perception also relies on “top-down” knowledge that observers bring to the scene, knowledge that aids in interpreting potentially ambiguous juxtapositions of visual attributes (such as the blue and red clothing on the girls). Both operate continuously in mature, sighted individuals who have had sufficient time and experience with which to learn about specific kinds of objects.

Young infants have not had as much time and experience with which to learn about objects, yet they inhabit the same world as adults and they encounter the same kinds of visual information. How do infants interpret visual scenes? Are they restricted to bottom-up processing, lacking the cognitive capacity to interpret visual information in a meaningful fashion, however that might be defined? Or might there be some capacity to perceive objects that is independent of visual experience? These questions have long been dominated by a tension between arguments for a learned or *constructed* versus an unlearned or *innate* system of object knowledge in humans. This article will examine the question of infants' object perception by considering these theoretical perspectives, and it attempts to answer the question with evidence from modeling of developmental processes and from empirical studies. I will restrict discussion to the developmental origins, in humans, of the ability to represent objects as coherent and complete across space and time—that is, despite partial or full occlusion—a definition of object perception akin to the *object concept* that was originally described by the eminent developmental psychologist Jean Piaget (more on this subsequently). The limited scope necessarily omits many interesting literatures on other topics related to object knowledge, such as object identity, numerosity, animacy, object-based attention, and so forth. Nevertheless, there has been a great deal of research effort directed at object concept development, and these investigations continue to bear on the question of infants' object perception by providing an increasingly rich base of evidence.

2. Theoretical considerations: Constructivism versus nativism

A perennial question of infant epistemology is the extent to which knowledge necessarily develops over time through learning and experience or whether some kinds of knowledge are available—“built in”—without any experience. This question has motivated innumerable experiments investigating infant knowledge of the physical and social environment (for recent reviews, see Carey, 2009; Johnson, 2010), and it has not yet been fully satisfied, although there have been suggestions to abandon it in light of mounting evidence from systems theory for multiple levels of developmental process (Spencer et al., 2009). I will illustrate these opposing views by outlining two theories of infant object perception: constructivist theory and nativist theory.

2.1. Constructivist theory

Piaget was the first to provide a detailed theory of development of object knowledge in humans, and he amassed a great deal of evidence to support it (Piaget, 1954). The evidence was derived from observations of infants and young children as he engaged them in child-friendly games, such as hiding a desired toy under a cover and then observing the child's attempts to retrieve it. Piaget suggested that knowledge of objects and knowledge of space were inseparable—without knowledge of spatial relations, there could be no knowledge of objects *qua* distinct entities. Instead, objects, early in life, were disconnected forms that moved capriciously and randomly.

Objectification was Piaget's term for knowledge of the self and external objects as distinct entities, spatially segregated, persisting across time and space, and obeying causal constraints. Piaget suggested that objectification is rooted in the child's recognition of her own body as an independent object and her own movements as movements of objects through space, corresponding to movements of other objects she sees. This recognition, in turn, was thought to stem exclusively from exploration of objects via manual activity—that is, object knowledge is constructed by the child.

Because skilled reaching and grasping of objects is not available until some time after 4–6 months, Piaget proposed that object knowledge likewise is beyond the ken of young infants. Until this time, objects consist of little more than fleeting images, as noted previously. Active, intentional search behavior marks the inception of objectification, beginning with infants' visual accommodation to dropped objects (by looking at the floor), and culminating with infants' systematic, organized search for hidden objects in multiple possible locations (in one of Piaget's more complicated games) some time during the second year. As the action systems develop, therefore, so develops the capacity to interact with objects and discover their properties, most importantly the maintenance of object properties across disruptions in perceptual contact, as when objects go out of sight.

Piaget's theory enjoys strong support for many of the details of behavior that he so assiduously captured, and for raising awareness of the problems faced by infants as they navigate the visual world. The reasoning behind the developmental changes in behavior, however, has not seen the same level of enthusiasm. Numerous experiments have revealed that by 2–4 months, infants appear to maintain representations of partly and fully hidden objects across short delays, somewhat younger than allowed for on Piaget's account, and inconsistent with the emphasis on manual activity as the principal agent of developmental change. These experiments have led to views of infant object knowledge as relying on unlearned, or innate, foundations, and some of these views are described in the following section.

2.2. Nativist theory

A central tenet of nativist theory is that some kinds of initial, unlearned knowledge form a central core around which more diverse, mature cognitive capacities are elaborated (Spelke, 1990; Spelke, Breinlinger, Macomber, & Jacobson, 1992). That is, some kinds of knowledge, including concepts of objects as coherent and continuous across occlusion, are *innate*. Philosophical discussions of innateness are ancient. Plato and Descartes, for example, proposed that some ideas, such as concepts of geometry or God or justice, were universal and available innately because they were unobservable or arose in the absence of any direct tutoring or instruction. With respect to infant knowledge, the focus of modern nativist theory is on *learnability*: According to this line of reasoning, in the absence of any viable theory of how humans come to understand object concepts so quickly after birth, in some cases well in advance of the manual search skills emphasized by Piaget, the assumption is that these concepts necessarily arise independent of experience.

Researchers of a nativist persuasion have offered three arguments for hypothesized innate object concepts. First, veridical object knowledge can be elicited in very young infants—as

young as 2 months of age, or perhaps even at birth—under a variety of circumstances, suggesting that early concepts emerge too quickly to have derived from postnatal learning. Second, infants' acquisition of object knowledge has been proposed to arise from *contrastive evidence*: opportunities to observe conditions under which an object behaves in a manner consistent or inconsistent with a particular concept (Baillargeon, 1994). On this view, a concept of persistence across occlusion, for example, must be innate, because there are no opportunities in the real world to observe objects going out of existence! Third, there is evidence from one nonhuman species—domestic chickens—for an unlearned capacity for *unity perception*, recognition of partly occluded shapes as similar to an unoccluded version of the same form (Regolin & Vallortigara, 1995). (In this experiment, newly hatched chicks were imprinted on a partly occluded cardboard triangle, and they subsequently chose to associate with a fully visible version of the triangle, rather than a version consisting of the previously seen triangle fragments. An experiment in which the imprinting/association objects were reversed showed the same result.)

As noted, there is compelling evidence from a variety of laboratories and experimental settings for representations of objects as solid entities that are spatiotemporally coherent and persistent by 2–4 months after birth. (Some of these experiments will be discussed in detail in subsequent sections of this article.) Nevertheless, the *developmental origins* of object concepts in human infants as a topic of investigation cannot be dismissed merely by noting competence in these experiments at a young age. Moreover, suggesting that infants learn about objects through only a single means, such as contrastive evidence, does not seem realistic. Unequivocal support for *innate* object concepts in humans would come from evidence for their emergence in the absence of visual experience—say, functionality at birth, or veridical object perception in blind individuals who have their sight restored. As we will see, experiments on infants' perception of partly occluded objects cast doubt on the viability of any of these varieties of innateness as the best descriptor of the development of object concepts. (Parenthetically, it is worth noting as well that the finding of unity perception in chicks remains an isolated phenomenon in the literature, inconsistent with experiments with another avian species—pigeons—which, as adults, apparently see partly occluded objects only in terms of their visible surfaces, not as having hidden parts; Sekuler, Lee, & Shettleworth, 1996.)

2.3. Summary

Piaget set the stage for decades of fruitful research that has established the availability of functional object concepts in the first year after birth, and Piaget's own theory of how knowledge may be constructed by the child's own behavior has been tremendously influential. Numerous experiments in the past few decades, however, have suggested that Piaget underestimated young infants' capacity to perceive object unity and boundaries under occlusion, so much so that alternate theories stressing innate contributions to object knowledge have appeared.

In the next section of this article, I describe evidence for developmental change early in postnatal life in how infants respond to partly and fully hidden objects. Evidence comes

from experiments that assess three kinds of perceptual completion in infancy: *spatial completion* (perception of partly hidden surfaces as continuous), *spatiotemporal completion* (perception of objects as continuing to exist upon becoming occluded), and *3D object completion* (perception of objects as solid in 3D space, with a back and sides that cannot be viewed from any one vantage point). Because evidence for developmental change requires an explanation, in the sections after that I describe models of development and empirical investigations of developmental mechanisms in infants. These models and investigations posit a central role for learning, and they suggest specific means by which learning—in models and in human infants—can lead to the kinds of object knowledge I have discussed.

3. Developmental change in infants' object perception

Piaget's observations led him to conclude that newborn infants have no true concepts of objects or space (Piaget, 1952). Neonates can discriminate among visible objects and track their motions, but when objects move out of sight or the baby's gaze falls elsewhere, previously encountered objects cease to exist for the infant. The first inklings of object concepts come from recognition memory, say when infants smile upon mother's return, beginning a few months after birth. Knowledge of objects as complete and coherent across gaps in perceptual contact imposed by occlusion did not come until later, after infants can grasp and reach objects and thereby come to more fully appreciate properties such as solidity, volume, and existence independent of the infant. These concepts did not come all at once. Piaget described one important concept as "reconstruction of an invisible whole from a visible fraction," and it was evinced by retrieval of an object from under a cover when only a part of the object was visible. An appreciation of continued existence despite complete occlusion was evinced by removal of an obstacle hiding a desired toy, or pulling away a cover from a parent's face during peekaboo. These behaviors were not seen consistently until 6–8 months or so, marking for Piaget the advent of a wholly new set of object and spatial concepts.

As noted previously, research over the past several decades has yielded a wealth of evidence from multiple measures (e.g., looking, reaching, and cortical activity) making it clear that infants represent object properties even when the objects are partly or fully hidden, and these findings have led, in turn, to nativist theoretical views that have sought to overturn Piagetian ideas about how these representations arise in infants. Yet two important facts remain, facts suggesting that Piagetian theory may not be so far off the mark as concerns developmental changes in infants' object knowledge. First, newborn infants do not seem to perceive partly occluded objects as having hidden parts. Instead, neonates construe such stimuli solely in terms of their visible parts, failing to achieve spatial completion (Slater, Johnson, Brown, & Badenoeh, 1996; Slater et al., 1990; but see Valenza, Leo, Gava, & Simion, 2006). There is a clear developmental progression in perceptual completion (Johnson, 2004), and this calls for an explanation of underlying mechanisms of change. And second, the majority of research on infants' perceptual completion in reality is broadly consistent with Piaget's observations: Infants provide evidence of representing partly occluded objects a few months after birth, and fully occluded objects by about the middle of the first

year. Some have claimed object permanence in infants on the basis of evidence from looking time studies (e.g., Baillargeon, 2008), but the short-term representations of hidden objects demonstrated in such experiments fall short of Piaget's criteria for full object permanence: accurate search in multiple locations for a hidden object, demonstrating knowledge of object persistence and the spatial relations between the object, the hiding locations, and the infant (Haith, 1998; Kagan, 2008).

In the remainder of this section, I will describe investigations of perceptual completion in infants. These investigations provide clear evidence for developmental change in how infants perceive occlusion events.

3.1. Spatial completion

Adults and 4-month-old infants construe the “rod-and-box” display depicted in Fig. 2, left, as consisting of two parts, a single elongated object moving back and forth behind an occluding box (Kellman & Spelke, 1983). Neonates, in contrast, construe this display as consisting of *three* parts: two distinct object parts and occluder (Slater et al., 1990, 1996). These conclusions stem from looking time experiments in which infants first view the rod-and-box display repeatedly until habituation of looking occurs, defined as a decline in looking times toward the display (judged by an observer) according to a predetermined criterion. Following habituation, infants see two new displays, and their posthabituation looking patterns are thought to reflect a novelty preference. The 4-month-olds and neonates showed opposite patterns of preference. Looking longer at the “broken” rod parts indicates that they were relatively novel compared to the rod-and-box display—the 4-month-olds' response, suggestive of unity perception. Looking longer at the “complete” rod indicates that infants likely construed the rod-and-box display as composed of disjoint objects—the newborns' response. These results led to the more general conclusion that neonates are unable to perceive occlusion, and that occlusion perception emerges over the first several postnatal months (Johnson, 2004). That is, “piecemeal” or fragmented perception of the visual environment extends from birth through the first several months afterwards.

Two-month-olds were found to show an “intermediate” pattern of performance—no reliable posthabituation preference—implying that spatial completion is developing at this

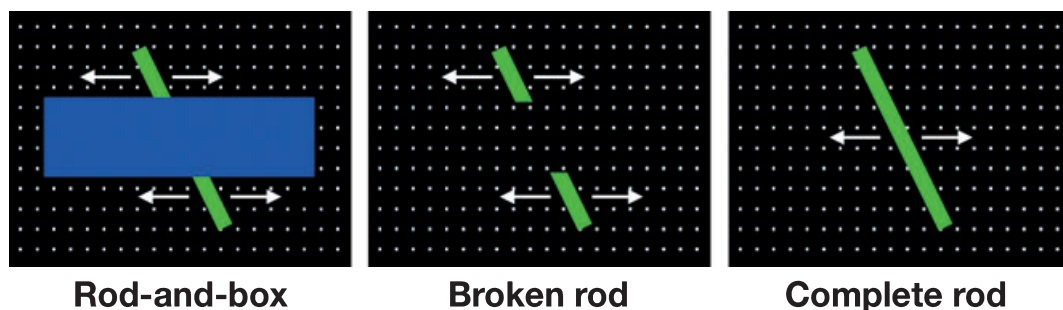


Fig. 2. Displays used in experiments that investigate spatial completion in young infants. Adapted from Johnson and Nájuez (1995).

point but not yet complete (Johnson & Náñez, 1995). Additional studies examined the possibility that 2-month-olds will perceive unity if given additional perceptual support. The amount of visible rod surface revealed behind the occluder was enhanced by reducing box height and by adding gaps in it, and under these conditions 2-month-olds provided evidence of unity perception (Johnson & Aslin, 1995). (With newborns, however, this manipulation failed to reveal similar evidence—even with enhanced displays, newborns perceived the moving rod parts as disjoint objects; Slater et al., 1996; Slater, Johnson, Kellman, & Spelke, 1994.) These experiments served to pinpoint more precisely the time of emergence of spatial completion in infancy: the first several weeks or months after birth under typical circumstances.

Additional experiments explored the kinds of visual information infants use to perceive spatial completion. Kellman and Spelke (1983) reported that 4-month-olds perceived spatial completion only when the rod parts, with aligned outer edges, moved in tandem behind a stationary occluder. We replicated and extended this finding, showing in addition that 4-month-olds provided evidence of completion only when the rod parts were aligned (Johnson & Aslin, 1996). Later experiments revealed similar patterns of performance in 2-month-olds when tested using displays with different occluder sizes and edge arrangements, as seen in Fig. 3 (Johnson, 2004). Infants provided evidence of spatial completion obtained only when rod parts were aligned across a narrow occluder; in the other displays, infants provided evidence of disjoint surface perception.

One possible interpretation of these findings is that alignment, motion, and occluder width (i.e., the spatial gap) are interdependent contributions to spatial completion, such that common motion is detected most effectively when rod parts are aligned (Kellman & Arterberry, 1998). I examined this possibility by testing 2-month-olds' discrimination of different patterns of rod motion with varying orientations of rod parts and occluder widths. Under all tested conditions, infants discriminated the motion patterns, implying that motion discrimination was neither impaired nor facilitated by misalignment or occluder width (Johnson, 2004). It might be that motion contributes to infants' spatial completion in multiple ways, first serving to segment the scene into its constituent surfaces, and then serving to bind moving surfaces into a single object (Johnson, Davidow, Hall-Haro, & Frank, 2008).

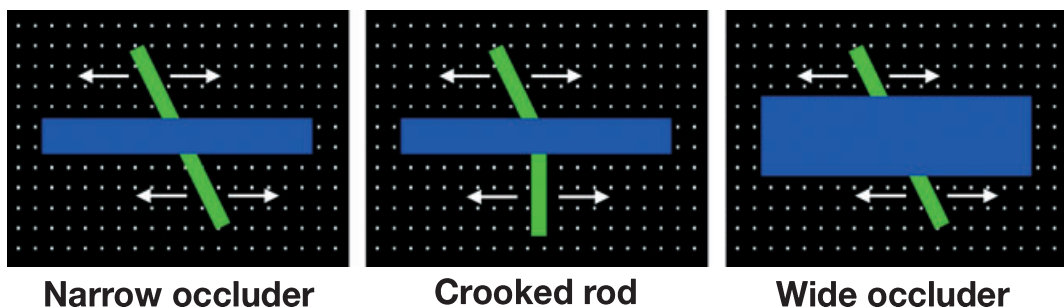


Fig. 3. Displays used to test the roles of occluder size and edge alignment in 2-month-olds' perception of spatial completion. Adapted from Johnson (2004).

In summary, experiments that explored development of spatial completion suggest that young infants analyze the motions and arrangements of visible surfaces. At birth, newborns perceive partly occluded surfaces as separate from one another and the background. Only later do infants integrate these surfaces into percepts of coherent, partly occluded objects. On this view, therefore, development of object knowledge begins with perception of visible object components, and it proceeds with increasing proficiency at representation of those object parts that cannot be discerned directly.

3.2. Spatiotemporal completion

A number of studies using different methods have shown that young infants can maintain representations for hidden objects across brief delays (e.g., Aguiar & Baillargeon, 1999; Berger, Tzur, & Posner, 2006; Clifton, Rochat, Litovsky, & Perris, 1991). Yet newborn infants provide little evidence of spatial completion, raising the question of how perception of *complete* occlusion emerges during the first few months after birth. Apart from Piaget's observations, this question has received little serious attention until recently, in favor of accounts that stress innate object concepts (e.g., Baillargeon, 2008; Spelke, 1990).

To address this gap in our knowledge, my colleagues and I conducted experiments with object trajectory displays, asking whether infants perceive the trajectory as continuous across occlusion—spatiotemporal completion. We reasoned that manipulation of spatial and temporal characteristics of the stimuli, and observation of different age groups, might provide insights into development of spatiotemporal completion, as they did in the case of spatial completion.

These investigations revealed a fragmented-to-holistic developmental pattern and revealed spatial and temporal processing constraints as well, both sets of results in parallel with the investigations of spatial completion described in the previous section. Spatiotemporal completion was tested using similar methods: habituation to an occlusion display (Fig. 4), followed by broken and complete test displays, different versions of the partly hidden trajectory seen during habituation. At 4 months, infants treat the ball-and-box display depicted in Fig. 4 as consisting of two disconnected trajectories, rather than a single, partly hidden path (Johnson, Bremner, et al. 2003); evidence comes from a reliable preference for the continuous version of the test trajectory. By 6 months, infants perceived this trajectory as unitary, as revealed by a reliable preference for the discontinuous trajectory test stimulus. When occluder size was narrowed, however, reducing the spatiotemporal gap across which the trajectory had to be interpolated, 4-month-olds' posthabituation preferences (and thus, by inference, their percepts of spatiotemporal completion) were shifted toward the discontinuous, partway by an intermediate width, and fully by a narrow width, so narrow as to be only slightly larger than the ball itself. In 2-month-olds, this manipulation appeared to have no effect.

Reducing the spatiotemporal gap, therefore, facilitates spatiotemporal completion. Reducing the *temporal* gap during which an object is hidden, independently from the *spatial* gap, also supports spatiotemporal completion. Increasing the ball size (Fig. 5) can minimize the time out of sight as it passes behind the occluder, and this led 4-month-olds to perceive

its trajectory as complete. Accelerating the speed of a smaller ball as it passed behind the occluder (and appeared more quickly) had a similar effect (Bremner et al., 2005). On the other hand, altering the orientation of the trajectory impaired path completion (Fig. 5), unless the edges of the occluder were orthogonal to the path; these findings are similar to outcomes of experiments on edge misalignment described in the previous section (Bremner et al., 2007).

This work leads to three conclusions. First, spatiotemporal completion proceeds from processing parts of paths to complete trajectories. Second, there may a lower age limit for trajectory completion (between 2 months and 4 months), just as there appears to be for spatial completion (between birth and 2 months). Third, young infants' spatiotemporal completion is based on relatively simple parameters. Either a short time or short distance out of sight leads to perception of continuity, and this may occur because the processing load is reduced by these manipulations. The fragile nature of emerging spatiotemporal completion is underscored as well by results showing its breakdown when either occluder or path orientation is nonorthogonal.

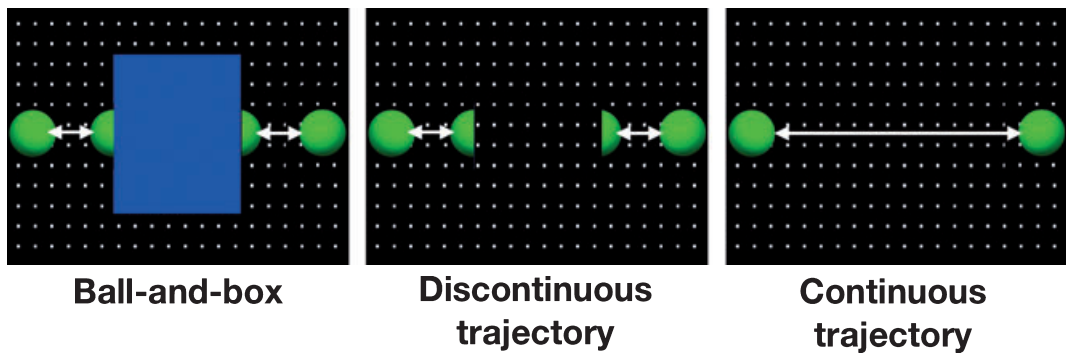


Fig. 4. Displays used in experiments that investigate spatiotemporal completion in young infants. Adapted from Johnson, Bremner, et al. (2003).

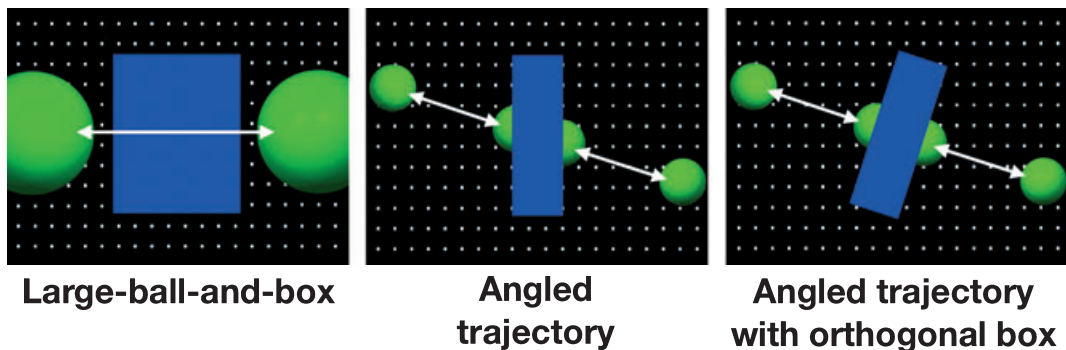


Fig. 5. Displays used to test the roles of occlusion duration and path orientation in 4-month-olds' perception of spatiotemporal completion. Adapted from Bremner et al. (2005, 2007).

3.3. 3D object completion

Spatial and spatiotemporal completion consist of filling in the gaps in object surfaces that have been occluded by nearer ones. Solid objects also occlude parts of themselves such that we cannot see their hidden surfaces from our present vantage point, yet our experience of most objects is that of filled volumes rather than hollow shells. Perceiving objects as solid in three-dimensional space despite limited views constitutes 3D object completion. In contrast to spatial and spatiotemporal completion, little is known about development of 3D object completion. We recently addressed this question with a looking time paradigm similar to those described previously (Soska & Johnson, 2008). Four- and six-month-olds were habituated to a wedge rotating through 15 degrees around the vertical axis such that the far sides were never revealed (Fig. 6). Following habituation infants viewed two test displays in alternation, one an incomplete, hollow version of the wedge, and the other a complete, whole version, both undergoing a full 360 degree rotation revealing the entirety of the object shape. Four-month-olds showed no consistent posthabituation preference, but 6-month-olds looked longer at the hollow stimulus, indicating perception of the wedge during habituation as a solid, volumetric object in 3D space.

In a follow-up study (Soska, Adolph, & Johnson, 2010), we used these same methods with a more complex stimulus: a solid “L”-shaped object with eight faces and vertices, as opposed to the five faces and six vertices in the wedge-shaped object described previously (Fig. 7). We tested 4-, 6-, and 9.5-month-olds. As in the Soska and Johnson (2008) study with the wedge stimulus, we found a developmental progression in 3D object completion: 4-month-olds’ posthabituation looking times revealed no evidence for completion, whereas 9.5-month-olds consistently looked longer at the hollow test display, implying perception of the habituation object as volumetric in 3D space. At 6 months, interestingly, only the male infants showed this preference; females looked about equally at the two test displays. At 9.5 months, the male advantage had disappeared: Both males and females looked longer at the hollow shape.

One interpretation of the sex difference at 6 months is that infants who were successful at 3D object completion engaged in *mental rotation* in this task: manipulation of a mental image of the object and imagining it from a different perspective. Mental rotation is a cognitive skill for which men have an advantage relative to women (Shepard & Metzler, 1971),

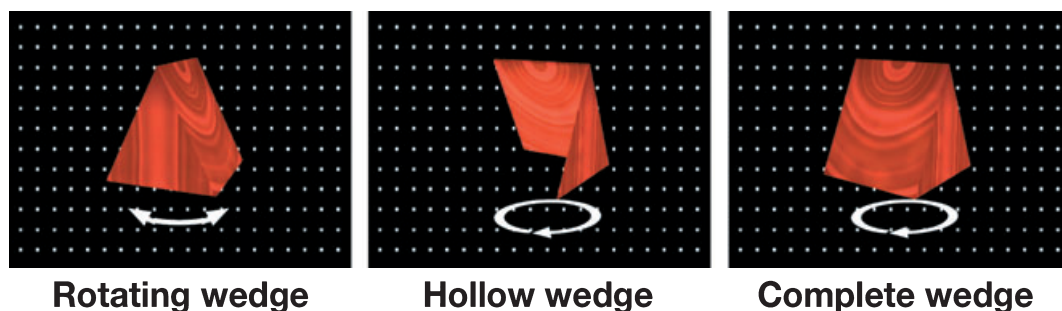


Fig. 6. Displays used in experiments that investigate 3D object completion in infants. Adapted from Soska and Johnson (2008).

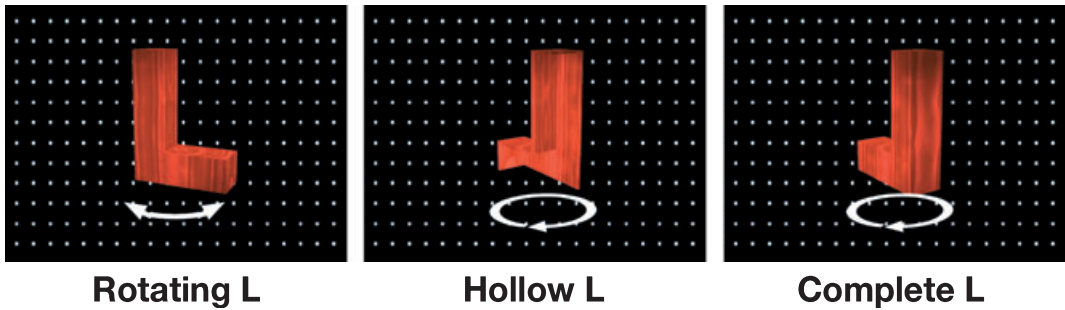


Fig. 7. Displays used in experiments that investigate 3D object completion in infants, with a more complex object relative to the Soska and Johnson (2008) study. Adapted from Soska and Johnson (unpublished data).

and two recent reports have provided evidence of a male advantage in young infants as well (Moore & Johnson, 2008; Quinn & Liben, 2008). It remains to be determined definitely whether mental rotation is involved in 3D object completion.

In summary, these data provide evidence for a developmental progression in infants' 3D object completion abilities, and for a role for stimulus complexity in infant performance. Both effects are consistent with the work on spatial and spatiotemporal completion described previously.

3.4. To be explained

The research described in this section can be summarized as follows. Infants are born with a functional visual system sufficient to identify distinct regions of the visual scene and discriminate different regions from one another. Newborns can detect edges and motion, and there is even a rudimentary capacity for depth perception (e.g., size and shape constancy; Slater, 1995). Yet newborns do not perceive objects as do adults, and therefore they do not “know” the world to consist of overlapping objects at different distances that have hidden parts. These kinds of knowledge arise over the first several postnatal months. I described three kinds of perceptual completion—spatial, spatiotemporal, and 3D object completion—and described evidence for a developmental progression in each.

How does this happen? One way to deal with this question is to ignore or deny it, which in essence is the nativist position as it is commonly presented in the infant cognition literature (e.g., Spelke, 1990; Spelke & Kinzler, 2007; Spelke et al., 1992). Considering cognition more broadly, this is not necessarily an illegitimate approach. There are phenomena in the literature that would appear to be impossible to explain otherwise, such as the emergence of linguistic structures in the absence of any input (Goldin-Meadow & Mylander, 1998; Senghas, Kita, & Özyürek, 2004). Such instances are rare, however, and this fact leads many of us to a second way to deal with the question of origins of object knowledge: Confront it and determine what is needed to account for the developmental progression I have presented.

In the remainder of the article, I consider arguments and evidence in favor of a learning account of how infants come to perceive object occlusion. The hypothesis is that infants learn certain features of the world such as edges (and perhaps faces) from prenatal

developmental mechanisms tantamount to a kind of “visual experience.” Occlusion, in contrast, must be learned postnatally.

4. Evidence for prenatal “visual experience” and memory

Newborn infants have two prerequisite skills for the ability to learn occlusion: They are born seeing and they are born with the capacity for short-term recall. The oculomotor system is sufficiently functional to guide the point of gaze to desired targets in the environment. Three kinds of stimulus are particularly salient: high-contrast edges, motion, and faces (Kessen, Salapatek, & Haith, 1972; Slater, Morison, Town, & Rose, 1985; Valenza, Simion, Macchi Cassia, & Umiltà, 1996). Newborn infants look at such stimuli preferentially—that is, they tend to look longer at high-contrast contours and patterns of motion relative to homogenous regions, for example, and at face-like patterns relative to arrangements of similar features that do not match human faces (Slater, 1995; Slater et al., in press). The developmental mechanisms that yield these behaviors, consequently, must be in effect prior to birth.

An interesting fact about prenatal visual development prior to the onset of patterned visual input is that there is spontaneous yet organized activity in visual pathways from early on, activity that contributes to retinotopic “mapping” (Sperry, 1963). Mapping refers to the preservation of sensory structure, for example, the relative positions of neighboring points of visual space, from retina through the thalamus, primary visual cortex, and higher visual areas. One way in which mapping occurs is by “waves” of coordinated, spontaneous firing of receptors in the retina, prior to eye opening, observed in some nonhuman species such as chicks and ferrets (Wong, 1999). Waves of activity are propagated across the retinal surface at a point in development after connections to higher visual areas have formed; the wave patterns are then systematically propagated through to the higher areas. This might be one way by which correlated inputs remain coupled and dissimilar inputs become dissociated, and a likely means by which edges (which can be defined as simple, local interactions in the input) can be detected upon exposure to patterned visual scenes (Albert, Schnabel, & Field, 2008). Retinal waves also can serve as a foundation for development of representations of more complex visual patterns as infants gain exposure to the environment (Bednar & Miikkulainen, 2007).

Evidence for short-term memory at birth was presented previously: Neonates habituate to repeated presentation of stimuli, implying recognition of familiar patterns, and they recover interest to new stimuli, implying recognition of novel patterns. The rudiments of memory, therefore, undergo developments prenatally sufficient to support recognition across brief delays. These developments include the emergence and strengthening of neural connections within and between regions of cortical and subcortical regions known in adult primates to maintain activity to visual and spatial input across temporary delays in presentation. These regions include areas of prefrontal cortex (e.g., the principal sulcus, anterior arcuate, and inferior convexity) and their connections to areas that are involved in object and face processing, such as the temporal lobe (Goldman-Rakic, 1987, 1996).

The newborn baby, therefore, is equipped with perceptual and cognitive mechanisms sufficient to begin the process of learning about objects—to detect edges in the visual scene, to track motion, to recognize familiar items, and to discriminate different items presented simultaneously or in sequence. These facts have motivated a number of attempts to explore the developmental process of coming to perceive and act on objects using connectionist or other kinds of computational models. Some of these models are described next.

5. Modeling developmental processes

How can computational models help us understand the developmental process of occlusion perception? Models can address questions of object perception development and other developmental phenomena by constraining hypotheses about preexisting skill sets, the necessary inputs from the environment, and specific learning regimens, and how these considerations influence learning outcomes. Models have five features in common: (a) specification of a starting point, (b) a particular kind of training environment, (c) a particular means of gathering and retaining information, (d) a particular means of expressing acquired knowledge, and (e) learning from experience—that is, modification of responses based on some change in internal representations of the external environment that stem from feedback. These features can be manipulated by the modeler to shed light on the developmental process, analogous to variations in experimental designs in the laboratory. (For a recent review of models of infant object perception, see Mareschal & Bremner, 2009.)

Models of object perception development that have been presented in the literature learn by association in a simple two- or three-dimensional environment. They are trained with particular inputs and are “queried” periodically for the state of their learning about hidden regions or continued existence, when occluded, of the objects they “see,” and performance interpreted in light of the starting points, environment, and so forth, as mentioned previously.

5.1. Modeling association learning

Mareschal and Johnson (2002) devised a connectionist model of unity perception based on three assumptions. First, infants can detect visual information relevant to object perception tasks prior to the effective utilization of this information. Second, experience viewing objects in motion is vital to perception of unity, because far objects move behind and emerge from near objects, providing support for the formation of associations between fully visible and partly occluded views. (In addition, young infants provide evidence of unity perception only when the visible parts undergo common motion; Kellman & Spelke, 1983.) Third, infants are equipped with short-term memory. None of these assumptions should reasonably be considered controversial or objectionable.

Given these assumptions, Mareschal and Johnson (2002) built a model using connectionist architecture: an input layer of inputs corresponding to a retina, hidden units whose computations formed representations of spatial completion, and an output layer that

provided a response in the form of a representation of a complete object, object parts, or an indeterminate response, after particular amounts of training in a specified environment. The architecture can be seen in Fig. 8, and an example of the training regimen can be seen in Fig. 9. Between the input and hidden units was a series of modules that processed the visual information in the training environment: the occluder, the motion patterns and orientation of rod parts, background texture, and points of intersection of the rod parts and occluder. The occluder remained stationary throughout each event, and the rod parts moved independently or in tandem. The model's task was to determine whether one or two objects (not including the occluder) were presented in the display. When the rod or rod parts were fully visible, the decision was accomplished directly, but it had to be inferred when there was an intersection of the rod and occluder. A memory trace of the previous portion of each event was stored and accumulated with increasing experience.

The model was set up to minimize the error between direct perception and the inference demanded by occlusion in response to training, and learned primarily by association: associations between views of partly hidden and fully visible "objects," and how unity perception was best predicted by the visual cues present in each display. Because human infants are especially sensitive to motion patterns and orientation of the rod parts and rely on these cues to perceive unity (Johnson, 1997), we trained the models with events in which these cues were available or absent, in different combinations, to examine their contributions to unity perception alone or in tandem with other cues.

The models were able to learn unity readily, and their performance was strongly affected by the cues available, in particular a combination of cues to which infants, likewise, are sensitive and use to perceive unity in the lab setting: common motion, orientation, and reliability of the rod parts, and T-junctions (cues that specify the intersection of the rod parts and occluder). These models, therefore, demonstrate that object knowledge, at some level, can

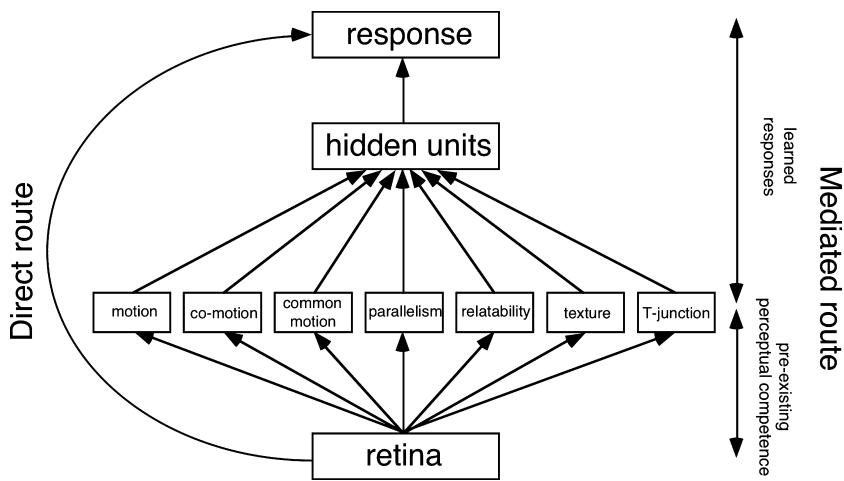


Fig. 8. Architecture of the Mareschal and Johnson (2002) model of perception of spatial completion. Adapted from Mareschal and Johnson (2002).

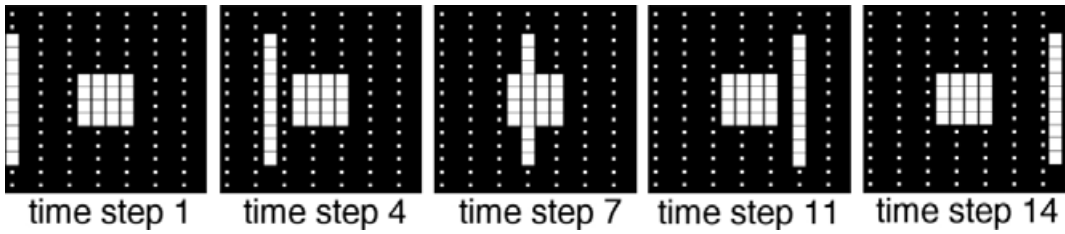


Fig. 9. Sample time steps illustrating the training regimen of the Mareschal and Johnson (2002) model.

be learned from proper experience in a structured environment, given appropriate lower-level perceptual skills—in this case, sensitivity to relevant visual information.

Models of fully hidden objects likewise have shown an important role for learning by association. Models described by Mareschal, Plunkett, and Harris (1999) were trained in a simple environment consisting of an occluder and a moving object that was small enough to be invisible for a number of time steps during an event in which it passed repeatedly back and forth along a horizontal trajectory. The models' task was to predict the location of this moving object in a future time step, and they learned to do this very quickly when it was fully visible. They were able to do so as well given repeated experience with a partly hidden trajectory—that is, a trajectory in which the object was briefly hidden. In other words, the models developed a representation of the moving object even in the absence of direct evidence for its existence—by virtue of a memory trace built from experience.

As noted previously, models are like experiments: A single modeling effort should not be taken to suggest that the specifics of the model's architecture or training provide any greater (or lesser) insights into human developmental processes than would a single set of experimental conditions tested in the lab. No doubt infants learn via association, but this is not all there is to how infants learn about the world at large. Other developmental phenomena are at work, and some of these phenomena have attracted the attention of modelers interested in exploring the origins of object knowledge. In the following section I will describe two recent models of visual development, each of which bears implications for infant object perception.

5.2. Modeling visual development

The scope and contributions of the models I have described are limited, in part because the human visual system does not work or develop in the same way: Our retinas have a fovea and we move our eyes to points of interest in the scene, and visual development in human infants consists of formation and strengthening of neural circuits within, to, and from visual areas of the brain (Atkinson, 2000), as opposed to updating of weights within fixed connections between hard-wired, fully operational modules characteristic of many models (Rumelhart, McClelland, and the PDP Research Group, 1986), including ours. With these caveats in mind, Schlesinger, Amso, and Johnson (2007) created a computational model of infants' gaze patterns based on the idea of "salience maps" produced by visual modules

tuned to luminance, motion, color, and orientation in an input image (Itti & Koch, 2000). Saliency was computed in part via a process of competition between visual features as the model received repeated exposures (or iterations) to the images, a strategy motivated by patterns of activity in the posterior parietal cortex that are suppressed in response to visual features that remain constant across exposure while increasing responses to features that change—thus highlighting their saliency (Gottlieb, Kusunoki, & Goldberg, 1998). The model had a simulated fovea and the ability to direct “gaze” toward the most salient region in the image. We input an image of a moving rod-and-box stimulus to the model. After several iterations, the model quickly developed a saliency map in which the rod segments were strongly activated, as activity for the edges of the occluder receded (Fig. 10). The model was intended to examine development of visual attention, not spatial completion per se, but given the success of this model and that of Mareschal and Johnson (2002), a model of “learning from scanning” seems feasible and likely to achieve important insights into human development.

The Schlesinger et al. (2007) model was designed to examine three kinds of cortical visual development in human infants, and their effects of these developments on scanning behavior and unity perception. The first was neural “noise,” uncorrelated activity among neurons within and across networks of cortical cells, which is characteristic of young infants’ cortical function (Skoczenski & Norcia, 1998), and might make pattern detection initially difficult across disparate regions of the visual scene. The second was *horizontal* connections in visual area V1, which serve to strengthen responses of neighboring cells that code for a common edge in the visual scene (Burkhalter, Bernardo, & Charles, 1993), and whose development is likely to facilitate perception of edge connectedness across a gap. The third was *recurrent processing* of visual information, akin to repetition of the input within working memory and accomplished in primate parietal cortex (Gottlieb et al., 1998), which is analogous to modulating the time spent covertly “comparing” two or more targets.

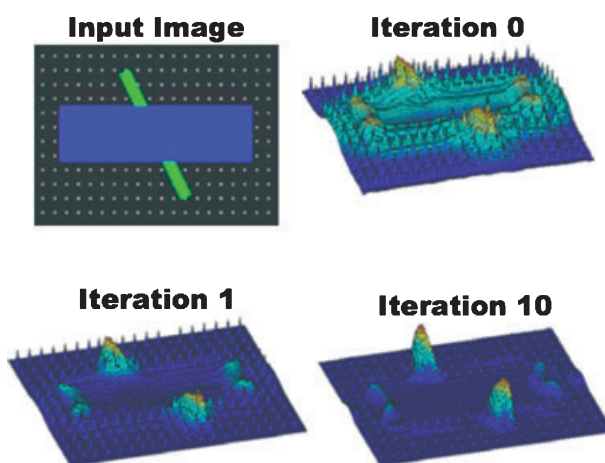


Fig. 10. Saliency map yielded by the model of visual development produced by Schlesinger et al. (2007). Adapted from Schlesinger et al. (2007).

The parameters representing neural noise had little effect on performance, but the interaction of developments in horizontal connections and recurrent processing had a substantial effect on unity perception, with an “ideal” value for the horizontal connections set to a fairly low number, and the addition of recurrent loops beneficial to the model’s ability to detect the unity of rod parts separated by the occluder. In other words, our model shows that growth of horizontal connections is neither necessary nor sufficient, whereas an increase in the duration of recurrent parietal activation is necessary and *almost* sufficient (i.e., it works for many values of horizontal connections).

5.3. Modeling vision and reaching

In humans, eyes are situated in a head, which is situated on a body, which can move through space and to which are attached arms and hands that serve to act on objects. Part of the exploration process involves moving to and around objects, bringing them closer and rotating them to produce new points of view for further visual inspection. Models of motor development have demonstrated that neural networks can learn to coordinate vision and reaching during bouts of exploration of the environment. Kuperstein (1988) described a model in which representations of posture and arm position emerged from correlations between self-produced movement and sensory information about external target positions in space. The model was endowed with two retinas located in a head on a trunk, and a hand on an arm, all of which were free to move in 3D space. The model was allowed to grasp an object as information about its position was input from both the motor system and the visual system. The developmental process was similar to the Piagetian notion of *circular reactions* (Piaget, 1954), in which a developing system’s behaviors are gradually honed after initial, sometimes lengthy, bouts of variability. The periods of variability were presumably used by the neural network to work out the coordination between sensorimotor feedback from different postures and positions of the limbs and the visual transformations they yielded. After correlations became more stable, the network learned to produce new, accurate patterns of reaching and grasping for objects that had not been encountered previously. In other words, the model had internalized the spatial mapping of limb position and visual coordinates—representations of the locations of the self and of external objects—from signals derived exclusively from sensory receptors and motor feedback, with no a priori knowledge of the objective “features” of objects.

Bullock, Grossberg, and Guenther (1993) introduced a model of eye-hand coordination with a similar goal: to examine emergence of correlations across sensory and motor systems without prespecified knowledge of how outputs from the two systems should be combined into a unitary representation. The phase of exploratory movements was termed *motor babbling* and constituted the principal learning period about the effects of movement on visual input. The goal of the model was to enact a reaching trajectory toward the object that was as direct as possible, on the basis of visual and sensory feedback. Following training, the model’s reaches for objects of different sizes and shapes were geared toward transforming visual information about the target and the effector (the hand) so as to produce maximally efficient, goal-oriented movements.

These models demonstrate that the visual input from objects is determined by the action capabilities of the system and the affordances of objects in the context of those actions. They illustrate in addition that cognitive development is constrained by expanding motor control over actions that provide increasingly detailed information about objects and events in the world.

6. How infants learn about objects

In the previous two sections, I described the capacities of newborns to detect and remember key information about the visual world, information that is important for specifying objects: their segregation from one another and their relative distances, and the retention of this information for brief intervals sufficient to support recognition upon repeated encounters. I also described models of development object perception. These models demonstrated that a naïve system, given appropriate perceptual, cognitive, and motor skills and a suitable environment in which to learn, can perceive objects as complete and persistent despite occlusion, and can act on objects by detecting relevant information about their properties. Does the developmental process in human infants accord with these findings?

6.1. Infants learn about objects via association

Consider first the possibility that infants learn about object occlusion via association. How might this work? The Mareschal and Johnson (2002) model learned to perceive partly hidden objects as complete in two ways: by associating objects with different visual cues in the input (i.e., texture, motion, junctions, and orientation), and by associating different views of objects with each other—a fully visible, complete rod that moved behind the occluder and then became partly occluded. The Mareschal et al. (1999) model was exposed to an object moving on a repetitive trajectory and quickly learned to predict its reappearance from behind an occluder. The model was set up to predict the location of the moving object based on its preceding position and trajectory, and the model maintained a memory trace of it when it was rendered invisible by the occluder. Is there evidence for similar processes in human infants?

To my knowledge, no one has tested the possibility that infants learn about objects by *associating* individual visual attributes with their coherence and persistence across occlusion, though the contributions of such visual attributes to perceptual completion have been investigated. Spatial completion has been observed in young infants (younger than 6 months) only when the rod parts are aligned, and moving in tandem behind the occluder (Johnson, 1997, 2004), in displays with the four visual cues examined in the Mareschal and Johnson (2002), but in the absence of one or more cues, spatial completion is disrupted (Kellman & Spelke, 1983). And spatiotemporal completion has been observed in young infants only when the trajectory is horizontal, not angled, and when the spatiotemporal gap imposed by the occluder is relatively short (Bremner et al., 2005, 2007; Johnson, Bremner, et al. 2003). But it is not clear that these studies provide evidence that association per se is an important mechanism of development in perceiving object occlusion.

Such evidence comes from experiments by Johnson, Amso, and Slemmer (2003), who examined 4- and 6-month-old infants' responses to object trajectory displays by recording predictive eye movements. We reasoned that a representation of a moving object would be revealed by a consistent pattern of fixations toward the far side of the occluder upon its occlusion. Infants were tested in one of four conditions. In the *baseline* condition, infants were shown the ball-box display depicted in Fig. 4 as eye movements were recorded with a corneal-reflection eye tracker. The display was presented for eight 30-s trials. In the *random* condition, infants viewed eight presentations of displays that were identical to the ball-box stimulus except the ball's point of reemergence after occlusion was randomized (left or right). In this case, anticipation offers no gain to the observer, who is just as likely to make perceptual contact with the ball if the point of gaze remains where the object moved out of view. (We hypothesized that anticipations in the random condition might be random eye movements themselves.) In the *training* condition, infants were first presented with four trials of the ball only, fully visible on its lateral trajectory (no occluder), followed by four trials with the ball-box display, as in the predictable condition. Finally, in the *generalization* condition, infants first viewed four trials with a vertical unoccluded trajectory, followed by four trials with a partly occluded horizontal trajectory.

In the baseline condition, 6-month-olds produced a significantly higher proportion of anticipatory eye movements than 4-month-olds, and a comparison of 4-month-olds' performance in the baseline versus random conditions revealed no reliable differences. This latter finding implies that any predictive eye movements we observed by 4-month-olds were actually not based on a mental representation of the occluded object and its motion, but instead were simply random eye movements scattered about the display that, by chance, happened to fit the criteria for categorization as predictive. Moreover, 4-month-olds' performance in the baseline condition did not improve across trials (as would be expected if the infants learned the repetitive sequence). In fact, there was a significant *decline* in anticipations across trials. These results indicate that eye movement patterns may have been driven more in the older age group by a veridical representation of the object on its path behind the occluder.

However, 4-month-olds in the training condition showed reliably more predictive eye movements relative to 4-month-olds in the baseline condition. Comparisons of the two 6-month-old groups, in contrast, revealed no significant differences. The boost in anticipation performance seen in the 4-month-old training group generalized from exposure to the vertical trajectory orientation, implying that infants in the training condition were not simply trained for facilitation of horizontal eye movements, but instead true representation-based anticipations.

How long does this effect of training last? Johnson and Shuwairi (2009) addressed this question with a replication of the Johnson, Amso, et al. (2003) experiment: Baseline and training conditions with 4-month-olds yielded similar results as the previous study. We extended these findings with three additional conditions: a *delay* condition, in which a 30-min wait was imposed between training (with an unoccluded trajectory) and test (with a partly occluded trajectory), and a *reminder* condition, identical to the delay condition except for the addition of a single additional training trial immediately before test. Performance in the delay condition was not significantly different from that of baseline, implying that the

gains produced by brief training did not survive the 30-min interruption prior to test. However, eye movement anticipations were facilitated by the reminder condition to the same extent as in the (immediate) training condition. (A fifth condition, *brief training*, consisted of a single training trial prior to immediate test, and this did not have any discernible effect on performance.)

Taken together, these findings suggest that there are consequential changes around 4 months after birth in representations of moving, occluded objects (Johnson, Bremner, et al. 2003). Such representations are sufficiently strong by 6 months to guide anticipatory looking behaviors consistently when viewing predictable moving object event sequences. Four-month-olds' anticipations under these conditions provided little evidence of veridical object representations. However, a short exposure to an unoccluded object trajectory induces markedly superior performance in our tracking task in this age group, and with a reminder, this training effect can last for a period of time outside the scope of short-term memory. These findings also help clarify the role of associative learning in object perception development. Infants did not seem to learn by viewing repetitive events that are perfectly predictable to adults (otherwise infants in the baseline conditions would have begun to show increased levels of anticipation after several trials viewing the occluded trajectory). Instead, infants learned by associating views of the fully visible object trajectory and the partly occluded object trajectory.

6.2. Infants learn about objects via “active assembly”

The Schlesinger et al. (2007) model discussed in the previous section highlighted two aspects of visual development that might have a key role in development of spatial completion: growth of horizontal connections among neurons and circuits in cortical visual area V1, and recurrent processing, which, we reasoned, served to compare aspects of the visual scene. How might these influence developing object perception skills?

Burkhalter and colleagues (Burkhalter, 1991; Burkhalter et al., 1993) have reported evidence for developments in horizontal connections in V1 from deceased fetuses and infants, across the period from 26 weeks postconception to 7 months after birth, but the precise role of these developments in object perception has not been documented. However, there are findings from experiments on spatial completion that bear on this question. Two-month-old infants have been found to perceive spatial completion in displays with a relatively narrow occluder, such that the rod parts are close together across the gap imposed by occlusion, but not when the occluder is wide (Johnson, 2004). (Similar findings were obtained in studies of spatiotemporal completion—reducing gap size facilitates perception of completion here as well.) Older infants are less susceptible to effects of widening this gap (Johnson, 1997). In addition, 4-month-old infants are more likely to look back and forth across a wide gap at the rod parts than are 2-month-olds (Johnson & Johnson, 2000). These results are to be expected if the visual system becomes better able to link aligned edges across a gap as connections between receptive fields become strengthened.

Other experiments examined the possibility that spatial completion develops from a constructive process—which I have termed *active assembly*—serving to integrate parts of

the visual scene into a coherent whole, in like fashion to recurrent processing discussed previously. Amso and Johnson (2006) and Johnson, Slemmer, and Amso (2004) observed 3-month-old infants in a spatial completion task and recorded infants' eye movements with a corneal reflection eye tracker during the habituation phase of the experiment. We found systematic differences in oculomotor scanning patterns between infants whose posthabituation test display preferences indicated unity perception and infants who provided evidence of perception of disjoint surfaces: "Perceivers" tended to scan more in the vicinity of the two visible rod segments, and to scan back and forth between them (Fig. 11). In a younger sample, Johnson et al. (2008) found a correlation between posthabituation preference—our index of spatial completion—and targeted visual exploration, operationalized as the proportion of eye movements directed toward the moving rod parts, which we reasoned was the most relevant aspect of the stimulus for perception of completion. (Precise localization of the point of gaze can be a challenge for these very young infants, attested by the fact that targeted scans almost always followed the rod as it moved, rarely anticipating its position.)

A relation between targeted visual exploration and spatial completion does not by itself pinpoint a causal role. Such evidence would come from experiments in which individual differences in oculomotor patterns were observed in both spatial completion and some other visual task, and this was recently reported by Amso and Johnson (2006). We found that both spatial completion and scanning patterns were strongly related to performance in an independent visual search task in which targets, defined by a unique feature (either motion or orientation) were placed among a large set of distracters. There were substantial individual differences in successful search, both in terms of detecting the target and the latency to do so, and these differences mapped clearly onto the likelihood of spatial completion. This finding is inconsistent with the possibility that scanning patterns were tailored specifically to perceptual completion, and instead suggests that a general facility with targeted visual behavior leads to improvements across multiple tasks.

Targeted visual exploration may make a vital contribution to the emergence of veridical object perception. As scanning patterns develop, they support binding of disparate visual features into unified percepts—active assembly of coherent objects from surface fragments, confirming the outcome of the Schlesinger et al. (2007) model of visual development. With the emergence of selective attention and other perception-action systems, infants become

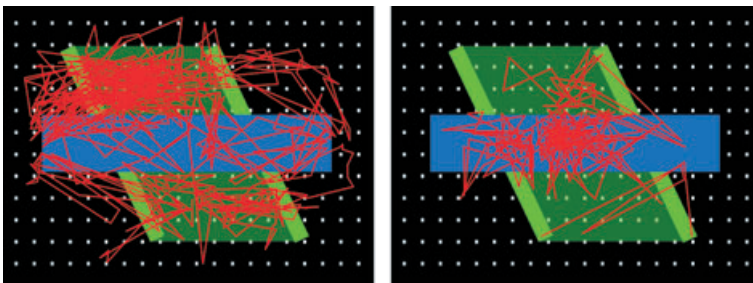


Fig. 11. Examples of individual differences in oculomotor scanning as 3-month-old infants view a rod-and-box display (Johnson et al., 2004).

increasingly active participants in their own perceptual development rather than passive recipients of information. Active engagement of the infant's visual attention is consistent with a key tenet of Piagetian theory—the central role of the child's own behavior in cognitive development—and with a constructivist view—the building of structure from constituent pieces. The following section describes another of these perception-action systems, visual-manual exploration, and its role in constructing volumetric objects.

6.3. *Infants learn about objects via visual-manual exploration*

The Kuperstein (1988) and Bullock et al. (1993) models demonstrated that when perception and action develop in tandem their coordination can be an emergent property, with each influencing the other to the benefit of the exploratory capacity of the organism. Developments in perception and action have long been of interest to developmental psychologists, and there has been recent evidence to show that 3D object completion emerges as a consequence of improvements in infants' motor skills. Two types of motor skill, both of which develop rapidly at the same time that 3D object completion seems to emerge—4 to 6 months—may play a particularly important role: self-sitting and coordinated visual-manual exploration. Independent sitting frees the hands for play and promotes gaze stabilization during manual actions (Rochat & Goubet, 1995), and, therefore, self-sitting might encourage coordination of object manipulation with visual inspection as infants begin to play with objects, providing the infants with multiple views. In addition, manipulation of objects—touching, squeezing, mouthing—may promote learning about object form from tactile information.

To examine these possibilities, Soska, Adolph, and Johnson (2010) observed infants between 4.5 months and 7.5 months in a replication of the Soska and Johnson (2008) habituation experiment with the rotating wedge stimuli (Fig. 6). In the same testing session we assessed infants' manual exploration skills by observing spontaneous object manipulation in a controlled setting and obtained parental reports of the duration of infants' sitting experience. We reasoned that infants who had more self-sitting experience would in turn show a greater tendency to explore objects from multiple viewpoints and therefore have more opportunities to learn about objects' 3D forms outside the lab. Thus, within this age range, individual differences in self-sitting experience and coordinated visual-manual exploration were predicted to be related to individual differences in infants' looking preferences to the complete and incomplete object displays, our index of 3D object completion.

Our predictions were supported. We found strong and significant relations between both self-sitting and visual-manual coordination, from parents' reports and the motor skills assessment, and 3D object completion performance, assessed with the habituation paradigm. We recorded a number of other motor skills to explore how widespread the relations were within the perception-action systems under investigation, such as grasping, holding, and manipulation without visual inspection, and none were related to 3D object completion.

Self-sitting experience and coordinated visual-manual exploration were the strongest predictors of performance on the visual habituation task. The results of a regression analysis

yielded evidence that the role of self-sitting was indirect, influencing 3D completion chiefly in its support of infants' visual-manual exploration. Self-sitting infants performed more manual exploration while looking at objects than did nonsitters, and visual-manual object exploration is precisely the skill that provides active experience viewing objects from multiple viewpoints, thereby facilitating perceptual completion of 3D form. These results provide evidence for a cascade of developmental events following from the advent of visual-motor coordination, including learning from self-produced experiences.

7. Concluding remarks

In principle, perceptual completion and other object perception skills available early in postnatal life might develop solely from "passive" perceptual experience, because natural scenes are richly structured and characterized by a high degree of redundancy (Graham & Field, 2007) and infants gain a great deal of exposure to the visual world—on the order of several hundred hours—by 2 months after birth (Johnson, Amso, et al. 2003). Thus—in principle—infants might learn about objects by observing the world and acquiring associations between views of objects when fully visible and partly or fully occluded. But the findings yielded by the Amso and Johnson (2006) and Soska, Adolph, and Johnson (2010) experiments indicate that passive experience may be insufficient to learn about the full range of occlusion phenomena, and, together with modeling accounts of object perception development, broaden our conceptions of how infants learn about the visual world. Active assembly and visual-manual exploration provide information to the infant about her own control of an event while simultaneously generating multimodal information to inform developing object perception skills. For complex kinds of perceptual completion, such as 3D object completion, the coordination of posture, reaching, grasping, and visual inspection seems to be critical: Only the visual-manual skills involved in generating changes in object viewpoint—rotating, fingering, and transferring while looking—were related to 3D object completion.

Careful consideration of the evidence I have described reveals that no one account, such as Piagetian or nativist theories, encompasses the full range of changes that underlie the emergence of object concepts in infancy. Significant advances, nevertheless, have been achieved. The rudiments of object knowledge are evident in the first 6 months after birth, revealed by detailed observations of information-gathering processes available to infants, from which more complex representations of objects are constructed. But, notably, there is no pure case of development caused in the absence of either intrinsic or external influences (Elman et al., 1996; Quartz & Sejnowski, 1997). The question is what mechanisms are responsible for perceptual and cognitive development.

Acknowledgments

Preparation of this article was supported by NIH grants R01-HD40432 and R01-HD48733.

References

- Aguiar, A., & Baillargeon, R. (1999). 2.5-month-old infants' reasoning about when objects should and should not be occluded. *Cognitive Psychology*, 39, 116–157.
- Albert, M. V., Schnabel, A., & Field, D. J. (2008). Innate visual learning through spontaneous activity patterns. *PLoS Computational Biology*, 4, 1–8.
- Amso, D., & Johnson, S. P. (2006). Learning by selection: Visual search and object perception in young infants. *Developmental Psychology*, 6, 1236–1245.
- Atkinson, J. (2000). *The developing visual brain*. New York: Oxford University Press.
- Baillargeon, R. (1994). How do infants learn about the physical world? *Current Directions in Psychological Science*, 3, 133–140.
- Baillargeon, R. (2008). Innate ideas revisited: For a principle of persistence in infants' physical reasoning. *Perspectives on Psychological Science*, 3, 2–13.
- Bednar, J. A., & Mäikkiläinen, R. (2007). Constructing visual function through prenatal and postnatal learning. In D. Mareschal, S. Sirois, G. Westermann, & M. H. Johnson (Eds.), *Neuroconstructivism: Perspectives and prospects*, Vol. 2. (pp. 13–37). New York: Oxford University Press.
- Berger, A., Tzur, G., & Posner, M. L. (2006). Infant brains detect arithmetic errors. *Proceedings of the National Academy of Sciences (USA)*, 103, 12649–12653.
- Bremner, J. G., Johnson, S. P., Slater, A., Mason, U., Cheshire, A., & Spring, J. (2007). Conditions for young infants' failure to perceive trajectory continuity. *Developmental Science*, 10, 613–624.
- Bremner, J. G., Johnson, S. P., Slater, A. M., Mason, U., Foster, K., Cheshire, A., & Spring, J. (2005). Conditions for young infants' perception of object trajectories. *Child Development*, 74, 1029–1043.
- Bullock, D., Grossberg, S., & Guenther, F. H. (1993). A self-organizing neural model of motor equivalent reaching and tool use by a multijoint arm. *Journal of Cognitive Neuroscience*, 5, 408–435.
- Burkhalter, A. (1991). Developmental status of intrinsic connections in visual cortex of newborn human. In P. Bagnoli & W. Hodos (Eds.), *The changing visual system* (pp. 247–254). New York: Plenum Press.
- Burkhalter, A., Bernardo, K. L., & Charles, V. (1993). Development of local circuits in human visual cortex. *Journal of Neuroscience*, 13, 1916–1931.
- Carey, S. (2009). *The origin of concepts*. New York: Oxford University Press.
- Clifton, R. K., Rochat, P., Litovsky, R. Y., & Perris, E. E. (1991). Object representation guides infants' reaching in the dark. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 323–329.
- Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.
- Gibson, J. J. (1950). *The perception of the visual world*. Boston: Houghton-Mifflin.
- Goldin-Meadow, S., & Mylander, C. (1998). Spontaneous sign systems created by deaf children in two cultures. *Nature*, 391, 279–281.
- Goldman-Rakic, P. S. (1987). Development of cortical circuitry and cognitive function. *Child Development*, 58, 601–622.
- Goldman-Rakic, P. S. (1996). The prefrontal landscape: Implications of functional architecture for understanding human mentation and the central executive. *Philosophical Transactions of the Royal Society of London B*, 351, 1445–1453.
- Gottlieb, J. P., Kusunoki, M., & Goldberg, M. E. (1998). The representation of visual salience in monkey parietal cortex. *Nature*, 391, 481–484.
- Graham, D. J., & Field, D. J. (2007). Statistical regularities of art images and natural scenes: Spectra, sparseness and nonlinearities. *Spatial Vision*, 21, 149–164.
- Haith, M. M. (1998). Who put the cog in infant cognition? Is rich interpretation too costly? *Infant Behavior & Development*, 21, 167–179.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40, 1489–1506.

- Johnson, S. P. (1997). Young infants' perception of object unity: Implications for development of attentional and cognitive skills. *Current Directions in Psychological Science*, 6, 5–11.
- Johnson, S. P. (2004). Development of perceptual completion in infancy. *Psychological Science*, 15, 769–775.
- Johnson, S. P. (Ed.) (2010). *Neoconstructivism: The new science of cognitive development*. New York: Oxford University Press.
- Johnson, S. P., Amso, D., & Slemmer, J. A. (2003). Development of object concepts in infancy: Evidence for early learning in an eye tracking paradigm. *Proceedings of the National Academy of Sciences (USA)*, 100, 10568–10573.
- Johnson, S. P., & Aslin, R. N. (1995). Perception of object unity in 2-month-old infants. *Developmental Psychology*, 31, 739–745.
- Johnson, S. P., & Aslin, R. N. (1996). Perception of object unity in young infants: The roles of motion, depth, and orientation. *Cognitive Development*, 11, 161–180.
- Johnson, S. P., Bremner, J. G., Slater, A., Mason, U., Foster, K., & Cheshire, A. (2003). Infants' perception of object trajectories. *Child Development*, 74, 94–108.
- Johnson, S. P., Davidow, J., Hall-Haro, C., & Frank, M. C. (2008). Development of perceptual completion originates in information acquisition. *Developmental Psychology*, 44, 1214–1224.
- Johnson, S. P., & Johnson, K. L. (2000). Early perception-action coupling: Eye movements and the development of object perception. *Infant Behavior & Development*, 23, 461–483.
- Johnson, S. P., & Nájuez, J. E. (1995). Young infants' perception of object unity in two-dimensional displays. *Infant Behavior & Development*, 18, 133–143.
- Johnson, S. P., & Shuwairi, S. M. (2009). Learning and memory facilitate predictive tracking in 4-month-olds. *Journal of Experimental Child Psychology*, 102, 122–130.
- Johnson, S. P., Slemmer, J. A., & Amso, D. (2004). Where infants look determines how they see: Eye movements and object perception performance in 3-month-olds. *Infancy*, 6, 185–201.
- Kagan, J. (2008). In defense of qualitative changes in development. *Child Development*, 79, 1606–1624.
- Kellman, P. J., & Arterberry, M. E. (1998). *The cradle of knowledge: Development of perception in infancy*. Cambridge, MA: MIT Press.
- Kellman, P. J., & Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology*, 15, 483–524.
- Kessen, W., Salapatek, P., & Haith, M. (1972). The visual response of the human newborn to linear contour. *Journal of Experimental Child Psychology*, 13, 9–20.
- Kuperstein, M. (1988). Neural model of adaptive hand-eye coordination for single postures. *Science*, 239, 1308–1311.
- Mareschal, D., & Bremner, A. J. (2009). Modeling the origins of object knowledge. In B. Hood & L. Santos (Eds.), *The origins of object knowledge* (pp. 227–262). Oxford, England: Oxford University Press.
- Mareschal, D., & Johnson, S. P. (2002). Learning to perceive object unity: A connectionist account. *Developmental Science*, 5, 151–185.
- Mareschal, D., Plunkett, K., & Harris, P. (1999). A computational and neuropsychological account of object-oriented behaviors in infancy. *Developmental Science*, 2, 306–317.
- Moore, D. S., & Johnson, S. P. (2008). Mental rotation in human infants: A sex difference. *Psychological Science*, 19, 1063–1066.
- Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.
- Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books.
- Quartz, S. R., & Sejnowski, T. J. (1997). The neural basis of cognitive development: A constructivist manifesto. *Behavioral and Brain Sciences*, 20, 537–556.
- Quinn, P. C., & Liben, L. S. (2008). A sex difference in mental rotation in young infants. *Psychological Science*, 19, 1067–1070.
- Regolin, L., & Vallortigara, G. (1995). Perception of partly occluded objects by young chicks. *Perception & Psychophysics*, 57, 971–976.

- Rochat, P., & Goubet, N. (1995). Development of sitting and reaching in 5- to 6-month-old infants. *Infant Behavior & Development*, 18, 53–68.
- Rumelhart, D. E., McClelland, J. L., & the PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press.
- Schlesinger, M., Amso, D., & Johnson, S. P. (2007). The neural basis for visual selective attention in young infants: A computational account. *Adaptive Behavior*, 15, 135–148.
- Sekuler, A. B., Lee, J. A. J., & Shettleworth, S. J. (1996). Pigeons do not complete partly occluded figures. *Perception*, 25, 1109–1120.
- Senghas, A., Kita, S., & Özyürek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science*, 305, 1779–1782.
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, 701–703.
- Skoczenski, A. M., & Norcia, A. M. (1998). Neural noise limitations on infant visual sensitivity. *Nature*, 391, 697–700.
- Slater, A. (1995). Visual perception and memory at birth. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research*, Vol. 9. (pp. 107–162). Norwood, NJ: Ablex.
- Slater, A., Johnson, S. P., Brown, E., & Badenoch, M. (1996). Newborn infants' perception of partly occluded objects. *Infant Behavior & Development*, 19, 145–148.
- Slater, A., Johnson, S. P., Kellman, P. J., & Spelke, E. S. (1994). The role of three-dimensional depth cues in infants' perception of partly occluded objects. *Early Development and Parenting*, 3, 187–191.
- Slater, A., Morison, V., Somers, M., Mattock, A., Brown, E., & Taylor, D. (1990). Newborn and older infants' perception of partly occluded objects. *Infant Behavior and Development*, 13, 33–49.
- Slater, A., Morison, V., Town, C., & Rose, D. (1985). Movement perception and identity constancy in the new-born baby. *British Journal of Developmental Psychology*, 3, 211–220.
- Slater, A., Quinn, P. C., Kelly, D. J., Lee, K., Longmore, C. A., McDonald, P. R., & Pascalis, O. (in press). The shaping of the face space in early infancy: Becoming a native face processor. *Child Development Perspectives*.
- Soska, K. C., Adolph, K. A., & Johnson, S. P. (2010). Systems in development: Motor skill acquisition facilitates 3D object completion. *Developmental Psychology*, 46, 129–138.
- Soska, K. C., & Johnson, S. P. (2008). Development of 3D object completion in infancy. *Child Development*, 79, 1230–1236.
- Spelke, E. S. (1990). Principles of object perception. *Cognitive Science*, 14, 29–56.
- Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, 99, 605–632.
- Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, 10, 89–96.
- Spencer, J. P., Blumberg, M. S., McMurray, B., Robinson, S. R., Samuelson, L. K., & Tomblin, J. B. (2009). Short arms and talking eggs: Why we should no longer abide the nativist-empiricist debate. *Child Development Perspectives*, 3, 79–87.
- Sperry, R. W. (1963). Chemoaffinity in the orderly growth of nerve fiber patterns and their connections. *Proceedings of the National Academy of Sciences (USA)*, 50, 703–710.
- Valenza, E., Leo, I., Gava, L., & Simion, F. (2006). Perceptual completion in newborn human infants. *Child Development*, 77, 1810–1821.
- Valenza, E., Simion, F., Macchi Cassia, V., & Umiltà, C. (1996). Face preference at birth. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 892–903.
- Wong, R. O. L. (1999). Retinal waves and visual system development. *Annual Review of Neuroscience*, 22, 29–47.
- Zeki, S. (1993). *A vision of the brain*. Cambridge, MA: Blackwell.