



Published in final edited form as:

*Am Psychol.* 2013 November ; 68(8): . doi:10.1037/a0034185.

## It's all connected: Pathways in visual object recognition and early noun learning

**Linda B. Smith**

Department of Psychological and Brain Sciences, Indiana University, Bloomington IN 47405

### Abstract

A developmental pathway may be defined as the route, or chain of events, through which a new structure or function forms. For many human behaviors, including object name learning and visual object recognition, these pathways are often complex, multi-causal and include unexpected dependencies. This paper presents three principles of development that suggest the value of a developmental psychology that explicitly seeks to trace these pathways and uses empirical evidence on developmental dependencies between motor development, action on objects, visual object recognition and object name learning in 12 to 24 month old infants to make the case. The paper concludes with a consideration of the theoretical implications of this approach.

### Keywords

Visual object recognition; noun learning; cognitive development; developmental theory

---

Development is activity-dependent change in a complex system. A tangle of successive causes and effects accumulate change over time and increase the structure and complexity of the developing system (e.g., Gottlieb, 1991; Lickliter & Honeycutt, 2003; Thelen & Smith, 1994). Research at many levels of analysis tells us that within this complex system, developmental pathways to specific outcomes are complex in two ways: they are multi-causal, each change dependent on multiple causes, and they are often degenerate, there is more than one route to the same functional end (Edelman & Gally, 2001; Whitacre, 2010). Degeneracy is believed to promote robustness in developmental outcomes. Because functionally redundant pathways can compensate for one another, they provide a kind of insurance against pathway failure.

This paper illustrates these ideas by considering the links between sensory-motor development, visual object recognition and word learning. The research connecting developments in these domains reveal the complicated cascade that is developmental process, offer useful ideas for exploiting these pathways in formulating effective interventions, and serve as a jumping off point for broader implications of this approach for how we do developmental science.

### Three principles

#### Principle 1: The past is prelude

Development, like evolution and culture, is a process that creates complexity by accumulating change. At any moment, the whole child is a product of all the previous

developments, and any new change begins with and must build on those previous developments. Theorists often refer to the far reach of early developments on later ones in terms of the “developmental cascade,” and do so most often when talking about atypical developmental process, about how, for example, motor deficits and limits on children’s ability to self-locomote cascade into the poor development of social skills (Galloway, Ryu & Argawal, 2008) or how disrupted sleep patterns in toddlers start a pathway to poor self regulation and conduct disorder (Bates, Viken, Alexander, Beyers & Stockton, 2002). But the cascade characterizes all aspects of typical and atypical development. One example pertinent to the present paper concerns the relation between sitting and the development of more view-independent visual representations of object shape. Given a view of just one side of a never-before-seen object, as in the top drawing of Figure 1, adults have strong expectations about the geometric structure of the whole (Tse, 1999). For example, adults expect a rotation of that object to yield (b) and not a shell (c). These expectations imply internal representations of three-dimensional objects and not 2-dimensional views. Using a preferential looking paradigm, Soska, Adolph, and Johnson (2010) tested 5- to 8-month old infants’ expectations about the unseen sides of simple drawings of objects. They found that infants’ ability to sit steadily, not age, was the best predictor of their expectations about the unseen sides of objects. They reasoned, and considerable data supports their reasoning (e.g., Pereira, James, Jones & Smith, 2010; Ruff, 1982), that babies that can sit steadily, they can hold and manipulate objects for sustained periods (without falling over) and that in so doing generate dynamic visual experiences of individual objects that in turn build up generalized expectations about the three-dimensional structure of visual things. In brief, sitting steadily and manipulating objects is part of the developmental pathway that leads to the processes of visual object recognition that characterize mature vision.

Clearly, sitting steadily is unlikely to be necessary or sufficient to these developments; it is easy to think of ways to generate these experiences without sitting or ways to support trunk control to foster stable sitting and thereby sustained manual play with objects. In the typical developmental pathway, sitting sets the stage for activity generated-experiences crucial to the visual object recognition system. The theoretical and practical relevance of understanding how development builds on itself is not diminished because there are multiple routes; instead, the complexity and degeneracy of developmental pathways is one reason that a pathways approach to developmental theory provides insight.

## **Principle 2: Overlapping tasks**

Developing organisms do not solve just one task; they solve many overlapping tasks (Thelen & Smith, 1994). Consider Piaget’s (1952) description of a secondary circular reaction: A rattle is placed in a 4 month old infant’s hands. The infant moves the rattle and so it comes into and out of sight and makes a noise. Piaget noted that this aroused and agitated the infant, causing more body motions, and thus causing the rattle to move more rapidly into and out of sight and to make more noise. Young infants have little organized control over hand and eye; yet over just minutes of interacting with the rattle, their activity becomes highly organized and goal-directed. Piaget believed this pattern of activity, involving multimodal perception–actions loops, held the key to understanding the origins of human intelligence.

Contemporary theorizing in computational neuroscience sees the importance of multiple modalities, heterogeneous subsystems, and their coordination in specific tasks in much the same way that Piaget did: functional systems of different neural components assembled in the service of specific physical tasks and time-locked to the same physical events drive neural change and build functional networks (McIntosh, Fitzpatrick, & Friston, 2001; Metta & Fitzpatrick, 2003; Sporns, 2011. Figure 2a illustrates these ideas from computational theory using Piaget’s example of a baby shaking a rattle. The figure shows three systems—

motor, vision, and audition—receiving qualitatively different sensory inputs from the very same physical event, a physical event driven by the motor system.. The qualitatively different patterns of activation in each system have their own dynamics but these internal dynamics are also time-locked to each other and to the activity in the physical world. Thus the activation patterns in each system are correlated in time. Computational theories suggest that these mutual dependencies among components in a system actively engaged with the physical world builds the flexible functional networks and higher order knowledge that comprise human intelligence (Lungarella, Pegors, Bulwinkle & Sporns, 2005; Lungarella & Sporns, 2006).

The human neural system is far more complex than the model system shown in Figure 2a. Each system is, itself, composed of many interconnected subsystems, each with their own sensitivities, properties, and intrinsic dynamics. Different subsets of components from this larger system will be recruited and will form different functional networks in different tasks, say in face-to-face play versus crawling versus object play. By several accounts, these overlapping coordinations are the engine of cognitive development (Barsalou, Simmons, Barbey, & Wilson, 2003; Edelman, 1987; Smith & Breazeal, 2007; Thelen & Smith, 1994). The theoretical idea is illustrated in Figure 2b: Systems A and B are coordinated in Task 1, creating change in both component systems and in their connections. Systems B and C are coordinated in the service of some other, second task. The key point is that the changes in System B wrought via coordination with System A in Task 1 will influence learning and performance in Task 2, constraining solutions to that task. This is a toy example, as children’s cognitive systems are not made from three systems and two tasks but from many systems and subsystems in many interleaved, variable, and repeated tasks. These overlapping co-ordinations –where changes wrought in one task are brought forward and may influence learning and adaptation in a very different task – will give rise to the cascading interactions characteristic of human development, wherein even seemingly far achievements may be developmentally related (Sheya and Smith, 2009). A pathways approach to developmental theory offers a framework in which to document and detail the mechanisms underlying both near and far developmental dependencies.

### Principle 3: Ordered developments

Biologically developing systems typically confront classes of experiences and tasks in a particular sequence. Research on the development of biological intelligence strongly suggests that a key ingredient of developmental process (see Turkewitz & Kenny, 1985) is a constrained ordering of experiences that is determined by development. One area in which this is seen is in the developmental ordering of the relative maturity of sensory-motor systems which is markedly different in different species –kittens, for example, hear, and walk and smell at birth but cannot see; humans, see and hear reasonably well, but are motorically very immature. There is a large experimental literature on the cascading developmental consequences of altering that natural order of sensory motor development in animals and several analyses about the origins of differences between near species in these terms (see, Winkowski & Knudsen, 2006; Turkewitz & Kenny, 1985; Lord, 2013). West and King (1987) extended these ideas with the broader proposal of “ontogenetic niches:” the environments in which development takes place change systematically with development itself and thus the developmental timing of environments may be exploited by evolutionary processes to ensure adaptive outcomes. For example, for many mammals and birds, early life is highly dependent on caretakers, a fact that tightly constrains early conspecific experiences. These constrained early experiences, in turn, have been hypothesized to canalize species-typical development (Gottlieb, 1991). For example, in the human context, early visual experiences might be expected to include relatively many faces –since young infants need near constant and close care from their parents, and these early face experiences

might be expected to engage, train, and tune specialized visual face processing. (e.g., Nelson, 2001). From a somewhat different perspective, cognitive theorists offered the “starting small” hypothesis: limits that arise from the immaturity of the neural system constrain the input and, rather than holding back development, play a role in fostering it (Elman, 1993; Fox, Levitt & Nelson, 2010; Newport, 1990).

Between birth and 2 years, human infants travel through a set of highly distinct developmental environments determined first by their early immaturity and then by their growing competencies. For example, each new motor skill achieved in the first two years - reaching, sitting, crawling, walking, - opens opportunities for new activities that yield new perceptual, cognitive and social experiences. A pathway approach to developmental theory provides a way of mapping and then understanding the ordered set of developmental tasks and experiences that build human intelligence.

## Pathways in visual object recognition and early noun learning

My colleagues and I have been working on understanding the behavioral pathways relevant to learning object names and to visual object recognition. Figure 3 provides an overview of a chain of successive developments that we have uncovered.

### Path 1: Hands, eyes, and visual object recognition

A fundamental problem in visual object recognition is how the snapshot 2-dimensional views that are the input are integrated to form expectations about and/or representations of 3-dimensional object shape. Several recent theoretical proposals posit that 3-dimensional views may be built from the dynamic experience of objects as they are rotated around the elongated axis (Farivar, 2009; Graf, 2006). This is why sitting steadily is part of the developmental pathway for visual object recognition. Once babies can hold and manipulate objects, they can show themselves dynamically organized views of 3-dimensional things, building up the principles of 3-dimensional shape and how to predict 3-dimensional shape from 2-dimensional projections (Graf, 2006; Farivar, 2009; Pereira et al, 2010; James, Swain, Jones & Smith, 2013).

When adults self-generate the views of 3-dimensional objects they see, either by holding and rotating the objects (Pereira et al., 2010) or by controlling the object views in virtual reality or through other means (James et al, 2002; Perrett, Harries & Looker, 1992), they systematically show themselves so-called planar views. These are object views in which the major axis of elongation is parallel or perpendicular to the line of sight as shown in Figure 4a. Pereira et al (2010) gave 12- to 36-month old children objects to hold and visually explore while recording the first-person views via a head camera. As children looked at, played with, and rotated the objects for viewing, they systematically generated many more planar views than would be expected by chance. This planar bias was found to be reliable even in the youngest children and the bias strengthened markedly between 18 and 24 months. These older children’s views were specifically biased to planar views that were elongated and to nonplanar views that were the rotations –around the most elongated axis – from one planar view to the next, dynamic views that have been proposed to provide the best support for forming 3-dimensional object representations (Cutzu & Tarr, 2007; Graf, 2006; see also Pereira et al., 2010). Subsequent research has shown that self-generated views that favor the planar sides predict better subsequent of recognition of novel objects by young children (James, et al, 2013) and that the bias obtains in 2 year olds even when strongly challenged by objects that are most easily held in ways that yield nonplanar views (James, Jones, Swain, Pereira & Smith, under revision). In brief, how children hold objects and the views they generate from holding those objects may be critical to the specific visual mechanisms that yield object-centered representations of 3-dimensional shape.

One class of theories about how the human visual system represents 3-dimensional object shape (Marr & Nishihara, 1978; Biederman, 1987) proposes that objects are represented in terms of their major parts and their relational organization of those major parts with respect to the major axis of elongation. Consistent with these ideas, adults readily recognize common objects from a few geometric components in their proper relation (Biederman, 1987). Recent studies indicate that the ability to recognize well-known objects – a chair, a dog – from similarly sparse information about object shape first emerges between the ages of 18 and 24 months (see Smith, 2009). For example, using both a name comprehension task and an action task, Smith (2003) examined 18- and 24-month-old children's ability to recognize 3-dimensional objects given only their sparse part structure or given richly detailed instances like those shown in Figure 4c. Older children recognized the sparse part stimuli as well as they did rich instances of the same objects. Younger children recognized the rich instances but not the part caricatures. Further studies have replicated this developmental trend (Augustine, Smith & Jones, 2011; Pereira & Smith, 2009; Son, Smith & Goldstone, 2008). Critically, the strength of individual children's planar bias in self-generated object views when manually engaged with objects predicts their ability to recognize objects in terms of sparse part representations (James, Swain, Jones & Smith, 2013). This fact suggests that manual actions on objects, actions that generate structured views around the most elongated axis, are developmentally linked to the development of object-centered representations of 3-dimensional shape. The path is from trunk control, to sitting steadily, to visual and manual interactions with objects, to sparse representations of 3 dimensional shape. In the end, the developmental story will require integrating Thelen's (1995) advances in the self-organization of motor development with Marr's (1982) computational-level theory of vision.

### **Paths 2 and 3: Visual object recognition and noun learning**

A considerable literature links the development of these sparse 3-dimensional representations of object shape to word learning: (a) young children's ability to recognize sparse geometric versions like those shown in Figure 4 is strongly correlated with productive vocabulary size and more strongly with vocabulary than with age (Smith, 2003; Pereira & Smith, 2009) (b) late talkers show deficits in recognizing sparse part caricatures of basic level categories (Jones & Smith, 2005); (c) representations of object shape in terms of the sparse part structure supports broad generalization of categories (Son, Smith, & Goldstone, 2008; Yee, Jones & Smith, 2012). All of these results suggest that the changes in the visual representation of object shape that occur between 12- and 36-months (see Smith 2009) – changes that appear to grow out of manual engagement with objects – set the stage for the rapid learning of object names. Manipulating and playing with objects prepares the visual system for forming shape based categories and thus for learning object names.

Other evidence suggests Path 3: learning object names also teaches children to attend to object shape and fosters more abstract and category relevant representations of shape (cite Smith 2009; Smith & Jones, 2011). Teaching children object names that refer to categories well organized by shape enhances attention to shape (and future object name learning, see, Perry, Samuelson, Malloy and Schiffer, 2010; Samuelson, 2002; Smith et al 2002). Other work shows that the recognition of the sparse part versions of 3-dimensional develops incrementally and is more advanced for better known than lesser known noun categories (Augustine, Jones, & Smith, 2011). So sitting engenders manual and visual exploration of 3-dimensional objects creating dynamic visual experiences that builds visual representations that support generalizing shape-based categories, and these representations support object name learning, and the learning object names (and basic level categories) feeds back on and refines those representations. It is all connected.

#### Path 4. Holding objects and stabilizing the head

Hands and eyes work together in goal directed action on objects (e.g., Land & Hayhoe, 2001; Pelz, Hayhoe & Loeber, 2001; Yoshida & Smith, 2008). Our recent work suggests that this perception action loop also plays a role in real time processes of stabilizing visual attention on an object and supporting the binding of a heard name to the seen thing. Newly moving toddlers (12 to 24 month olds) move their heads more often and nearly twice as fast as 3 year olds (Shen et al, 2010). This is primarily because they have not yet learned to compensate for the physical forces generated by their own body movements, so many movements are big movements (and potentially yield to falls, see, e.g, Bertenthal, Rose & Bai, 1997). Head stabilization is a particular problem when infants first begin to sit independently (Bertenthal & von Hofsten, 1998) and to walk (e.g., Ledebt, 2000). Recent evidence shows that holding objects stabilizes the head of newly walking infant (Claxton, Melzer, Ryu, & Haddad, 2012). Head stabilization may also stabilize and localize visual attention and in so doing support visual learning (Kerr, Condon, & McDonald, 1985).

The mechanistic basis for the proposal of a developmental pathway from holding objects to sustained attention begins with Posner's (1980) classic paper on attention as a spatial spotlight. Since then extensive research has documented the importance of localized attention for visual processing (e.g., Luck & Vecera, 2002; Yantis, 2008), for binding elements into a unified object representation (Treisman, 2004), for indexing and keeping track of objects in working memory (e.g., Makovaki & Jiang, 2009), and for the rapid detection and processing of objects (e.g., Ling & Carrasco, 2006; Yantis, 2008). Experimental tasks show that adults can readily attend to one specific location (and more rapidly detect objects at that location) without moving the eyes and while eye gaze is fixated elsewhere (e.g., Shepherd, Findlay & Hockey, 1986). Thus spatial attention in adults is internal and does not require moving the sensors toward the attended object. However, attention is also tied to the body. Adults typically orient eye gaze to the attended location. Moreover, eye movements (Grosbras et al, 2005; Rizzolatti et al, 1987), head movements (Colby & Goldberg, 1999), and even hand movements (Hagler, Riecke & Sereno, 2007; Knudsen, 2007; see also Thura et al, 2008) bias visual attention in the direction of the movement. Visual attention thus appears coupled to mechanisms of directional action – perhaps, because, more often than not, we direct attention in preparation for action. Consistent with these ideas are current discoveries about the involvement of motor planning regions in cortical attentional networks (Hagler et al, 2007, Knudsen, 2007; Kelley, et al, 2008, Collins, Heed, & Roder, 2010). Infants and young children's attentional systems may be more tied to bodily action and develop in part through developments in motor planning systems.

A core theoretical problem in understanding spatial action (and thus perhaps also in understanding spatial attention) is the coordination of frames of reference that specify the location of targets with respect to the body. The problem is the body has many different reference frames (e.g., Schlicht, R. Schrater, 2007). For example, reaching to an object typically involves turning the eyes or head to the object and then moving the hand in the direction of eye-gaze (see, Jeannerod, 1997). But the spatial coordinates of the object with respect to the eye, the head, and the hand are all different and require integration (Mullette-Gillman et al, 2005) or remapping into a common reference frame (Cohen & Andersen, 2002) if eye, head, and hand are to smoothly move to the same location. Laboratory studies of reaching often fix head position and most contemporary theories of adult reaching assuming that the common reference frame for both action and attention is eye-centered. However, a number of studies suggest that head direction plays a strong role in stabilizing eye-gaze direction in natural action contexts (e.g., Einhauser et al, 2007; Flanders et al, 1999). Other studies show that reaches are more precise when hand, head and eye are aligned and disrupted when hand, head and/or eyes point in opposite directions (see

Jeannerod, 1997; Vercher et al, 1994). Such findings suggest online interactions among multiple reference frames for action. Research on motor development makes clear that very young children have trouble aligning multiple frames of reference and often solve motor-planning problems by clamping degrees of freedom by keeping eyes, hands, head, trunk all aligned and moving together. If the reference frames for visual attention overlap (or interact with) the same reference frames for planning action, then the prediction is that young children would attend best with stabilized and aligned eye, head, and hand. This is just what holding an object does for toddlers.

### Path 5. Stabilized head, sustained attention, and object name learning

Our recent head-camera studies provide compelling evidence for a link between holding an object, sustained visual attention to that object, and learning an object name (Pereira, Smith & Yu, 2013; Smith, Yu & Pereira, 2011; Yu et al, 2009; Yu & Smith, 2012). In all the experiments, we asked parents to play with their toddlers with three toys at time. We recorded the child's first person view using a wide lens head camera. The head camera captures a head-centered view –the moment-to-moment *available* visual information and its changes as the child moves and changes the view of objects. Figure 4a shows the dynamic real time changes in the image size of objects in the head camera view for one typical toddler in the toy play task (Smith et al, 2011). A large image size means the object is unoccluded and close to the head and eyes; when image size drops to zero, the object is not in view. The child's whole-body action –and grasping and holding objects close–creates a view that is highly dynamic: Objects go rapidly in and of view and at any moment in time, there is often just one object dominating the head-camera view (Smith et al, 2010). However, amidst these dynamic switches from one object to another, there are moments of visual stability. These occur when children are holding an object (Yu et al, 2009; Yu & Smith, 2012). During holding, the held object is (1) large in image size (closer to the child than other objects and often obstructing the view of other objects), (2) near the center of the head camera image, and (3) in terms of low level saliency properties, highly salient. Moreover, the head-centered view maintains these properties stably over time while the object is being held.

In two experiments (Yu & Smith, 2012, Pereira et al, 2013), we asked parents and their 18 month olds to play with novel objects, and prior to play we taught parents the names of those objects. We asked parents to name the objects, when it seemed suitable, naturally as they played. After play we tested the toddlers in a name comprehension task to determine if they had learned any of those names. We then went back and looked at the images from the toddler's head-camera images from the play session to determine how parent naming events that led to learning by the toddler differed from those that did not.

There were always 3 objects in play (of roughly the same real size and bottom up saliency). Therefore, when a parent named an object there was the one target object (the intended referent) and two potential competitors for attention. We analyzed the sensory properties of the naming target and competitors for a times series from 10 sec before to after the parent's naming of the object during play. Figure 5 shows the key findings: Parent naming events that led to learning had a unique visual signature: infants learned the object name when the named object, the target in the figure, dominated the visual field in image size relative to other objects in the infant's view, the competitors. Parent naming also led to learning when the named target was centered in the head camera image (which implies aligned head and eyes), and more centered than the competitor objects. Critically, naming events that led to learning (but not other naming events that did not) showed *enduring* and significant differences in these properties for the named object relative to visual competitors. Finally, these visual signatures of learning coincided with the *toddler's* holding of the named object

(Yu & Smith, 2012). Holding brings the selected object close, blocking the view of competitors and holding stabilizes and aligns eyes, heads and hands and by hypothesis, these alignments may localize and sustains visual attention, leading to learning.. Notice how this developmental pathway integrates across usually disparate subfields in psychology: This pathway takes us from Posner (1980) though Jeannerod (1997) to what your first-grade teacher knew: sit up straight and still with hands clasped at midline to pay attention.

### **Paths 6 and 7: Why pretend play in toddlers is diagnostic of later language**

Two-year-old children often play with objects in a way that has been of special interest to researchers of early language. In this play, children substitute one object for another – for example, using a pot as a hat, a stick as a sword, or a cardboard box as a boat (Bergen, 2002; Bretherton et al, 1994; McCune, 1995; Piaget, 1962). These object substitutions are linked to early language development, with their absence being a diagnostic marker of significant language delay (e.g., Bergen, 2002; Rescorla & Goosens, 1992; Rutherford et al., 2006) that is used in clinical assessments of language and other developmental disorders (e.g., Johnson et al, 2008; Lewis et al, 2000). This form of play emerges in typically developing children between 18 and 30 months – the same age range in which children begin to recognize basic level categories from the sparse part structure (Smith, 2003) and at the same time that object name vocabularies are rapidly expanding (Bergen, 2002; Lewis et al, 2000; McCune-Nicolich, 1981; Shore et al, 1984).

The tie between object substitutions and language development is classically attributed to a shared “symbolic function”: for example, the pot on the child’s head and the word hat both “stand for” a real hat (e.g., Lilliard, 1993; Piaget, 1962). Consistent with this idea, a number of researchers (McCune-Nicolich, 1981; Shore et al, 1984; Striano, Tomasello & Rochat, 2001) have noted constraints on the types of objects that children substitute for others. The substituted objects tend to be simple in shape and to have minimal surface details, and thus perhaps are symbol-like. Thus, a banana might be substituted for a phone, but a richly detailed toy truck would not be. Critically, the shape of the substituted object is also geometrically similar to the shape of the replaced object (Bretherton et al, 1984). This observation led us to test the hypothesis that the emergence of object substitutions in play was a product of developmental changes in visual object recognition and specifically in the sparse representation of 3-dimensional object shape (Smith & Jones, 2011). We found that the ability to recognize basic level categories from sparse 3-dimensional shape representations strongly predicted object substitutions in play even with language and vocabulary size controlled. The link makes sense, because to see a bucket as like a hat, or a banana as like a phone, one has to see the common abstract shape across these different categories.

Why then is pretend play, and specifically objects substitutions in play, predictive not just of current language but of future language? The pathways in Figure 4 provides an explanation of why failure to develop object substitution is diagnostic of future language delay. Object substitutions in play are like the canary in the coal mine: they are not causally related to language delay, but their absence is an easily detected signal of a problem in language acquisition. As shown in the figure, early learning of object names promotes (and is supported by) the formation of increasingly abstract models of 3-dimensional shape. These newly formed representations invite and support the substitution of geometrically appropriate objects for one another in play. These substitutions are predictive of later language development because later language is causally dependent on early language development. Early language development (which consists mostly of learning object names) supports changes in visual object recognition and these changes in object recognition lead, along with other developments, to symbolic play. The absence of object substitutions in children’s play is thus a surface sign of a weakness in language learning.



The developmental links between object name learning, visual object recognition, and pretend play highlight the cascade that is developmental process – that development consists of many interacting and mutual dependencies across systems that may seem at first unrelated (Thelen & Smith, 1994). The results also focus attention on object recognition as a component of developmental change in what on the surface appears to be an unrelated competency. In this case, changes in visual object recognition matter to the emergence of object substitutions in play, and may be the source of the link of these object substitutions to language development. There may be other unsuspected consequences of ongoing changes in object perception and representation. The development of visual object recognition has not been well studied, particularly outside of infancy (see, Nishamura, Scherf & Berhman, 2009; Smith 2009), and many researchers of cognitive development assume that the infants' and toddlers' visual recognition is like that of adults. But emerging evidence suggests that it is not, and instead is not fully mature until adolescence (Juttner et al, 2012; Nishamura, et al, 2009; Rentschler et al, 2004). Humans are visual animals and the present results suggest that the increasing sophistication of children's visual object recognition is likely to be part of the pathways producing developmental changes in many cognitive domains.

## A pathways approach

A pathways approach is relevant to the big questions that motivate much of current research in cognition: What does it mean to be human? To what extent is human cognition determined (or predetermined) by evolutionary history and the innate constraints of our genes (and the more opportunistic epigenesis of gene action)? Certainly, when one looks about there is wonderful universality to the humans, in social behavior, in language, in ways of thinking. But there are also clear differences among individuals (Baldassarre et al, 2012) and among individuals living in different cultures (Markus & Kitayama, 1991). Understanding developing pathways (and the underlying mechanisms) provides insights into the underlying truths of human universality and variability.

Here is the unifying idea: Each infant and child is an individual and develops and changes as an individual. Both the intrinsic biology and the environment may be systematically constrained and thus canalize development outcome. But the developing organism has to travel that path. Because developmental process is degenerate and opportunistic, depending –at each moment in time that creates change –on the current state and abilities of the organism and on the idiosyncracies of the environment at that moment, development itself must be highly specific to the individual. Different children will follow different developmental paths that depend on the specific tasks they discover and the intrinsic dynamics of their own system, even if they end up, more or less, with the same set of human competencies.

One elegant demonstration of the individual nature of developmental trajectories is Thelen, Corbetta, Kamm, Spencer, Schneider, and Zernicke's (1993) week-by-week study of the transition from not-reaching to reaching for visually presented objects. Thelen et al. studied four babies and found four different patterns of activity, and thus, four different patterns of development. The basic developmental pattern was: The presentation of an enticing toy is arousing and elicits all sorts of nonproductive actions, and very different actions in individual babies. These actions are first, quite literally, all over the place with no clear coherence in form or direction. But by acting, each baby in its own unique fashion, sooner or later makes contact with the toy—banging into or brushing against it or swiping it. These moments of contact select some movements, carving out patterns that are then repeated with increasing frequency. Over weeks, the cycle repeats—arousal by the sight of some toy, action, and occasional contact. Over cycles, increasingly stable, more efficient and more effective forms of reaching emerge.

As infants produce different movements—in their uncontrolled actions initiated by the arousing sight of the toy—they each discover initially different patterns and different developmental tasks to be solved. Some babies in the non-reaching period hardly lift their arms at all. Other babies flail and flap and are always moving. These different babies must solve different problems to grasp an object. The flailer needs to become less active lowering the hands to bring them to midline and create balance. The placid baby needs to be more active, to raise her hands and to lift them up.

What is remarkable in the developmental patterns observed by Thelen and collaborators is that each infant found a solution by following individual developmental pathways that eventually converged to highly similar outcomes. Because action defines the task and because action—through the coordination of heterogeneous sensory systems—finds the solution, development is very much an individual and context-dependent matter, and not pre-defined prior to action itself. The fact that each infant must follow its own path, and makes its own perhaps unique way to maturity, is grounds for optimism for building effective strategies for children born with neural, bodily, and environmental limitations (Ansari & Karmiloff-Smith, 2002; Galloway et al, 2008; Ulrich, Ulrich, Angulo-Kinzer & Yun, 2001). Because developmental pathways are degenerate, because development builds on itself, one can—at particular junctures in these paths—create workarounds, alternative routes. If we know the pathways, we can create scaffolds for development at just the right points. For most typically developing children, given the constraints of the world, of human bodies, and of the heterogeneous and multimodal system out of which intelligence is made, these different individuals will develop broadly similar systems (what one might summarize as “universals”) but at its core, development (like evolution) is opportunistic, individualistic, and local in its causes. Developing organisms solve a series of overlapping tasks in time, following pathways that are constrained, but perhaps not determined, by many redundancies created by the properties of the developing system and the environments that the developing system plays an active role in creating.

## Acknowledgments

The research reviewed in this paper and the preparation of the paper were supported in part by R01HD 28675 and R21HD068475 from NICHD.

## References

- Ansari D, Karmiloff-Smith A. Atypical trajectories of number development: A neuroconstructivist perspective. *Trends in Cognitive Sciences*. 2002; 6(12):511–516. [PubMed: 12475711]
- Augustine E, Smith LB, Jones SS. Parts and relations in young children’s shape based object recognition. *Journal of Cognition and Development*. 2011; 12(4):556–572.
- Baldassarre A, Lewis CM, Committeri G, Snyder AZ, Romani GL, Corbetta M. Individual variability in functional connectivity predicts performance of a perceptual task. *Proceedings of the National Academy of Sciences*. 2012; 109(9):3516–3521.
- Bates JE, Viken RJ, Alexander DB, Beyers J, Stockton L. Sleep and adjustment in preschool children: Sleep diary reports by mothers relate to behavior reports by teachers. *Child Development*. 2002; 73(1):62–75. [PubMed: 14717244]
- Bergen D. The role of pretend play in children’s cognitive development. *Early Childhood Research and Practice*. 2002; 4:1–13.
- Bertenthal BI, Rose JL, Bai DL. Perception-action coupling in the development of visual control of posture. *Journal of Experimental Psychology: Human Perception and Performance*. 1997; 23(6): 1631. [PubMed: 9425672]
- Bertenthal B, Von Hofsten C. Eye, head and trunk control: the foundation for manual development. *Neuroscience & Biobehavioral Reviews*. 1998; 22(4):515–520. [PubMed: 9595563]

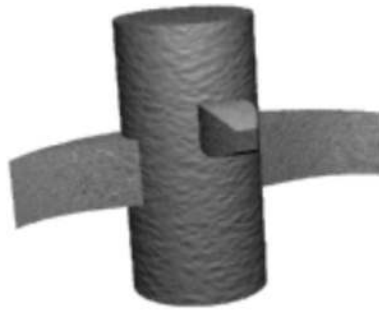
- Biederman I. Recognition-by-components: A theory of human image understanding. *Psychological Review*. 1987; 94:115–117. [PubMed: 3575582]
- Bretherton, I.; O'Connell, B.; Shore, C.; Bates, E. The effect of contextual variation on symbolic play: Development from 20 to 28 months. In: Bretherton, I., editor. *Symbolic Play: The Development of Social Understanding*. Orlando, FL: Academic Press; 1984.
- Claxton LJ, Melzer DK, Ryu JH, Haddad JM. The control of posture in newly standing infants is task dependent. *Journal of Experimental Child Psychology*. 2012
- Cohen YE, Andersen RA. A common reference frame for movement plans in the posterior parietal cortex. *Nature Reviews Neuroscience*. 2002; 3(7):553–562.
- Colby C, Goldberg M. Space and attention in the parietal cortex. *Annu. Rev. Neurosci*. 1999. 1999; 22:319–49.
- Collins T, Heed T, Röder B. Visual target selection and motor planning define attentional enhancement at perceptual processing stages. *Frontiers in Human Neuroscience*, 4. 2010 Mar 5.
- Cutzu, F.; Tarr, MJ. Representation of three dimensional object similarity in human vision. Paper presented at the SPIE Electronic Imaging: Human Vision and Electronic Imaging II; San Jose, CA. 2007.
- Edelman, GM. *Neural Darwinism: The theory of neuronal group selection*. New York: Basic Books; 1987. p. 167-188.
- Edelman GM, Gally JA. Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences, USA*. 2001; 98(24):13763–13768.
- Elman JL. Learning and development in neural networks: The importance of starting small. *Cognition*. 1993; 48(1):71–99. PMID: 8403835. [PubMed: 8403835]
- Einhauser W, et al. Human eye-head co-ordination in natural exploration. *Network*. 2007; 18(3):267–97. [PubMed: 17926195]
- Farivar R. Dorsal-ventral integration in object recognition. *Brain Research Reviews*. 2009; 61(2):144–153. [PubMed: 19481571]
- Flanders ML, Daghestani A. Berthoz, Reaching beyond reach. *Experimental Brain Research*. 1999; 126(1):19–30.
- Fox SE, Levitt P, Nelson III CA. How the timing and quality of early experiences influence the development of brain architecture. *Child Development*. 2010; 81(1):28–40. PMID: PMC2846084. [PubMed: 20331653]
- Graf M. Coordinate transformations in object recognition. *Psychological Bulletin*. 2006; 132(6):920–945. [PubMed: 17073527]
- Galloway JCC, Ryu JC, Agrawal SK. Babies driving robots: self-generated mobility in very young infants. *Intelligent Service Robotics*. 2008; 1(2):123–134.
- Gottlieb G. Experiential canalization of behavioral development: Theory. *Developmental Psychology*. 1991; 27(1):4–13.
- Grosbras M, Laird AR, Paus T. Cortical regions involved in eye movements, shifts of attention, and gaze perception. *Human Brain Mapping, Special Issue: Meta-Analysis in Functional Brain Mapping*. 2005; 25(1):140–154.
- Hagler DJ, Riecke L, Sereno MI. Parietal and superior frontal visuospatial maps activated by pointing and saccades. *NeuroImage*. 2007; 35:1562–1577. [PubMed: 17376706]
- Hayward WG, Tarr MJ. Testing conditions for viewpoint invariance in object recognition. *Journal of Experimental Psychology-Human Perception and Performance*. 1997; 23(5):1511–1521. [PubMed: 9411023]
- James KH, Swain SN, Jones SS, Smith LB. Young Children's Self-Generated Object Views and Object Recognition. *Journal of Cognition and Development*. 2013 online view.
- James KH, Humphrey GK, Vilis T, Corrie B, Baddour R, Goodale MA. “Active” and “passive” learning of three-dimensional object structure within an immersive virtual reality environment. *Behavior Research Methods, Instruments & Computers*. 2002; 34:383–390.
- James KH, Jones SS, Swain S, Pereira A, Smith LB. Some views are better than others: Evidence for a visual bias in object views self-generated by toddlers. *Developmental Science*. (under revision).
- Jeannerod, M. *The Cognitive Neuroscience of Action*. New York: 1997. 1997

- Johnson KC, DesJardin JL, Quittner AL, Winter ME. Assessing joint attention and symbolic play in children with cochlear implants and multiple disabilities: two case studies. *Otology and Neurotology*. 2008; 29:246–250. [PubMed: 18223452]
- Jones SS, Smith LB. Object name learning and object perception: a deficit in late talkers. *Journal of Child Language*. 2005; 32(1):223–240. [PubMed: 15779885]
- Jüttner M, Wakui E, Petters D, Kaur S, Davidoff J. Developmental trajectories of part-based and configural object recognition in adolescence. *Developmental Psychology*. 2012
- Kelley TA, Serences JT, Giesbrecht B, Yantis S. Cortical mechanisms for shifting and holding visuospatial attention. *Cerebral Cortex*. 2008; 18(1):114–125. [PubMed: 17434917]
- Kerr B, Condon SM, McDonald LA. Cognitive spatial processing and the regulation of posture. *Journal of Experimental Psychology: Human Perception and Performance*. 1985; 11(5):617. [PubMed: 2932533]
- Knudsen EI. Fundamental components of attention. *Annual Review of Neuroscience*. 2007; 30:57–78.
- Land M, Hayhoe M. In what ways do eye movements contribute to everyday activities? *Vision Research*. 2001; 41(25–26):3559–3565. [PubMed: 11718795]
- Ledebt A. Changes in arm posture during the early acquisition of walking. *Infant Behavior and Development*. 2000; 23(1):79–89.
- Lewis V, Boucher J, Lupton L, Watson S. Relationships between symbolic play, functional play, verbal and nonverbal ability in young children. *International Journal of Communication Disorders*. 2000; 35:117–127.
- Lilliard A. Pretend play skills and the child's theory of mind. *Child Development*. 1993; 64:348–371. [PubMed: 8477622]
- Lord K. A Comparison of the Sensory Development of Wolves (*Canis lupus lupus*) and Dogs (*Canis lupus familiaris*). *Ethology*. 2013; 119(2):110–120.
- Lickliter R, Honeycutt H. Developmental dynamics: toward a biologically plausible evolutionary psychology. *Psychological Bulletin*. 2003; 129(6):819. [PubMed: 14599279]
- Ling S, Carrasco M. Sustained and transient covert attention enhance the signal via different contrast response functions. *Vision Research*. 2006; 46(8–9):1210–1220. [PubMed: 16005931]
- Luck, SJ.; Vecera, SP. Attention. In: Pashler, H.; Yantis, S., editors. *Steven's handbook of experimental psychology Sensation and perception*. 3rd ed. Vol. 1. Hoboken, NJ, US: John Wiley & Sons Inc.; 2002. p. 235-286.
- Lungarella M, Pegors T, Bulwinkle D, Sporns O. Methods for quantifying the informational structure of sensory and motor data. *Neuroinformatics*. 2005; 3:243–262. [PubMed: 16077161]
- Lungarella M, Sporns O. Mapping information flow in sensorimotor networks. *PLoS Computational Biology*. 2006; 2:1301–1312.
- Makovski T, Jiang YV. The role of visual working memory in attentive tracking of unique objects. *Journal of Experimental Psychology: Human Perception and Performance*. 2009; 35(6):1687–1697. [PubMed: 19968429]
- Markus HR, Kitayama S. Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*. 1991; 98(2):224.
- Marr, D. *Vision: A computational investigation into the human representation and processing of visual information*. New York, NY: Henry Holt and Co. Inc.; 1982.
- Marr D, Nishihara HK. Representation and recognition of spatial-organization of 3-dimensional shapes. *Proceedings of the Royal Society of London, Series B: Biological Sciences*. 1978; 200:269–294.
- McCune-Nicolich L. Toward symbolic functioning: Structure of early pretend games and potential parallels with language. *Child Development*. 1981; 52:785–797.
- McIntosh AR, Fitzpatrick SM, Friston KJ. On the marriage of cognition and neuroscience. *Neuroimage*. 2001; 14:1231–1237. [PubMed: 11707079]
- Metta G, Fitzpatrick P. Early integration of vision and manipulation. *Adaptive Behavior*. 2003; 11:109–128.

- Mullette-Gillman OA, Cohen Y, Groh J. Eye-centered, head-centered, and complex coding of visual and auditory targets in the intraparietal sulcus. *Journal of Neurophysiology*. 2005; 94(4):2331–2352.
- Nelson CA. The development and neural bases of face recognition. *Infant and child development*. 2001; 10(1–2):3–18.
- Newport EL. Maturational constraints on language learning. *Cognitive Science*. 1990; 14(1):11–28.
- Nishimura M, Scherf S, Behrmann M. Development of object recognition in humans. *F1000 Biology Reports*. 2009:56. [PubMed: 20948628]
- Pelz J, Hayhoe M, Loeber R. The coordination of eye, head, and hand movements in a natural task. *Experimental Brain Research*. 2001; 139(3):266–277. [PubMed: 11545465]
- Pereira AF, James KH, Jones SS, Smith LB. Early biases and developmental changes in self-generated object views. *Journal of vision*. 2010; 10(11)
- Pereira AF, Smith LB. Developmental changes in visual object recognition between 18 and 24 months of age. *Developmental science*. 2009; 12(1):67–80. [PubMed: 19120414]
- Pereira AF, Smith LB, Yu C. A bottom-up view of toddler word learning. *Psychological Bulletin & Review*. 2013
- Perrett DI, Harries MH, Looker S. Use of preferential inspection to define the viewing sphere and characteristic views of an arbitrary machined tool part. *Perception*. 1992; 21:497–497. [PubMed: 1437467]
- Perry LK, Samuelson LK, Malloy LM, Schiffer RN. Learn locally, think globally exemplar variability supports higher-order generalization and word learning. *Psychological science*. 2010; 21(12):1894–1902. [PubMed: 21106892]
- Piaget, J. *The origins of intelligence in children*. Oxford, England: International Universities Press; 1952.
- Piaget, J. *Play, Dreams, and Imitation in Childhood*. New York: Norton and Company; 1962.
- Posner MI. Orienting of attention. *The Quarterly Journal of Experimental Psychology*. 1980; 32(1):3–25. [PubMed: 7367577]
- Rentschler I, Jüttner M, Osman E, Müller A, Caelli T. Development of configural 3D object recognition. *Behavioral Brain Research*. 2004; 149(1):107–111.
- Rescorla L, Goosens M. Symbolic play development in toddlers with expressive specific language impairment. *Journal of Speech Hearing Research*. 1992; 35:1290–1303. [PubMed: 1283611]
- Rizzolatti G, Riggio L, Dascola I, Umiltá C. Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia*. Special Issue: Selective Visual Attention. 1987; 25(1-A):31–40.
- Ruff HA. Role of manipulation in infants' responses to invariant properties of objects. *Developmental Psychology*. 1982; 18(5):682.
- Rutherford M, Young G, Hepburn S, Rogers S. A longitudinal study of pretend play in autism. *Journal of Autism and Developmental Disorders*. 2006; 37:1024–1039. [PubMed: 17146707]
- Samuelson LK. *Statistical regularities in vocabulary guide*. 2002
- Schlicht EJ, Schrater PR. Impact of coordinate transformation uncertainty on human sensorimotor control. *Journal of Neurophysiology*. 2007; 97(6):4203–4214. [PubMed: 17409174]
- Shen, H.; Baker, TJ.; Candy, TR.; Yu, C.; Smith, LB. Development and Learning (ICDL), 2010 IEEE 9th International Conference on. IEEE; Aug. 2010 Using the head to stabilize action: Reaching by young children; p. 108-113.
- Shepherd M, Findlay JM, Hockey RJ. The relationship between eye movements and spatial attention. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*. 1986; 38(3-A):475–491.
- Sheya, A.; Smith, LB. Development through sensory-motor coordinations. In: Stewart, J.; Gapenne, O.; Di Paolo, E., editors. *Enaction: Towards a new paradigm for cognitive science*. Cambridge, MA: MIT Press; 2009.
- Shore C, O'Connell B, Bates E. First sentences and symbolic play. *Developmental Psychology*. 1984; 20:872–880.

- Smith LB. Learning to recognize objects. *Psychological Science*. 2003; 14(3):244–250. [PubMed: 12741748]
- Smith LB. From fragments to geometric shape: Changes in visual object recognition between 18- and 24-months. *Current Directions in Psychological Science*. 2009; 18:290–294.
- Smith LB, Breazeal C. The dynamic lift of developmental process. *Developmental Science*. 2007; 10:61–68. [PubMed: 17181701]
- Smith LB, Jones SS. Symbolic play connects to language through visual object recognition. *Developmental Science*. 2011; 14(5):1142–1149. [PubMed: 21884329]
- Smith LB, Jones SS, Landau B, Gershkoff-Stowe L, Samuelson L. Object name learning provides on-the-job training for attention. *Psychological Science*. 2002; 13(1):13–19. [PubMed: 11892773]
- Smith LB, Yu C, Pereira AF. Not your mother's view: The dynamics of toddler visual experience. *Developmental science*. 2011; 14(1):9–17. [PubMed: 21159083]
- Son JY, Smith LB, Goldstone RL. Simplicity and generalization: Short-cutting abstraction in children's object categorizations. *Cognition*. 2008; 108(3):626. [PubMed: 18565504]
- Soska KC, Adolph KE, Johnson SP. Systems in development: Motor skill acquisition facilitates 3D object completion. *Developmental Psychology*. 2010; 46:129–138. [PubMed: 20053012]
- Sporns, O. *Networks of the Brain*. MIT press; 2011.
- Striano T, Tomasello M, Rochat P. Social and object support for early symbolic play. *Developmental Science*. 2001; 4:442–456.
- Thelen, E.; Smith, LB. *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press; 1994.
- Thelen E. Motor development: A new synthesis. *American psychologist*. 1995; 50(2):79. [PubMed: 7879990]
- Thelen E, Corbetta D, Kamm K, Spencer JP, Schneider K, Zernicke RF. The transition to reaching: Mapping intention and intrinsic dynamics. *Child development*. 1993; 64(4):1058–1098. [PubMed: 8404257]
- Thura D, Hadj-Bouziane F, Meunier M, Boussaoud D. Hand position modulates saccadic activity in the frontal eye field. *Behavioural Brain Research*. 2008; 186(1):148–153. [PubMed: 17881066]
- Treisman, A. Psychological issues in selective attention. In: Gazzaniga, MS., editor. *The cognitive neurosciences*. 3rd ed. Cambridge, MA, US: MIT Press; 2004. p. 529-544.
- Tse PU. Volume completion. *Cognitive Psychology*. 1999; 39(1):37–68. [PubMed: 10433787]
- Turkewitz G, Kenny PA. Limitations on input as a basis for neural organization and perceptual development: A preliminary theoretical statement. *Developmental Psychobiology*. 2004; 15(4): 357–368. PMID: 11677607. [PubMed: 7106395]
- Ulrich DA, Ulrich BD, Angulo-Kinzler RM, Yun J. Treadmill training of infants with Down syndrome: evidence-based developmental outcomes. *Pediatrics*. 2001; 108(5):e84–e84. [PubMed: 11694668]
- Vercher JL, et al. Eye-head-hand coordination in pointing at visual targets: spatial and temporal analysis. *Experimental Brain Research*. 1994; 99(3):507–23. [PubMed: 7957730]
- West MJ, King AP. Settling nature and nurture into an ontogenetic niche. *Developmental Psychobiology*. 1987; 20(5):549–562. PMID: 3678619. [PubMed: 3678619]
- Whitacre JM. "Degeneracy: a link between evolvability, robustness and complexity in biological systems". *Theoretical Biology and Medical Modelling*. 2010; 7(6):6. [PubMed: 20167097]
- Winkowski DE, Knudsen EI. Top-down gain control of the auditory space map by gaze control circuitry in the barn owl. *Nature*. 2006; 439(7074):336–339. [PubMed: 16421572]
- Yantis S. The neural basis of selective attention: Cortical sources and targets of attentional modulation. *Current Directions in Psychological Science*. 2008; 17(2):86–90. [PubMed: 19444327]
- Yee M, Jones SS, Smith LB. Changes in visual object recognition precede the shape bias in early noun learning. *Frontiers in psychology*. 2012:3. [PubMed: 22291676]
- Yoshida H, Smith LB. What's in view for toddlers? Using a head camera to study visual experience. *Infancy*. 2008; 13(3):229–248. [PubMed: 20585411]
- Yu C, Smith LB, Shen H, Pereira AF, Smith T. Active information selection: Visual attention through the hands. *Autonomous Mental Development, IEEE Transactions on*. 2009; 1(2):141–151.

Yu C, Smith LB. Embodied attention and word learning by toddlers. *Cognition*. 2012



a. original view



b. same object rotated

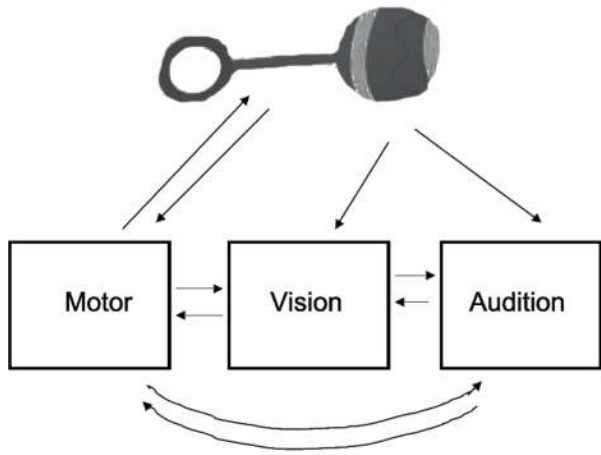


c. different object

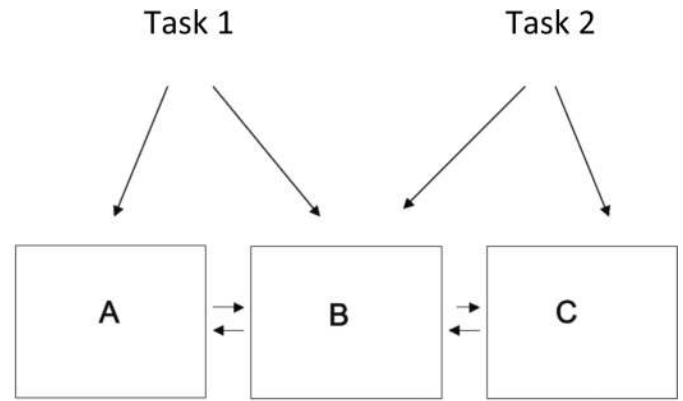
**Figure 1.**

Example of task requiring prediction of novel views of an object from a single view.  
(Illustrated volumes from CNBC object bank, Hayward & Tarr, 1997).





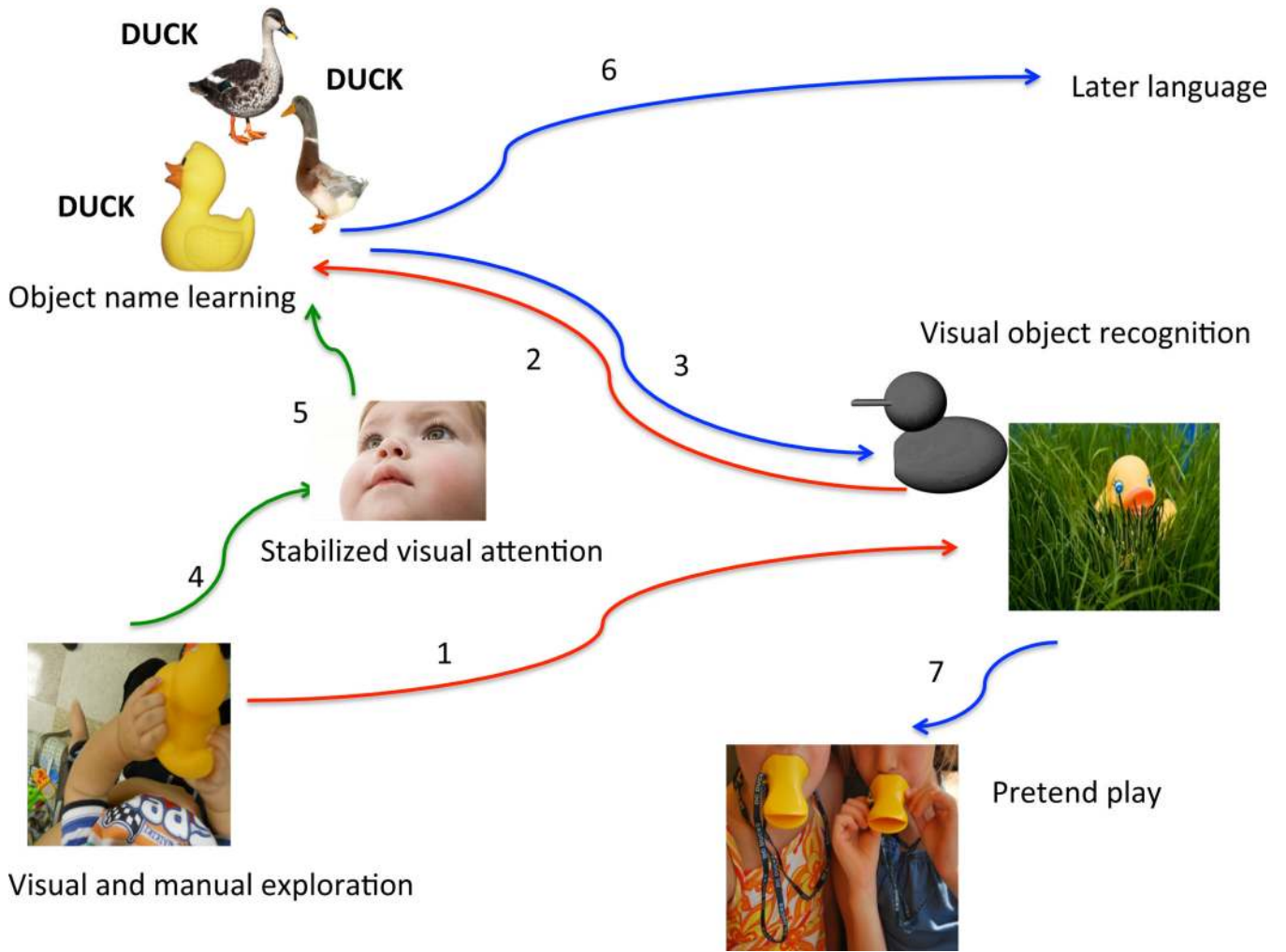
a. coordination of multiple systems in a single task



b. overlapping coordinations across tasks

**Figure 2.**

a. An illustration of how a task such as shaking a rattle recruits and coordinates multiple systems, setting up an in-task functional network that can lead to change in the system as a whole and the individual components. b. An illustration of overlapping coordinations across tasks.



**Figure 3.** Behavioral pathways relevant to the early object name learning and visual object recognition. See text for clarification of individual paths.

a. views of a rectangular block

nonplanar (3/4) view



planar views -elongated



axis of elongation  
*perpendicular* to LOS



planar view  
foreshortened



axis of elongation  
*parallel* to LOS

b. head-camera view  
of object exploration

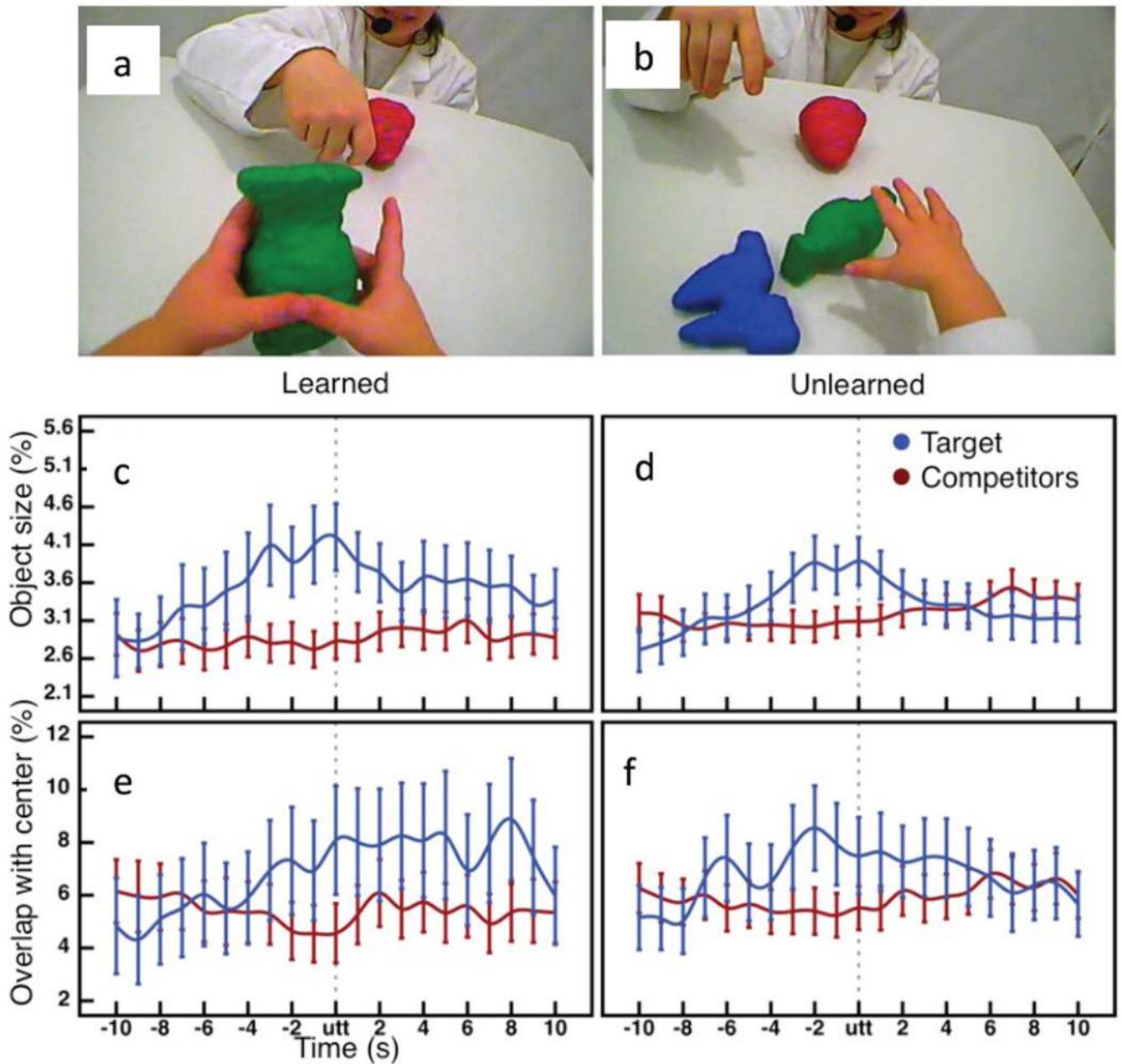


c. rich and sparse-part representations  
of common categories



**Figure 4.**

a. An illustration of planar and nonplanar views of a rectangular block. b. a head camera view of a toddler visually and manually exploring a novel object. c. Rich and sparse-part representations of common categories.



**Figure 4.**

a and b. Example Head camera images during two naming moments when later testing showed the child had learned the name (a) and not learned the name (b). c and d. The image size (5 pixels) of the named target (blue) and the mean of other in view, competitor, objects (red) for the 20 second window around the naming utterance (utt) for naming moments that lead to the learning of the object name (c) or did not (d). e and f. The overlap of the image of the named target and competitors with the center of the head camera image for the 20 second window around the naming utterance (utt) for naming moments that lead to the learning of the object name (e) or did not (f). See Yu & Smith (2012) and Pereira, Smith & Yu (2013) for technical details and related graphs.