## Chapter 16

# More than concepts: How multiple integrations make human intelligence

Linda B. Smith

### Editors' Preview

Chapter 16, 'More than concepts: How multiple integrations make human intelligence', provides a departure from the other chapters in the book because it questions from the start whether the term concept is a necessary theoretical construct. It questions the traditional cognitive framework, that is, sense–think–act, for explaining behaviour. In the standard cognitive account, behaviour arises as a consequence of mental representations or concepts of experience. However, in the framework advocated in Chapter 16, mental representations are more the 'ghost in the machine', and are perhaps better regarded as byproducts of more primary systems such as perception, action, emotion, and social interaction, each of which is closely coupled to the others in its activity, and all of which are connected to the world in which they operate. In other words, in this more 'dynamic-systems' framework, concepts are an epiphenomenal product of the coordination of multiple sensory–motor systems that allow us to respond to the world in intelligent ways.

Development plays a crucial role in the dynamic-systems framework and Chapter 16 provides illustrative examples of how multimodal subsystems become coordinated during real-time change to create concepts. For instance, infants who wear sticky mittens and are provided with the opportunity to act on (i.e. reach for) objects while looking at them (before they develop mental exploration skills) subsequently show increased object exploration through both vision and oral manipulation (i.e. mouthing behaviours). The enhanced sensory–motor coordination provided by the sticky mittens manipulation may thus promote the formation of object concepts by infants. In addition, infants who are provided with the opportunity to explore transparent containers through both vision and touch display subsequent strong

> performance in object-permanence tasks involving transparent boxes and in visual cliff tasks that involve traversing across a transparent surface that offers support, but which is visually specified as a falling-off point. By the dynamic-systems account, the exploration of the containers allows infants to coordinate sight and touch experiences, which in turn enables the development of a concept of a solid surface of support. Note that the concept in this case displays a hallmark characteristic of human intelligence which is that it is transportable and can be generalized from one context to another (from object search to spatial locomotion).
>
> Chapter 16 also considers the ways in which human intelligence may be unique, at least relative to different species of nonhuman animals. One candidate difference is that animal species do not always have the same type of coupling of various sensory and action systems. In the dynamic-systems account, such coupling is what allows the systems to learn from each other, thereby promoting the development of abstract concepts. For example, birds may have difficulty acquiring a generalizable concept of solid surface (that can be extended to transparent surfaces) because they cannot engage in the kind of eye–hand coordination that infants by 9 months of age are capable of. Another candidate difference is that of language. Hearkening back to the theme previously discussed in Chapters 6 and 13, words may equip humans with a more discrete, categorical, rule-like form of intelligence that provides a complement to a continuous, graded, more commonsense way of knowing the world. One final candidate (suggested also in Chapters 14 and 15) is the fact that our perceptions of and actions on the world are constrained and guided by the presence of conspecifics (who possess the same machinery for carrying out intelligent behaviour). Thus, according to Chapter 16, what may be uniquely human about concepts or intelligence (the latter being the preferred term in the chapter) are: (1) the overlapping operations of various perception–action systems (which in turn provide the redundant structure from which concepts are formed); (2) language; and (3) the socially embedded nature of our interaction with the world.

Understanding how and why human intelligence has the properties it does is certainly one of the most compelling questions in all of science. The phenomena in need of explanation are vast and varied, including behaviours of categorization, language and communication, imitation and learning from example, tool use, the invention of advanced symbol systems such as mathematics, as well as art and architecture. However, the topic of this book, 'concepts', is *not* a phenomenon in need of explanation. 'Concepts' are hypothetical constructs proposed by theorists to explain the sorts of phenomena listed above, and the jury is very much still out on whether such a theoretical construct is even needed (for differing views, see Beer, 2000; Brooks, 1991; Fodor, 1998; Keil, 1994; Smith & Katz, 1996) As venerable a figure as William James (1890) considered them misleading 'figments' and suggested psychologists abandon the construct altogether.

The construct of a 'concept' derives from the classic view in which mental life is divided into the discrete steps of 'sense–think–act'. Cognition, by definition, is about the 'think' part, the knowledge and processes that mediate perceiving and acting. Concepts, from this perspective provide the content for cognition and are amodal and propositional, consisting of relatively fixed and compositional representations. The work the concepts do as theoretical constructs, the argument for hypothesizing their existence in the first place, is that they provide stability and structure in the cognitive system. How can we recognize all varieties of dogs as 'dogs'? In this view, it is because all varieties of dogs activate the same fixed concept (e.g. Keil, 1994). How do we recognize the common structure of the solar system to an atom? In this view, it is because the componential and relational structure of the underlying concepts share significant similarities (Gentner, 1983). This approach has long dominated the study of human intelligence and without doubt has been the source of many significant insights. However, there are serious limitations to this approach, including explanations of the flexibility of human intelligence, its capacity to do the truly new (e.g. Smith & Katz, 1996), the unsolved problem of how anything like propositions emerge in the dynamic events of neural processing (Sporns, 2000), the problem of how new propositional representations (that are not new combinations of innate primitives) can emerge (see Barsalou, 1993; Fodor, 1999; Thelen & Smith 1994), and the disconnect between propositional representations and the real time and decidedly physical aspects of behaviour, learning, and development (e.g. Elman, 2004; O'Regan & Noe, 2001; Port & van Gelder, 1995; Smith & Gasser, 2005).

Accordingly, there are increasing discussions of an alternative view that does not segregate mental life into cognitive versus noncognitive processes and in which the construct of a 'concept' has a minimal role (or perhaps no role at all, see, Barsalou et al., 2003, 2007; Beer, 2000; Spencer & Schutte, 2004; Zwaan, 2004). This chapter considers human intelligence from this alternative 'more-than-concepts' point of view. There are several evolving forms of this newer approach (see Anderson, 2003; Wilson, 2002) but there are also common principles across these variants that are foundational to the ideas presented here:

1. *Knowledge has no existence separate from process*, but is instead embedded in, distributed across, and thus inseparable from real-time processes (e.g. Ballard et al., 1997; O'Regan & Noe, 2001; Samuelson & Smith, 2000; Spivey, 2006). From this perspective, there is no fixed and separate representation of anything, no concepts at all in the usual sense.
2. *Intelligence is made out of what have traditionally been viewed as noncognitive systems*, perception, action, emotion, and the coupling of these processes to the world and across individuals in social interactions (Hutchins, 1995; Rogoff, 2003; Tomasello et al., 1993). Rather than cognition being distinct from noncognitive processes, cognition *may simply be* the operation of this complex system as a whole.

3. *To understand intelligence is to know how to 'build' a behaving system that behaves intelligently.* One way to do this is to study how cognition develops from birth over time in biological systems. Another way is to pursue the evolutionary processes (and cross-species differences) that exploit developmental process. Alternatively, one could try to engineer intelligence by creating developing robots or by evolving artificial cognitive systems (Beer, 2000; Smith & Breazeal, 2007).
4. *Intelligence resides in a complex dynamic system that includes brain, body, and world* and can be understood only by understanding the continuous, real-time and closed-loop interactions of the brain through the body to the world and *back again.*

Given these starting principles, how does one explain the 'specialness' of human intelligence? Evolution provides an important clue. Many species do very smart things – navigate, calculate, detect, and smartly use subtle forms of regularities. But in many species, this intelligence is limited to specific tasks and contexts and is not transportable, nor inventive (Rozin, 1976). In contrast, the species we think of as having the most advanced forms of biological intelligence are advanced precisely because they are open systems, influenced by many sources of information, generalizing broadly and inventing new solutions. These same species also have long post-birth periods of immaturity. If this long period of immaturity is key to the inventive and open nature of human intelligence, then we need to understand just what is happening over that long period of development.

The theme of this chapter is that human intelligence emerges in the accrued effects of many different integrations of heterogeneous processes, each integration occurring as the consequence of *doing* some task. A complex system that must find solutions to many different overlapping tasks creates the specialness of human intelligence, an intelligence capable of inventing alphabets, writing poetry, building bridges, and jerry-rigging a broken door lock. The developmental processes that make human intelligence are bound to a structured world that includes conspecifics with a shared language, culture, and enduring artefacts and thus these contexts in which human development occurs are also important parts of the process. In brief, what is 'special' about human intelligence is precisely opposite of those accounts that propose specially evolved, innate, dedicated, and encapsulated systems.

## Integration

In his book, *The Origins of Intelligence in Children*, Piaget (1952) described a pattern of infant activity that he called a secondary circular reaction. A rattle would be placed in a 4-month-old infant's hands. As the infant moved the rattle, it would both come into sight and also make a noise, arousing and agitating the infant, causing more body motions, and thus causing the rattle to move into and out of sight and to make more noise. Infants at this age have very little organized control over hand and eye. They cannot yet reach for a rattle and if given one, they do not necessarily shake it. But if the infant accidentally moves it, and sees and hears the consequences, the infant will become captured by the activity – moving and shaking, looking and listening – and incrementally through this repeated action gain control over the shaking of the rattle and, indeed, acquire the intention to shake to produce noise. This pattern of activity – an accidental action that leads to an interesting and arousing outcome and to the intention to re-experience of the outcome – is foundational to intelligence.

In his book, *Neural Darwinism*, Edelman (1987) also pointed to the coupling of heterogeneous sensory–motor systems in the creation of cognition. Edelman's theory, like that of Piaget, starts by recognizing the multimodal nature of the brain at birth; it is – from the start – a complex system made up of many heterogeneous subsystems with their own intrinsic dynamics; but these subsystems are also highly interactive and densely interconnected within themselves and also over longer pathways to other subsystems. Like Piaget, Edelman proposed that development occurs through the coupled interactions of these subsystems to each other and to the physical world as they are engaged in real-time tasks. In the context of a specific task, distinct processes are time-locked to each other *and* to the world, and this creates adaptive change in the internal operating characteristics of those subsystems and in their connections to each other.

To illustrate this, Reeke and Edelman (1984) built a simple computational device. The device's task was to learn to recognize all varieties of the letter A, from the mere experience of looking at As. Figure 16.1 provides a schematic illustration. The feature-analysis subsystem consists of line detectors excited by corresponding patterns of stimulation. The tracing subsystem gathers information about shape through 'eye-movements' as the letter is scanned. The developmental power arises because these activation patterns in these two subsystems are time-locked to each other and to the same physical world enabling straightforward Hebbian learning to create systematic and adaptive change.

That is, at the same time that the feature analyser is analysing features, the shape tracer is extracting a global description of shape. The outputs of these two heterogeneous processes, at every step in time, are mapped to each other. There are actually seven mappings being accomplished simultaneously in real time. One mapping, the feature analysis map, maps an input letter to a list of features. The second mapping, the tracing map, maps the input letter to the action sequences of scanning. The third map – from the tracing process to the physical world – selects moment by moment the input (the specific letter part) to both subsystems. The fourth and fifth maps are the recurrent activity within each subsystem: at any moment in time, the activity in the feature analysis subsystem, for example, depends not only on the current input but also on its just preceding state. The sixth and seventh maps are what Edelman calls re-entrant maps; they map the activities of the two subsystems to each other. Thus, two independent mappings of the stimulus to internal activity take qualitatively different glosses on the perceptual information and through their re-entrant connections, by being correlated in real time and by being coupled to the same physical world, they educate each other. Reeke and Edelman's simulation successfully taught itself to recognize all varieties of A, generalizing to novel fonts and handwriting, merely from actively looking at As.

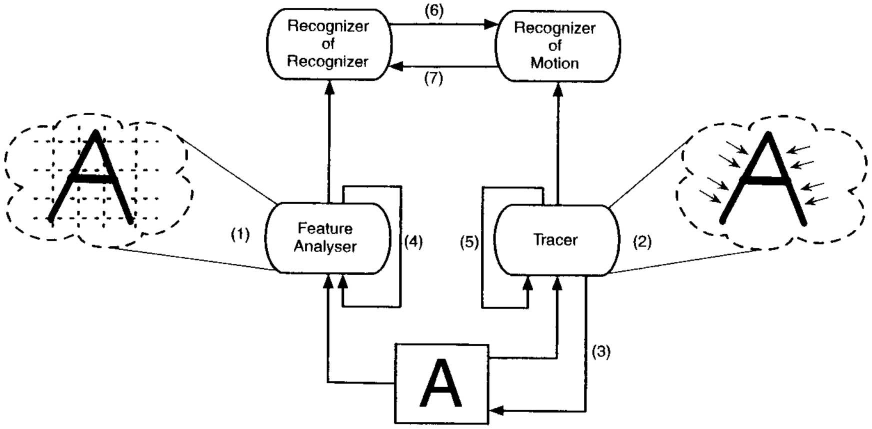

**Fig. 16.1** A schematic of Reeke's and Edelman's (1984) network model of letter recognition. The letter A at the bottom of the figure depicts the two-dimensional input array. This input is connected to both a feature analysis system and a tracing system. The recurrent connection for each of these systems represents the system's dependence not only on input but on its own history. The feature analysis system is composed of feature detectors which track the local structure of the input array, like an oriented line segment. This system outputs to a more abstract detector which integrates information across the local detectors capturing the global structure of the input array. The tracing system scans the input array and detects the contour of objects. This system, like the feature analysis system, outputs to a higher-level network which captures shared characteristics of related input arrays. The two higher-level networks are connected to each other enabling the two subsystems (feature analysis and tracing) to work together to classify letters.

The idea that coupled perception and action drives perceptual development is well supported by classic work such as Held and Hein's (1963) demonstration of changes in the visual system of kittens who actively explored their world and the lack of change in the visual system of kittens who passively viewed the same visual events (see also, Gonzalez et al., 2005; Harman et al., 1999; Hein & Diamond, 1972; Landrigan & Forsyth, 1974). The newer idea – with growing supporting evidence in cognitive neuroscience (Barsalou et al., 2005; Martin & Chao, 2001; Pulvermüller et al., 1999, 2005) and in computational studies of brain systems and learning (Lungarella & Sporns, 2006; Lungarella et al., 2005; McIntosh et al., 2001; Tononi, 2004) is that this self-generated activity through the coordination of multiple sensory–motor systems also creates the higher-level cognition characteristic of human intelligence.

## Multiple modalities

The human sensory motor (i.e. cognitive) system is far more complex than the model system shown in Figure 16.1. There are many more component subsystems and variable and complex patterns of connectivity among them. For example, many cortical brain regions associated with specific modalities are comprised of densely connected subregions that capture specific feature states within that single modality and there are also systems of connections among those feature areas. And, there are integrations across modalities (e.g. Martin & Chao, 2001; Pulvermüller et al., 2005; Rogers et al., 2007). Rapidly advancing work in systems neuroscience and in cognitive neuroscience increasingly document the pervasiveness of these cross-area integrations and indeed the greater interconnectivity of different sensory systems early in development (see Stiles, 2008). Overlapping, interacting, and redundant ways of knowing are a fundamental aspect of the human brain and may well be the source of abstract and transportable knowledge.

Multimodality may be key because of what Edelman (1987) calls *degeneracy*. Originally from the mathematical sense of the word, degeneracy in neural structure means that any single function can be carried out by more than one configuration of neural signals and that different neural clusters also participate in a number of different functions. As indicated in Figure 16.2, degeneracy is an intermediate level of complexity in a network between modularity and complete (or random) connections. A modular system is such that inputs in one module are unaffected by the activity in other modules, yielding highly stable but also highly limited patterns of activity. In a completely connected or random network, everything affects everything with the consequences of limited stability and considerable variability. In between is a degenerate network, with partially overlapping patterns of connectivity that include dense local connections and sparser longer pathways. The consequence is considerable complexity and many dynamically stable states (Sporns, 2002). The human brain is generally understood to be degenerate in this sense.

One consequence of degeneracy is that the whole can function even with the loss of one component. There are multiple routes to the same behavioural outcome. For example, because we encounter space through sight, sound, movement, touch, and even smell, we can know space even if we lack one modality. Being blind, for example, does not wipe out spatial concepts; instead, as studies of blind children show (Landau & Gleitman, 1985), comparable – and highly abstract – spatial concepts can be developed, presumably through the coordination of remaining redundant and overlapping subsystems.

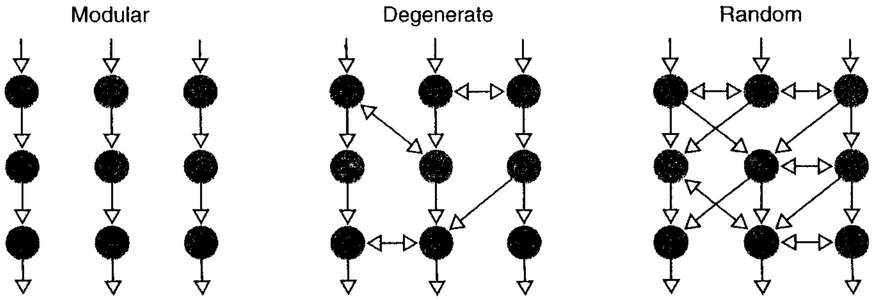

**Fig. 16.2** A schematic illustration of three kinds of connectivity with very different properties with respect to stability and inventiveness: a modular system, a degenerate system, and a random and densely connected system.

The partial redundancies in functional connectivity also mean that different tasks can recruit different but overlapping consortiums of subsystems, a fact of considerable *developmental* importance. In one suggestive experiment, Needham et al. (2002) fit 2- to 5-month-old infants with Velcro-covered 'sticky mittens'. These mittens enabled the infants to grab objects merely by swiping at them, enabling them to precociously coordinate vision and reaching. Infants who were given 2 weeks of experiences with 'sticky' mittens subsequently showed more sophisticated object exploration even with the mittens off. They looked at objects more, made more visually coordinated swipes at objects than did control infants who had no exploratory experiences with 'sticky mittens'. Needham et al. found that the sticky-mitten task not only facilitated the development of reaching for objects but also visual–oral exploration. That is, infants who had experience with sticky mittens looked at objects more – even in nonreaching tasks – and also mouthed and orally explored objects in more advanced ways.

Figure 16.3 provides a schematic illustration of what may be the profound significance of these results. Two subsystems – reaching and looking – are coordinated in the sticky-mitten task and in so doing educate each other. But these components are also involved in other coordinations, in other tasks that recruit other coalitions of subsystems. Thus, extra experience in the coordination of reaching and looking with sticky mittens ends up not being just about looking and reaching but potentially about other developments, other coordinations, generating cascading developmental consequences in other tasks in which some of the same subsystems are involved.

Infants' learning about transparent surfaces presents another example of generalization from overlapping integrations, and perhaps also, evidence for an 'abstract idea' made through such integrations. Transparent surfaces are interesting precisely because they violate the usual hand–eye correlations in the world. In most cases, one can directly reach to a seen object by following the line of sight. In contrast to this usual case, transparent yet solid surfaces block direct line of sight-reaching paths. There can be no evolutionary programme for learning about transparency since transparent and solid surfaces are not common in nature but are artefacts, invented by people. And indeed, some species (birds) seem unable to resolve this problem of transparency (see Malakoff, 2004).

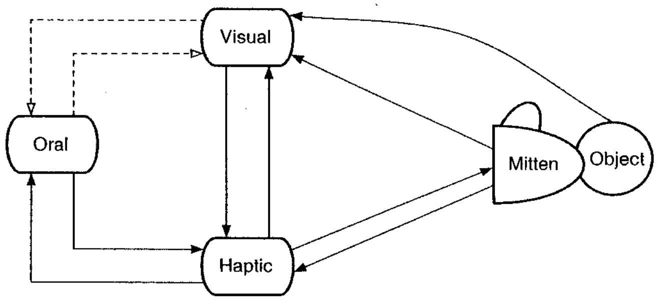

**Fig. 16.3** A schematic illustration of affect of 'sticky' mittens on the visual, haptic, and oral systems. The use of 'sticky' mittens during manual exploration reorganizes the coordination of the visual and haptic systems. The oral system is greyed in the figure because it is not directly involved in the activity with the 'sticky' mittens. However, it is connected to the haptic system – infants often manually and orally explore objects – and through this connection is potentially influenced by the visual–haptic reorganization. The visual–haptic reorganization leads to the coupling of the visual and oral systems – depicted by empty arrows with dotted lines – which are not directly linked.

Babies (like birds) do not do well with this violation of usual expectations. The experimental demonstration of infants' difficulties is provided by Diamond (1990) who presented 9-month-old infants with toys hidden under boxes. The boxes were either opaque – hiding the toy – or transparent, enabling the infants to see the toy under the box. As illustrated in Figure 16.3, the boxes were open on the side, so that infants, by reaching to that side, could retrieve the object. Diamond found that infants were able to retrieve the toy from the opaque container, reaching around to the side opening. However, they frustratingly failed in the same task with a transparent container. The problem with the transparent container is that infants attempt to reach for the toy in the usual way (trying to put their hand through the transparent surface following the line of sight) and thus fail. They do not search for and find the opening.

But infants, unlike birds, are *not* stuck with what is a maladaptive approach in the modern artefact-filled world. Infants easily learn to solve this problem through their ordinary interactions with transparent containers. In a demonstration of this point, Titzer (1997, see also Smith & Gasser, 2005) conducted a microgenetic study in which 8-month-old infants were given either a set of opaque or transparent containers to play with at home. Parents were given no instructions other than to put these containers in the toy box, making them available to the infants during play. When the infants were 9 months old, they were tested in Diamond's task. The babies who had played at home with opaque containers failed to retrieve objects from transparent ones just as in the original Diamond study. However, infants who had played at home with the transparent containers sought out and rapidly found the openings and retrieved the object from the transparent boxes. Infants' at-home explorations of the transparent containers did *not* include the specific task of sideways retrieval of objects, although it seems likely that in their spontaneous play, objects were both put into and retrieved from the openings of the containers. Titzer proposed that during play – through the coordination of seeing and touching as they put objects in and out of the containers – infants learned to recognize the subtle visual cues that distinguish solid transparent surfaces from openings and had learned that surfaces with the visual properties of transparency are, in fact, solid and thus to reach around to the openings. In the framework of Figure 16.1, the haptic cues from touching the transparent surfaces educated vision, and vision educated reaching and touch, enabling infants when subsequently tested in Diamond's task to find the openings in transparent containers.

Critically, these coordinations of touch and sight had broader cascading consequences beyond retrieving objects from small transparent containers, cascading effects that some might want to summarize under the umbrella of a 'concept of solid and supporting surface'. The result was obtained in an additional transfer task, the visual cliff. The visual cliff was originally designed to study infant depth perception (Gibson & Walk, 1960; Walk & Gibson, 1961). It consists of a transparent but solid surface placed over a visual 'drop off'. In the classic version of the task, 8- and 9-month-old infants are placed on the 'safe' side of the surface. Infants this age typically avoid the visual drop, not moving onto the transparent surface that extends over the vertical drop.

Titzer tested the babies who had participated in their training study with transparent and opaque containers on the visual cliff and found that the infants who played with the transparent containers at home did not avoid the visual cliff. Instead, they happily crawled onto the transparent surface over the drop off, showing no apprehension whatsoever. The babies who had played with opaque containers, in contrast, avoided the 'edge' refusing – even when called by their parent – to approach the visually (but not tactually) apparent cliff. The infants, who had extensive play with transparent containers, were apparently both sensitive to the subtle visual cues that specify the solidity of a transparent surface and to the cues felt from their hands as they felt the surface and thus were confident of its support. Again, two subsystems – seeing and touching – are coordinated when playing with transparent containers, each system educating the other in the discovery of relevant regularities to that coupling. The changes in these component subsystems – the regularities found in one task will be transported to other tasks that also recruit these same subsystems (for another compelling example, see Bertenthal & Campos, 1990). In this way, the coordination of multimodal subsystems in specific tasks may create abstract and transportable ideas.

One might ask at this juncture: what do babies have that birds do not, that babies but not birds, so readily learn to distinguish transparent surfaces from openings in those surfaces? The correct answer is probably many, many different processes (and body parts). But in the case of transparency, one relevant difference may be the sensory–motor couplings and time-locked dynamic visual and haptic experiences that infants can generate with their hands as they manually and visually explore objects. Each sensory–motor system has its own processes and computational mechanisms. These unique takes on the world are a critical starting point to outcome and without doubt matter greatly. But what also matters is the *functional* coupling these heterogeneous processes *in multiple tasks*, such that the component systems change the internal workings of each other, finding higher-order regularities that transcend specific modalities and specific tasks (see Barsalou et al., 2003). What is unique in the human system relative to other species may not be so much any one bit of computational machinery but rather the higher-order integrative architecture, that is open to new and multiple solutions (see also Honey et al., 2007).

## An open system

Human beings learn and do things that have never been done before. Even young children exhibit cognitive skills that were not imaginable generations ago, learning about transparency, pictures, and videos, about drag and click, how to program, and the rhythm and syntax of texting, as well as those older intellectual artefacts of reading and multiplication. There seems no limit to what we can potentially incorporate, make into everyday cognition, even into child's play. In this way, human cognition is decidedly not predetermined and not preset, but is rather open-ended. This open-endedness begins (but does not end) with our ability to discover goals, through our own activity.

The lesson from rattle-shaking and sticky mittens is this: goals are not prior to the task but are themselves emergent in the infant's engagement with the world. Prior to shaking the rattle, or catching a toy with the sticky mittens, infants can have no specific goal to shake to make noise, or to swat to snatch an object. Another example of goal discovery through action is 'infant conjugate reinforcement' (Argulo–Kinzler et al., 2002; Rovee-Collier & Hayne, 1987). Infants (as young as 3 months) are placed on their backs and their ankles are attached by a ribbon to a mobile which is suspended overhead. The mobile, which produces interesting sights and sounds, provides the infant with many time-locked patterns of correlations. More importantly, infants themselves discover these relations through their own movement patterns. The faster and harder infants kick, the more the mobile moves fuelling the infants' kicking. This is a highly engaging task for infants; they smile and laugh, and often become angry when the contingency is removed. This experimental procedure, like the world, provides complex, diverse, and never exactly repeating events but critically events perfectly time-locked with the infant's own actions. It is spontaneous nontask-related movement that starts the process off by creating the opportunity for the coordination of the infant's action with that of the mobile. It is this coordination which ultimately defines the task and thus becomes the goal.

Because the goals that drive intentional action (and learning) are not pregiven but are discovered by the individual through action, development itself is highly specific to the individual. Different children will follow different developmental paths that depend on the specific tasks they encounter and the intrinsic dynamics of their own system. One elegant demonstration of the individual nature of developmental trajectories is Thelen's et al. (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, but different actions for different 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 – 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 nonreaching 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. As action defines the task and as 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. Given the constraints of the world, of human bodies, and of the heterogeneous and multimodal system out of which intelligence is made, 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.

## Doing with images

Actions create perceivable outcomes, some of which are stable and sit as new things in the world. For example, a child through a series of actions, stacks one thing onto the other, and then another, and then another, not only sees and feels the repetitive actions as they unfold but also in the end can also sit back, look at, and reflect on a tower of blocks that did not exist before. This link between actions and stable products in the world may be profoundly important to human intelligence and in particular, to the development of symbols. Alan Kay, one of the founding fathers of object-oriented programming and graphical-user interfaces (the idea behind the original MacIntosh drag-and-click we now all use) is a visionary computer scientist who gave a talk in 1987 with a cult-like following. The talk's title is 'Doing with images makes symbols'. (One can google the title of the talk to find hundreds of copies of it.) Inspired by Piaget (1952), Bruner (1956), and Vygotsky (1978), Kay proposed that abstract ideas (and symbolic thought) were built out of real-time sensory–motor interactions with images, that is, with the stable *perceivable consequences* of our own actions. This is like the idea of a closed loop in active vision in which every action creates perceivable consequences that also guide action at the next step. But the key idea here is that some perceivable consequences are special in being *image-like,* in that they are stable and enduring, a perceivable constant that is coupled to the messier context-specific and continuous dynamics of perception and action. These external stabilities – 'artefacts' that may be initially produced without plan or goal – are a profound force on human cognition.

The development of spatial classification provides an interesting phenomenon with which to consider these ideas. Between their first and third birthdays, children begin to use space to represent similarity, putting like things close together (Sugarman, 1983). Indeed, during this period, they become almost compulsive spatial sorters. Confronted with an array of four identical cars and four identical dolls, they physically group them – moving all the cars spatially close to each other and spatially apart from the groups of dolls even though there is no explicit task to do so. They are so reliable at doing this that many developmental psychologists use the task as a way to measure young children's knowledge of categories (e.g. Mandler et al., 1991; Nelson, 1973; Rakison & Butterworth, 1998). Their reasoning is that if a 2-year-old knows that two objects are the same kind of thing, she should spatially group them together. A perhaps just-as-interesting question is why the child spatially groups objects at all.

The developmental evidence suggests a progressive discovery of spatial classification. Nine- to 10-month-old infants when given sets of objects of like kinds do not systematically group them into categories. However, they do – more often than expected by chance – pick up like objects but not unlike objects, one in each hand, and bang them together (Forman, 1982). By 12 months of age these manipulations – like manipulations of like kinds – become more systematic and extended to all available like instances (Sugarman, 1983). For example, given four cars and four dolls, the child may systematically push each of the four cars. Around 18 months of age, children will not only manipulate objects from one category in sequence but also systematically manipulate in different ways objects from two different categories, for example, first pushing each of four cars, one after another and then touching each of four dolls in turn. Sometime after 24 months, the sorting seems more purposeful with all of one kind gathered to form one group and the other kind left unorganized. Later the systematic formation o two spatial groups emerges.

Sheya and Smith (Sheya, 2005; Sheya & Smith, 2008, in press) propose that this developmental pattern emerges though the child's own actions, actions that at first have no goal of creating a classification. Four behavioural tendencies are proposed to drive the process. The first is that infants reach to objects in which they are interested. The second is that infants have a tendency to repeat just-performed motor acts, and in particular to repeat reaches to nearby locations (e.g. Smith et al., 1999). The third is that perceptually similar objects are similarly enticing to infants. The fourth is that infants notice the outcomes of their own actions. These four tendencies can be understood in terms of a dynamic salience map, a map that determines where infants look to and reach next.

Imagine an array of eight toys, five of one kind and three of another as illustrated in Figure 16.4. Attention to and the touching of one toy alters the salience map, by

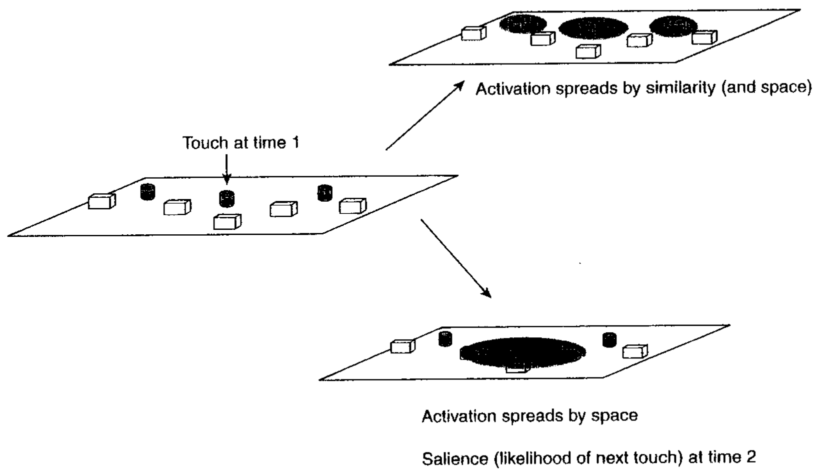

**Fig. 16.4** An illustration of the two ways in which perceptual-motor activity at one moment in time may organize attention and behaviour at the next. A touch to one objec may increase the salience of objects at locations near to that touch or a touch to one object may increase the salience of objects similar in their properties to the first object.

activating the spatial location of that toy. This activation can spread along two potential dimensions – physical space or a feature space according to the feature similarity of the objects. In their behavioural experiments, Sheya and Smith showed that activation in this salience space (as measured by the next toy touched) spreads mostly by space for younger infants (12-month-olds) but by feature similarity for older infants (18-month-olds). However, for all infants, both closeness and similarity interacted with *close* and *similar* things being more salient than close and dissimilar things.

These tendencies can create the sequential touching of like objects. As children are drawn to nearby and similar things, they are likely – through just these processes alone – to drop similar things near other, with the interactive effects of spatial proximity and physical similarity increasing the salience of reaching, again and again, to like and near things. A system whose activity is biased to both reach to similar locations and to reach to similar objects, will as consequence of reaching and dropping those things, end up with similar things near each other. It is here that Alan Kay's idea enters in. This unplanned consequence of similar things ending up near each other creates *an image*, a stable array of like things in proximity and apart from different things.

Namy et al., (1997) conducted a microgenetic study with the goal of encouraging the development of spatial classification in toddlers who did not yet spatially group like objects. The children's 'training' was a fun task of putting objects into a shape sorter. As illustrated in Figure 16.5, the shape sorter was a transparent container structured so that children could see the objects once they had been dropped inside. Children were given two different kinds of objects (e.g. blocks and bolls) that might be put into the container. The opening on the top of the shape container only allowed one type of object to fit inside the hole. Children at this age have strong perseverative tendencies to repeat the same action, and so they (quite happily) attempted to put all the objects into the container – the kind that fit and the kind that did not. But their actions led to only one kind actually being in the container together and thus in spatial proximity. Their actions thus produced a stable image of like things being near each other and apart from different things.

This experience turned these children in to spatial classifiers, advancing them several months in this developmental progression. In the transfer task, children were given sets of eight objects – four of one kind or four of another and no shaper sorter. Children who had previously sorted with the transparent shape sorter – that enabled the children to see the product of their activity – were much more likely to sort the objects into spatially organized groups than children in the control conditions. Namy et al. (1997) suggest that their experimental procedure intensified the experiences characteristic of children's every day activities. Children are attracted to objects that are physically similar, they tend to repeat similar acts and this leads to similar things often ending up close in space. As children play with things in their world, a world in which physically like things have similar propensities, like things will end up near each other. Moreover, as children interact with a world structured by adults with a similar psychology, they will encounter cups near cups on the shelves and socks all together in the sock drawer. In these ways, spatial proximity becomes *the* foundational metaphor for both informal ideas and formal mathematical theories about similarity.

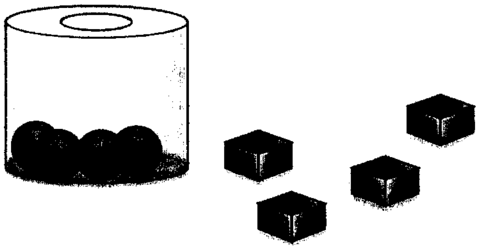

**Fig. 16.5** Depicts the transparent sorter used in Namy et al. (1998). The sorter enabled young children to spatially segregate objects of different kinds.

Many (though not all) of the spatial classifications that children encounter and do in the world – the cups near the cups, the socks in the sock drawer – are man-made creations, artefacts. And they are everywhere, at least in our Western world. But it seems likely that if some child never saw a spatial classification of like-near-like that they would spontaneously produce it nonetheless. Spatial classification is an artefact that is likely to spontaneously emerge over and over again, even if unsupported by convention. This is because spatial classification is encouraged by the physical fact that like things have like propensities and the psychological fact of an attentional and action system governed by both nearness in space and similarity in features. Not all artefacts that shape human cognition are so easily or so readily re-invented by each of us. Instead, they are passed down (and incrementally adapted) across generations within a culture. In our culture, children grow up with several powerful systems of artefacts – tools, letters, books, stories, machines. In all human cultures, children grow up with language. These systems may literally make thought more computationally powerful (see Clark, 1998; 2003; Greenspan & Shanker, 2004; Premack, 2004; Vygotsky, 1978; Wertsch, 1985). Relevant examples that have been discussed by others include how the invention of zero made everyday arithmetic easy (Seife, 2000), how learning to read an alphabetic system changes how we perceive and process spoken sounds (e.g. Treiman & Kessler, 2007), how social interactions with tools potentiate the invention and use of tools (e.g. Call & Tomasello, 1994). The focus in the next section is how one shared symbol system – language – may empower intelligence.

## Doing with symbols

Children grow in a sea of symbols that include spoken words, marks on paper, and gestures. These physical and thus perceivable symbols are a very special form of regularity with four characteristics potentially crucial to the 'specialness' of human cognition. First, symbol systems (including language and mathematics) are in-the-world regularities that are *shared* (e.g. Freyd, 1983; Hutchins & Hazelhurst, 1995). This shared aspect means that these systems are very stable, continually constrained by many local communicative acts. Second, symbol systems are *arbitrary*. For example, the form of the word *dog* gives us no hints about the kinds of thing to which it refers and there is nothing in the similarity of the forms of *dig* and *dog* that conveys a similarity in meaning. It is interesting to ask *why* language and other symbol systems are this way, and why invented symbol systems that often begin with strong iconicity between form and referent become increasingly noniconic. One might expect that a multimodal, grounded, sensorimotor sort of learning would favour a more iconic, pantomime-like language in which symbols were similar to referents. But symbol systems are decidedly not like this (see DeLoache, 2002). Third, the distinct physical tokens that comprise effective symbol systems are rapidly, automatically, and *categorically perceived*. For some such systems (e.g. spoken words) this may be because of evolved sensory systems. The shared and repeated use of symbols may also encourage them to take on easily perceived and produced forms (e.g. Christiansen & Kirby, 2003; Ohala, 1993). Finally, some symbol systems (e.g. orthographies, mathematics) are made easy to perceive through extensive training. In sum, effective symbol systems require that the symbols themselves not take too much cognitive processing so that they can be mere pointers to meaning. Finally, symbols *refer*. Symbols systems are not about themselves but are about their referents. These properties of symbols confer computational power through re-representation, by creating and enabling higher-order correlations, and by creating dynamically complex systems that can jump from one highly stable state to another.

## Re-representation through replacement

The physical symbols that surround children – spoken words, letters, drawings – are dual entities: (1) physical things with their own properties and (2) entities which stand for other things. Operating on physical symbols as physical entities may empower cognition, simply as a consequence of replacing one entity (the referred to thing) with another (the physical symbol (see Son et al., 2008, for further discussion and evidence). This point can be illustrated by considering what we mean when we say that some thing is the 'same.' The problem with 'same' is that one can be talking about *any* kind of thing and *any* kind of similarity – how two books are alike, how whales are sort of the same as fish but not really, how two smells are alike, how the relative quality of two cars are the same, and so forth. Explaining this ability might actually require some underlying and innately given machinery that can compute sameness over anything (what some might call a concept, see Smith, 1993; Smith et al., 1997 for a discussion of this possible machinery).

The present concern, however, is how symbols as physical entities might leverage thought and cognitive development. A task invented by Premack (1976) to measure the transcendent and abstract nature of sameness judgements provides an interesting case. Understanding Premack's problem begins with a simpler problem, matching to sample. In matching-to-sample, the learner must find the two objects in a set of three that are the same, no matter what those two objects are. Various animal species can do this (see Cook & Wasserman, 2006). These judgements suggest an 'abstract concept of same', although such a conclusion from matching-to-sample tasks alone is highly controversial as success can also be explained in other more perceptual ways.

Premack's problem takes this matching-to-sample task a level higher. Table 16.1 shows the required responses. Given the joint presentation of the first two cards, the correct response is 'same'; given the joint presentation of the next two cards, the correct response is again 'same'; the next two cards when jointly presented require the response 'different' and the final two cards when jointly presented require the response different. The rule is that they are the 'same' when the *relation* in the two cards is the same. This is not easily learnt by children, nor most animals.

**Table 16.1** Examples of the kinds of trials that make up Premack's problem

| Presentation | Cards | Correct response |
|---|---|---|
| Card 1 | X-X | |
| Card 2 | Y-Y | |
| | | Same |
| Card 1 | X-Y | |
| Card 2 | O-R | |
| | | Same |
| Card 1 | X-X | |
| Card 2 | O-R | |
| | | Different |
| Card 1 | X-R | |
| Card 2 | O-O | |
| | | Different |

Chimpanzees can be taught to do this (see also, Oden, Thompson, & Premack, 1990). The method begins by first teaching them to label pairs of things that stood in an identity relation with one arbitrary physical token, such as a heart-token for same (e.g. ♥ → AA) and to label pairs of things that were different with another token (e.g. # → AB). Having learnt this, the chimpanzees can make second-order matches, judging AA to be related to BB. The informative idea is this Chimpanzees could potentially do this task, not by abstracting the relation of sameness across instances, but by knowing that AA→ ♥ and BB → ♥. Then they can simply respond to the sameness of ♥ and ♥. Notice that this second-order relation (the same relation between 'same' and 'same') can result from *direct computations over the labels*. By replacing AA with a heart-token and BB with a heart-token, these chimpanzees can reduce the problem to matching-to-sample. In this way, symbols can enable solutions to what might seem intractable problems, enabling manipulations of concrete entities (the symbols). Recent research on learning mathematical symbol systems such as algebra, suggest that even highly advanced forms of human intelligence might be tied to perceptual (and indeed motor) aspects of manipulating physical symbols as entities in the world (Landy & Goldstone, 2007).

## Higher-order correlations

Symbols may not just replace their referents in computations but they may also enable the discovery of higher-order regularities (Colunga & Smith, 2005; Smith & Gasser, 2005). Young children's progressively more rapid learning of object names provides a case in point. Children's learning of such common names such as *chair*, *dog*, and *milk* begins by their learning the specific similarities that organize these specific categories. Considerable evidence suggests that this initial category-by-category learning *creates* higher-level knowledge that enables children to rapidly and correctly map a name to *the whole category* given just one single instance of that category (see Colunga & Smith, 2005, for review). Indeed, by the time children are 3 years of age, they generalize the names for novel artefacts by shape, names for novel animals by multiple similarities, and names for substances by material (Smith et al., 2002a).

Experimental and computational studies suggest four steps in children's learning. These are illustrated in Figure 16.6 (Smith et al., 2002b) which shows just one of the regularities that children learn – that artefact categories are organized by shape. Step 1 in the learning process is the mapping of names to objects – the name 'ball' to a particular ball and the name 'cup' to a particular cup, for example. This is done multiple times for each name as the child encounters multiple examples. And importantly, in the early lexicon, solid, rigidly shaped things are in categories typically well-organized by similarity in shape (Samuelson & Smith, 1999). This learning of individual names sets up step 2 – first-order generalizations about the structure of individual categories, that is, the knowledge that balls are round and cups are cup-shaped. The first-order generalization should enable the learner to recognize novel balls and cups.

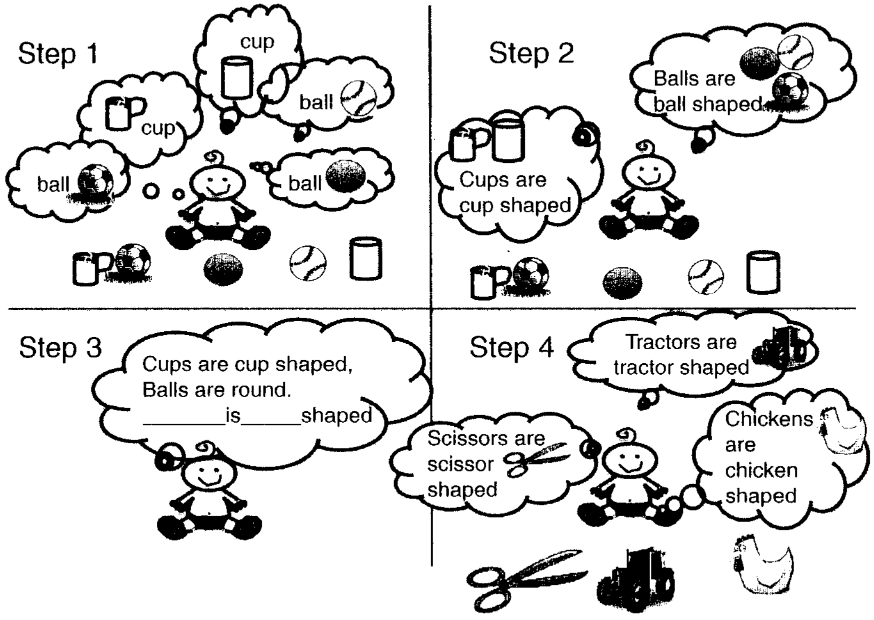

**Fig. 16.6** Steps in forming a higher-order generalizations: (1) learn specific word object associations; (2) make first-order generalizations for each of the specific categories; (3) make higher-order generalizations across the regularities discovered in those first-order generalizations; and (4) use these higher-order generalizations to rapidly learn new categories.

But the key to the rapid learning of *new* words is the formation of a higher-order generalization. As most of the solid and rigid things that children learn about are named by their shape, children may also learn the second-order generalization that names for artefacts (solid, rigid things) in general span categories of similar-shaped things. As illustrated in step 3 of the figure, this second-order generalization requires generalizations over specific names and specific category structures. Once such a higher-order generalization is formed, the child behaves as if it has an abstract and variablized rule: for any artefact, whatever its individual properties or individual shape, form a category by a shape. Step 4 illustrates the potential developmental consequence of this higher-order generalization – attention to the right property, shape – for learning new names for artefacts. The plausibility of this account has been demonstrated in experimental studies that effectively accelerate the vocabulary-acquisition function by teaching children the relevant correlations (Smith et al., 2002b) and in simulation studies with neural nets (Colunga, 2003). Although we describe the process here in step-by-step fashion to illustrate logic of what is learned, the system is actually and continuously making multiple mappings and this may be important to answering the question of just what enables a system to build higher-order, almost rule-like generalizations (see Colunga & Smith, 2005).

Critically, it may be symbols, or more correctly a learning process that makes generalizations over a *large* system of associations in which one class of associates are arbitrary and orthogonal to each other (that is having the properties of words) that makes these higher order generalizations possible. Words – in contrast to other kinds of associates of objects (e.g. the actions we perform on them, or their typical locations) are special because they are arbitrary and orthogonal. As Colunga (2003) has argued, arbitrary and nonoverlapping symbols will pull apart first-order categories in ways critical to forming the second-order generalizations (step 3) by connectionist networks. Indeed, her simulations indicate that networks that readily form second-order generalizations and yield accelerating rates of vocabulary acquisition do not do this if the labels, the words, are not orthogonal.

A second property of words that may also be critical to the formation of higher-order generalizations is that words present systems of correlations *among themselves* not just to things in the world. Yoshida and Smith (2005) demonstrated the potential importance of these word–word associations in a study that attempted to teach young children the higher-order generalizations that solid and artefact-like things are named by shape and that nonsolid substances are named by material. They attempted to teach this to Japanese-speaking children *at an age earlier* than when these children normally make this distinction (see Yoshida & Smith, 2003). They did this by intensively teaching children names for novel solid things in categories well-organized by shapes and names for novel nonsolid things in categories well organized by material, and they did this in two ways: without correlating linguistic cues or *with* correlating linguistic cues (roughly like count–mass determiners in English for which there are no counterparts in Japanese). The main result is that only children taught the correlations between the perceptual cues of solid → shape-based and nonsolid → material based *in the context of redundant linguistic cues that linguistically marked the novel nouns as being of two kinds* made the higher-order generalizations enabling generalizations of learning in transfer to the formation of novel lexical categories for novel solid and nonsolid things.

We know from corpus analyses and other forms of data-mining that there is considerable derivable latent structure in the co-occurrence probabilities and associations among words – that is the physical symbols – themselves. Importantly, word–word associations are *ungrounded* in the sense that associations and co-occurrence relations exist among symbols that can be considered in isolation from their referents and meaning. Considerable structure is derivable from the statistical regularites among the words themselves including phonological, semantic, and syntactic categories (Chater & Manning, 2006; Landauer & Dumais, 1997; Li et al., 2000; Mintz et al., 2002; Monaghan et al., 2005; Redington et al., 1998; Steyvers & Tenenbaum, 2005).

As illustrated in Figure 16.7, the human cognitive system has three kinds of very different but connected data sets out of which to build knowledge: (1) correlations among the multimodal sensory–motor processes time-locked to each other and the world in tasks; (2) statistical regularities – and the latent structure derivable from those regularities – that characterize the symbols themselves; and (3) then the regularities with which (some of) these symbols map to sensory–motor experiences.

## Analogue and digital

Degenerate, integrative, and graded sensory–motor processes are inherently inventive and adaptive, not brittle. As many theorists have pointed out (e.g. Barsalou & Prinz, 1997; Clark, 2003; Pfeiffer & Scheier, 1999) these qualities of the human perception–action system lead to an intelligence that is very different from the standard digital computer, precisely because of the adaptive, graded, and blended nature of sensory–motor representations. Digital computers that route arbitrary symbol structures between memory structures and a central processor fail to capture the more subtle, complex, and graded regularities that comprise everyday intelligence. However, digital systems have some decided advantages. O'Reilly (2006) notes in a recent review in *Science* that digital systems have the advantage of being able to rapidly move from one highly stable solution to a very different but also stable one, without blending or passing through intermediate states (see also Clearfield et al., 2009; Hanania & Smith, in press; Munakata, 2001; Rougier et al., 2005).

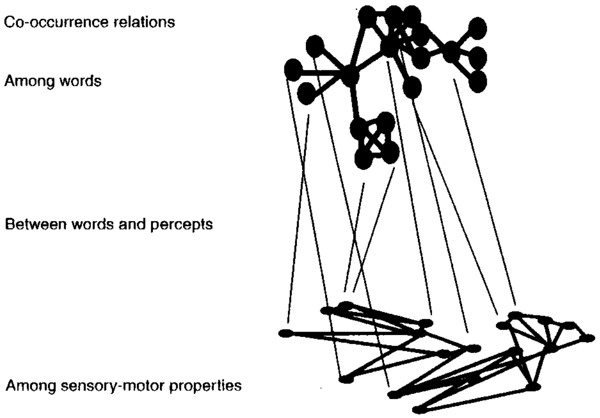

**Fig. 16.7** A schematic illustration of three distinct kinds of associative systems critical to human cognition: associations (co-occurrence regularities) among the symbols themselves; associations among perceivable properties of things in the world; and associations from symbols to perceivable properties in the world.

One possibility is that language helps creates a digital-like function on top of the more graded and continuous and common sense way of knowing, giving human cognition two complementary ways of solving problems. One domain in which there is both evidence and theoretical models to support these ideas is the development of selective attention and attention shifting (Hanania & Smith, in press; Morton & Munakata, 2007; O'Reilly, 2006; Rougier et al., 2005). The key theoretical idea is that sufficient experience with words creates near-discrete and nondistributed activation patterns and these are necessary to bistable systems that exhibit 'rule-like' all-or-none gating of goals. Two different and independent simulations of these processes, each motivated by somewhat different concerns and data, came to the same conclusion (Rougier et al., 2005; Smith et al., 1997; see also Hanania & Smith, in press).

In the Rougier et al. simulations, the rapid shifting of attention in an all-or-none manner from one dimension to another required highly selective patterns of activation that corresponded to the input on one dimension independent of the other dimension. In these simulations, Rougier et al. (2005) specifically attempted to model the core properties thought to characterize the prefrontal cortex (PFC) and executive control: (1) recurrent and thus stabilizing patterns of activation; (2) an adaptive gating mechanism which maintains a rule so long as the outcome is predicted but that destabilizes the rule when the predictability of the outcome changes; and (3) modulation of processing in other areas, thus enabling the actively maintained rule to orchestrate and capture supporting processes in other systems (the so-called executive function of the PFC). The network was trained in three overlapping tasks: (1) labelling attributes on single dimensions (red, blue); (2) making judgements of sameness and difference on a single dimension (same colour, same shape); and (3) making comparisons on a single dimension (bigger, littler). Networks trained in all three of these tasks (but not just any one of them) developed highly abstract and orthogonal patterns of activation, not just to individual attributes (red, blue) but sameness along a single dimension independent of the particular attributes on that dimension (same colour, same shape). This result replicates a previous recurrent connectionist model by Smith et al. (1997) which also showed that highly abstract dimensional representations required training on both labelling (or selective responding to) attributes *and* same–different judgements on the dimension across a variety of different attributes and in the context of irrelevant variation on a different dimension. Both sets of simulations highlight two important contributions to the developmental processes that create the specialness of human cognition – overlapping tasks and language.

## Doing with people

The highest forms of biological intelligence reside in a social world; development takes place among conspecifics with similar internal systems and similar external bodies. The importance of the social embeddedness of human cognition is well recognized in

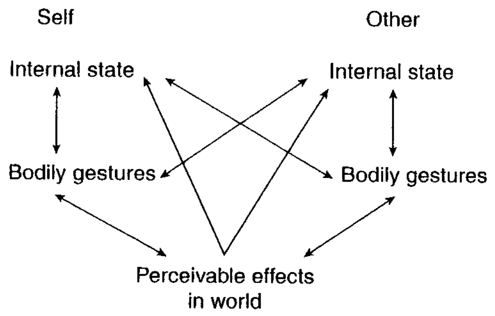

**Fig. 16.8** The system of correlations added by being a social being in a social world. Coupled regularities between internal states, bodily gestures, and their perceivable effects in the world within and across individuals.

the literature (e.g. Markova & Legerstee, 2006; Rogoff, 2003; Striano & Reid, 2006). Perhaps less well recognized (but see, Smith, 2000a, 2000b; Yu et al., 2005) is how this social embeddedness is itself made manifest in sensory–motor coordinations (see Pereira et al., 2008; Smith & Breazeal, 2007).

Figure 16.8 illustrates the couplings between one's own internal states, one's outward bodily behaviours, the bodily behaviours of others, and their internal states. Crucially, the body and its behaviours are observable by others. But because observable behaviours are also linked to the actor's internal state, observable bodily behaviours also provide (albeit imperfect) information to others about those internal states. As one's bodily actions also influence the internal states of others, one's own actions are also (albeit indirectly) linked to the internal states of others. In brief, development occurs within a complex system of coupled behaviours, coupled bodies, and coupled cognitive systems.

These couplings generate a network of learnable correlations: between the appearance of the self and the appearance of others, between the behaviour of the self and the behaviour of others, between one's own bodily behaviours and one's internal states, between the external states of others and one own's internal states. In ways deeply analogous to the interactions of perception and action as illustrated in Figure 16.1, these correlations can yield transcendent higher-order regularities, in this case, about intentions and motivations (see Smith, 2000b; Smith & Breazeal, 2007). The dynamic socially embedded coupling of two intelligent systems – to each other through similar bodies and behaviours – is a potent agent of change and, is indeed, one component of developmental process that makes humans special.

## What is special about human intelligence?

The task the contributors to this book were given by the editors was to specify what, if anything, is unique about human concepts and the role of evolution and development in this uniqueness. From the present perspective, this is an ill-formed question since 'concept' is an ill-defined theoretical construct that may or may not be needed in the end to explain human behaviour. This chapter began with the newer idea that intelligence does not reside in cognitive processes that are separate and distinct from the real-time sensory–motor processes which actively engage the world but that, instead, human intelligence can only be understood in terms of a complex system of brain, body, and world. A broad overview of what we know about the *making* of intelligent behaviour in human development from this perspective suggests the following:

- The open and inventive character of human cognition is made through a protracted period of immaturity in which children perceive and act in the world in the service of many different overlapping and discovered tasks.
- This developmental process (like evolution) is local, individualistic, and opportunistic and, therefore, highly creative.
- Many heterogeneous processes time-locked to each other and to the world change their internal dynamics (and thus the computations they can perform) as a consequence of these couplings.
- Many different tasks create different overlapping soft assemblies of these different component processes and cascading effects of learning from one task to others.
- Many different overlapping soft-assemblies NS integrations across many different tasks create abstract processes that transcend the specifics of specific modalities and the hear-and-now of specific tasks.
- Each action creates a new opportunity for perception and learning and some actions also create perceivable and stable images.
- A learning context with such self-created and also culturally provided images, artefacts, and symbols powers the cognitive system through re-representation, the creation of latent structure, and internal patterns of activation with multiple (discrete-like) stabilities.
- A learning context filled with mature conspecifics, with similar bodies and similar cognitive systems, enables systems of correlations across self and other, and ultimately creates predictions about the intentions and attentional states of others.

In brief, there is no simple answer to what makes human intelligence what it is, no one silver bullet, no one special thing, that makes human cognition. We are just kidding ourselves if we offer simple answers to these questions.

Finally, what are the roles of developmental process and evolution in all this? The bullet points above describe *developmental process*. It is complex, multilevelled, multicomponent, and in its entirety, it makes human intelligence what it is, and what it *can be*. Evolution selected this developmental process at all its levels of analysis from molecules to behaviour and including the multimodal sensory and motor systems (each with its own intrinsic dynamics and biases), the body's morphology, the possibilities for actions, the affiliative ties to conspecifics, and an integrative cognitive architecture that dynamically couples and uncouples components in functional tasks. In this complex system and in this developmental process, there will be many continuities with other organisms – no one magical difference anywhere – but the whole, and what it can create through its own activity, is uniquely human.

## References

Angulo-Kinzler, R. M., Ulrich, B., & Thelen, E. (2002). Three-month-old infants can select specific leg motor solutions. *Motor Control*, **6**, 52–68.

Anderson, M. (2003). Embodied cognition: A field guide. *Artificial Intelligence*, **149**, 91–130.

Ballard, D., Hayhoe, M., Pook, P., & Rao, R. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, **20**, 723–767.

Barsalou, L. W. (1993). Challenging assumptions about concepts. *Cognitive Development*, **8**(2), 169–180.

Barsalou, L. W. & Prinz, J. J. (1997). Mundane creativity in perceptual symbol systems. In T. B. Ward, S. M. Smith, & J. Vaid (Eds.), *Creative Thought: An Investigation of Conceptual Structures and Processes*, pp. 267–307, Washington, DC: American Psychological Association.

Barsalou, L. W., Pecher, D., Zeelenberg, R., Simmons, W. K., & Hamann, S. B. (2005). Multimodal simulation in conceptual processing. In W. Ahn, R. L. Goldstone, B. C. Love, A. Markman, & P. Wolff (Eds.), *Categorization Inside and Outside the Lab: Festschrift in Honor of Douglas L. Medin*, pp. 249–270, Washington, DC: American Psychological Association.

Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, **7**, 84–91.

Barsalou, L., Breazeal, C., & Smith, L. B. (2007). Cognition as coordinated noncognition. *Cognitive Processing*, **8**, 79–91.

Beer, R. D. (2000). Dynamical approaches to cognitive science. *Trends in Cognitive Science*, **4**, 91–99.

Bertenthal, B. I. & Campos, J. J. (1990). A systems approach to the organizing effects of self-produced locomotion during infancy. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in Infancy Research*, Vol. 6, pp. 1–60, Norwood, NJ: Ablex.

Brooks, R. A. (1991). 'Intelligence Without Representation'. *Artificial Intelligence Journal*, **47**, 139–159.

Bruner, J. S. (1996). *The Culture of Education* Cambridge, MA: Harvard University Press.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking* Oxford, England: John Wiley and Sons.

Call, J. & Tomasello, M. (1994). The social learning of tool use by orangutans (*Pongo pygmaeus*). *Human Evolution* **9**, 297–313.

Chater, N. & Manning, C. D. (2006). Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences. Special Issue: Probabilistic Models of Cognition*, **10**(7), 335–344.

Christiansen, M. &. Kirby, S (2003). *Language Evolution* Oxford: Oxford University Press.

Clark, A. (1998). Magic words: How language augments human computation. In P. Carruthers & J. Boucher (Eds.), *Language and Thought: Interdisciplinary Themes*, pp. 162–183, Cambridge: Cambridge University Press.

Clark, A. (2003). *Natural-Born Cyborgs: Minds, Technologies and the Future of Human Intelligence*, New York: Oxford University Press.

Clearfield, M. W., Dineva, E., Smith, L. B., Diedrich, F. J., & Thelen, E. (2009). Cue salience and infant perseverative reaching: tests of the dynamic field theory. *Developmental Science*, **12**(1), 26–30.

Colunga, E. & Smith, L. B. (2005). From the lexicon to expectations about kinds: A role for associative learning. *Psychological Review*, **112**(2), 342–382.

Cook, R. G. & Wasserman, E. A. (2006). *Relational Discrimination Learning in Pigeons* New York: Oxford University Press.

DeLoache, J. (2002). The symbol-mindedness of young children. In W. W. Hartup & R. A. Weinberg (Eds.), *Child Psychology in Retrospect and Prospect: The Minnesota Symposia on Child Psychology*, Vol. 32, pp. 73–101, Mahwah, NJ: LEA Publishing.

Diamond, A. (1990). Developmental time course in human infants and infant monkeys and the neural bases of inhibitory control in reaching. In A. Diamond (Ed.) *The Development and Neural Bases of Higher Cognitive Functions*, pp. 637–676, New York: New York Academy of Sciences.

Edelman, G. (1987). *Neural Darwinism* New York: Basic Books.

Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Science*, **8**, 301–306.

Forman, G. (1982). A search for the origins of equivalence concepts through a microanalysis of block play. In G. Forman (Ed.), *Action and Thought: From Sensorimotor Schemes to Symbolic Operations*, pp. 97–135, New York: Academic Press.

Freyd, J. (1983). Shareability: The social psychology of epistemology. *Cognitive Science*, 7(3), 191–210.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155–170.

Greenspan, S. I. & Shanker, S. G. (2004). *The First Idea: How Symbols, Language and Intelligence Evolved from Our Primate Ancestors to Modern Humans* New York: Da Capo Press.

Gibson, E. J. & Walk, R. D. (1960). The 'visual cliff'. *Scientific American*, **202**(4), 64–71.

González, J. C., Bach-y-Rita, P., & Haase, S. J. (2005). Perceptual recalibration in sensory substitution and perceptual modification. *Pragmatics & Cognition. Special Issue: Cognitive Technologies and the Pragmatics of Cognition*, **13**(3), 481–500.

Hanania, R. & Smith, L. B. (in press). Selective attention and attention switching: towards a unified developmental approach. *Developmental Science*, in press.

Harman, K. L., Humphrey, G. K., & Goodale, M. A. (1999). Active manual control of object views facilitates visual recognition. *Current Biology*, **9**(22), 1315–1318.

Hein, A. & Diamond, R. M. (1972). Locomotory space as a prerequisite for acquiring visually guided reaching in kittens. *Journal of Comparative and Physiological Psychology*, **81**(3), 394–398.

Held, R. & Hein, A. (1963). Movement-produced stimulation in the development of visually guided behavior. *Journal of Comparative and Physiological Psychology*, **56**(5), 872–876.

Honey, C. Kotter, R., Breakspear, M., & Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Science*, **104**, 10240–10245.

Hutchins, E. (1995). *Cognition in the Wild* Cambridge, MA: MIT Press.

Hutchins, E. & Hazelhurst, B. (1995). *How to Invent a Lexicon: The Development of Shared Symbols in Interaction* London: UCL.

James, W. (1890). *The Principles of Psychology*, Vol. 1, New York: Holt.

Keil, F. C. (1994). The birth and nurturance of concepts by domains: The origins of concepts of living things. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the Mind: Domain*

*Specificity in Cognition and Culture*, pp. 234–254, New York: Cambridge University Press.

Landau, B. & Gleitman, L. (1985). *Language and Experience* Cambridge, MA: Harvard University Press.

Landauer, T. K. & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, **104**(2), 211–240.

Landrigan, D. T. & Forsyth, G. A. (1974). Regulation and production of movement effects in exploration–recognition performance. *Journal of Experimental Psychology*, **103**(6), 1124–1130.

Landy, D. & Goldstone, R. L. (2007). Formal notations are diagrams: Evidence from a production task. *Memory & Cognition*, 35, 2033–2040.

Li, P., Burgess, C., & Lund, K. (2000). The acquisition of word meaning through global lexical co-occurrences. In E. V. Clark (Ed.), *Proceedings of the 30th Annual Child Language Research Forum*, pp.167–178, Stanford, CA: Center for the Study of Language and Information.

Lungarella, M. & Sporns, O. (2006). Mapping information flow in sensorimotor networks. *PLoS Computational Biology*, **2**(10): e144–150.

Lungarella, M., Pegors, T., Bulwinkle, D., & Sporns, O. (2005). Methods for quantifying the informational structure of sensory and motor data. *Neuroinformatics*, 3, 243–262.

Malakoff, D. (2004). Clear and present danger. *Audubon*, **106**(1), 65–68.

Mandler, J. M., Bauer, P. J., & McDonough, L. (1991). Separating the sheep from the goats: Differentiating global categories. *Cognitive Psychology*, **23**(2), 263–298.

Markova, G. & Legerstee, M. (2006). Contingency, imitation, and affect sharing: Foundations of infants' social awareness. *Developmental Psychology*, **42**(1), 132–141.

Martin, A. & Chao, L. L. (2001). Semantic memory and the brain: Structure and processes. *Current Opinion in Neurobiology Special Issue: Cognitive Neuroscience*, **11**(2), 194–201.

McIntosh, A. R., Fitzpatrick, S. M., & Friston, K. J. (2001). On the marriage of cognition and neuroscience. *Neuroimage*, **14**, 1231–1237.

Mintz, T. H., Newport, E. L., & Bever, T. G. (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive Science: A Multidisciplinary Journal*, **26**(4), 393–424.

Monaghan, P., Christiansen, M. H., & Chater, N. (2007). The phonological–distributional coherence hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive Psychology*, 55(4), 259–305.

Morton, J. B. & Munakata, Y. (2002). Active versus latent representations: A neural network model of perseveration, dissociation, and decalage. *Developmental Psychobiology. Special Issue: Converging Method Approach to the Study of Developmental Science*, **40**(3), 255–265.

Munakata, Y. (2001). Graded representations in behavioral dissociations. *Trends in Cognitive Sciences*, 5(7), 309–315.

Namy, L. L., Smith, L. B., & Gershkoff-Stowe, L. (1997). Young children discovery of spatial classification. *Cognitive Development*, **12**(2), 163–184.

Needham, A., Barrett, T., & Peterman, K. (2002). A pick me up for infants' exploratory skills: Early simulated experiences reaching for objects using 'sticky' mittens enhances young infants' object exploration skills. *Infant Behavior and Development*, 25(3), 279–295.

Nelson, K. (1973). Some evidence for the cognitive primacy of categorization and its functional basis. *Merrill Palmer Quarterly*, **19**(1), 21–39.

Oden, D. L., Thompson, R. K., & Premack, D. (1990). Infant chimpanzees spontaneously perceive both concrete and abstract same/different relations. *Child Development*, **61**(3), 621–631.

Ohala, J. J. (1993). Sound change as nature's speech perception experiment. *Speech Communication*, **13**, 155–161.

O'Regan, J. K. & Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, **24**, 939–1031.

O'Reilly, R. C. (2006). Biologically based computational models of high-level cognition. *Science*, **314**(5796), 91–94.

Pereira, A., Smith, L. B., & Yu, C. (2008) Social coordination in toddler's word learning: Interacting systems of perception and action. *Connection Science*, **23**, 73–89.

Pfeifer, R. & Scheier, C. (1999). *Understanding Intelligence* Cambridge, MA: MIT Press.

Piaget, J. (1952). *The Origins of Intelligence in Children* Oxford: International Universities Press.

Port, R. & Van Gelder, T. (1995) *Mind as Motion* Cambridge, MA: MIT Press.

Premack, D. (1976). *Intelligence in Ape and Man* Oxford: Lawrence Erlbaum.

Premack, D. (2004). Is language the key to human intelligence? *Science*, **303**(5656), 318–320.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, **22**(2), 253–336.

Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience*, **21**(3), 793–797.

Rakison, D. H. & Butterworth, G. E. (1998). Infants' attention to object structure in early categorization. *Developmental Psychology*, **34**(6), 1310–1325.

Redington, M., Chater, N., & Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science: A Multidisciplinary Journal*, **22**(4), 425–469.

Reeke, G. & Edelman, G. (1984). Selective networks and recognition automata. *Annals of the New York Academy of Sciences*, **426**, 181–201.

Rogers, T., Patterson, K., & Graham, K. (2007). Colour knowledge in semantic dementia: It is not all black and white. *Neuropsychologica*, **45**, 3285–3298.

Rogoff, B. (2003). *The Cultural Nature of Human Development* New York: Oxford University Press.

Rougier, N. P., Noelle, D. C., Braver, T. S., Cohen, J. D., & O'Reilly, R. C. (2005). Prefrontal cortex and flexible cognitive control: Rules without symbols. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(20), 7338–7343.

Rovee-Collier, C. & Hayne, H. (1987). Reactivation of infant memory: Implications for cognitive development. In H. Reese (Ed.), *Advances in Child Development and Behavior*, Vol. 20, pp. 185–238, San Diego, CA: Academic Press, Inc.

Rozin, P. (1976). The evolution of intelligence and access to the cognitive unconscious. In J. M. Sprague & A. N. Epstein (Eds.), *Progress in Psychobiology and Physiological Psychology*, pp. 245–277, New York: Academic Press.

Samuelson, L. (2002). Statistical regularities in vocabulary guide language acquisition in connectionist models and 15–20-month olds. *Developmental Psychology*, **38**, 1011–1037.

Samuelson, L. K. & Smith, L. B. (1999). Early noun vocabularies: Do ontology, category structure and syntax correspond? *Cognition*, 73(1), 1–33.

Samuelson, L. & Smith, L. B. (2000). Grounding development in cognitive process. *Child Development*, 71, 98–106.

Seife, C. (2000). *Zero: The Biography of a Dangerous Idea* New York: Penguin.

Sheya, A. (2005, April). The origins of classification: Proximity, perseveration, and similarity. *Presented at the Biennial Meeting of the Society for Research on Child Development*, Atlanta, GA.

Sheya, A. & Smith, L. B. (in press). Changing priority maps in 12- to 18-month olds: An emerging role for object properties. *Psychonomic Bulletin & Review.*

Smith, L. B. (1993). The concept of same. *Advances in Child Development and Behavior*, **24**, 216–253.

Smith, L. B. (2000a). How to learn words: An associative crane. In R. Golinkoff & K. Hirsh-Pasek (Eds.), *Breaking the Word Learning Barrier*, pp. 51–80, Oxford: Oxford University Press.

Smith, L. B. (2000b). Avoiding association when its behaviorism you really hate. In R. Golinkoff & K. Hirsh-Pasek (Eds.), *Breaking the Word Learning Barrier*, pp. 169–174, Oxford: Oxford University Press.

Smith, L. B. & Breazeal, C. (2007) The dynamic lift of developmental process. *Developmental Science*, **10**, 61–68.

Smith, L. B. & Gasser, M. (2005). The development of embodied cognition: Six lessons from babies. *Artificial Life*, **11**, 13–30.

Smith, L. B. & Katz, D. B. (1996) Activity-dependent processes in perceptual and cognitive development. In R Gelman & T. Au (Eds.), *Handbook of Perception and Cognition*, Vol. 13, pp. 413–445, San Francisco, CA: Academic Press.

Smith, L.B., Gasser, M., & Sandhofer, C. (1997). Learning to talk about the properties of objects: A network model of the development of dimensions. *The Psychology of Learning and Motivation*, **36**, 220–256.

Smith, L. B., Thelen, E., Titzer, R., & McLin, D. (1999). Knowing in the context of acting: The task dynamics of the A-not-B error. *Psychological Review*, **106**(2), 235–260.

Smith, L. B., Colunga, E., & Yoshida, H. (2002a). Making an ontology: Cross-linguistic evidence. In L.M. Oakes & D. H. Rakison (Eds), *Early Category and Concept Development: Making Sense of the Blooming Buzzing Confusion*, pp. 275–302, New York, NY: Cambridge University Press.

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, S. (2002) Early noun learning provides on-the-job training for attention. *Psychological Science*, **13**, 13–19.

Son, J. Y., Smith, L. B., & Goldstone, R. L. (2008). Simplicity and generalization: Short-cutting abstraction in children's object categorizations. *Cognition*, **108**, 626–638.

Spencer, J. P. & Schutte, A. R. (2004). Unifying representations and responses: Perseverative biases arise from a single behavioral system. *Psychological Science*, **15**, 187–193.

Spivey, M. (2006). *The Continuity of Mind* New York: Oxford University Press.

Sporns, O. (2002). Networks analysis, complexity, and brain function. *Complexity*, **8**, 56–60.

Steyvers, M. & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, **29**(1), 41–78.

Stiles, J. (2008). *The Fundamentals of Brain Development* Cambridge, MA: Harvard University Press.

Striano, T. & Reid, V. M. (2006). Social cognition in the first year. *Trends in Cognitve Sciences*, **10**(10), 471–476.

Sugarman, S. (1983). *Children's Early Thought: Developments in Classification* New York: Cambridge University Press.

Thelen, E. & Smith. L. B. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action* Cambridge, MA: MIT Press.

Thelen, E., Corbetta, D., Kamm, K., Spencer, J., Schneider, K., & Zernicke, R.F. (1993). The transition to reaching: Matching intention and intrinsic dynamics. *Child Development*, **64**, 1058–1098.

Titzer, R. C. (1997). Infants' understanding of transparency: A reinterpretation of studies using the object retrieval task and visual cliff. *Dissertation Abstracts International*, **58**, 1570–1586.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition* Cambridge, MA: Harvard University Press.

Tomasello, M., Kruger, A., & Ratner, H. (1993). Cultural learning. *Behavioral and Brain Sciences*, **16**, 495–552.

Tononi, G. (2004). An information integration theory of consciousness. *BioMed Central Neuroscience*, **5**, 42.

Treiman, R. & Kessler, B. (2007). Learning to read. In M. G. Gaskell (Ed.), *Oxford Handbook of Psycholinguistics*, pp. 657–666, Oxford: Oxford University Press.

Vygotsky, L. (1978). *Mind in Society* Cambridge, MA: Harvard University Press.

Walk, R. D. & Gibson, E. J. (1961). A comparative and analytical study of visual depth perception. *Psychological Monographs*, 75(15), 1–44.

Wertsch, J. V. (1985). *Culture, Communication and Cognition: Vygotskian Perspectives* NewYork: Cambridge University Press.

Wilson, M. (2002). Six Views of embodied cognition. *Psychonomic Bulletin & Review*, **9**(4), 625–636.

Yoshida, H. & Smith, L. B. (2005). Linguistic cues enhance the learning of perceptual cues. *Psychological Science*, **16**(2), 90–95.

Yoshida, H. & Smith, L. B. (2003). Shifting ontological boundaries: How Japanese- and English-speaking children generalize names for animals and artifacts. *Developmental Science*, **6**, 1–34.

Yu, C., Ballard, D.H., & Aslin, R.N. (2005). The role of embodied intention in early lexical acquisition. *Cognitive Science*, **29**(6), 961–1005.

Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory*, Vol. **44**, pp. 35–62, New York: Elsevier Science.