Preliminary Thoughts on a Rational Constructivist Approach to Cognitive Development

PRIMITIVES, SYMBOLS, LEARNING, AND THINKING

For Lin

INTRODUCTION

This chapter considers a newly emerging view of cognitive development: rational constructivism. I will attempt to sketch the view as I see it, in broad strokes. I will draw on arguments and evidence to see if an overall picture will emerge. Two key developmental issues are discussed: how to characterize the initial state, and how to characterize mechanisms of learning and developmental change. I will argue for the following theses: (1) Infants are much smarter and much more sophisticated learners than what William James, Piaget, or Quine had thought; infants’ world is not “a blooming, buzzing confusion.” However, it remains unclear how best to characterize the initial state. Instead of sensorimotor primitives or core knowledge, the initial state may perhaps be best characterized as a set of proto-conceptual primitives. (2) Over the last several decades of research on cognitive development three types of learning mechanisms have been uncovered: language and symbol learning as a vehicle for conceptual development; Bayesian learning as a tool for belief revision; and explanation, analogy, and related processes as ways to organize factual knowledge and generate new hypotheses that drive genuine conceptual change. These mechanisms may be considered both rational and constructive.
of inferences (Carey, 2009). For example, the object concept is not limited to the visual modality; similar principles guide infants' perception of objects in the visual and tactile modalities (e.g., if two pieces move together—common fate—infants perceive them as parts of one object). Infants' number sense not only allows them to perceive approximate numerosities in vision and audition, but it also allows them to perform simple arithmetic operations, such as approximate addition and subtraction, an example of rich inference. I have no doubt that the progress in developmental psychology warrants a serious reconsideration of the Piagetian accounts of development. I also share the confidence that the last several decades of infancy research has been illuminating. Here I raise a few questions about the core knowledge thesis and consider a revision. The perceptual–conceptual distinction has been central to the debate about the initial state of a human infant. For various reasons, more developmental psychologists are more sympathetic to and more comfortable with the view that infants start life with perceptual capacities and primitives. They can perceive color, size, shape, and motion and, indeed, there are well-studied brain areas that are designated for representing these basic perceptual primitives (e.g., Elman et al. 1999; Karmiloff-Smith, 1990; Piaget, 1954). These views are in agreement with the long-standing philosophical tradition of empiricism (e.g., Locke, 1690/1975; Hume, 1748/1999) and Piaget's characterization of the initial state in terms of sensorimotor primitives. Spelke, Baillargeon, R. Gelman, and their colleagues were among the first to challenge the Piagetian view, marshaling both theoretical arguments as well as new empirical evidence. The avalanche began with the seminal work of R. Gelman and Baillargeon (1987), Spelke (1989), and Baillargeon, Spelke, and Wasserman (1985). New methodological advances—in particular the development of the violation-of-expectation looking time paradigm—allowed researchers to ask questions of infants that could not be asked before. New studies with young infants barely 4 months of age suggest they may already represent persisting objects; even more astonishingly, these young infants may already have a set of principles that guide their reasoning about medium-sized objects, and these are the very same principles that adults use still—for example, contiguity, solidarity, and contact. Spelke et al. (1994) articulated the core knowledge view as an alternative to the standard Piagetian theory. They focused specifically on the centrality of these early concepts. Not only is the object concept embedded in a system of reasoning, it is also amodal—similar evidence is found in visual as well as tactile tasks (e.g., Sterri & Spelke, 1988). More recently, in an elegant, well-written, and ambitious synthesis of recent work on infant cognition and cognitive development in general, Carey (2009) argued forcefully against the Piagetian characterization, as well as also of the British empiricist, of the initial state. She marshaled evidence from a large body of literature supporting the claim that human infants are endowed with at least four "core cognition" systems: object, number, agents, and cause. Although Carey suggests that the format of representation for each of these concepts may be iconic, she also argues that these primitives are conceptual, using the criterion that these early representations play important informational roles in larger
conceptual structures. Within the core knowledge system of object cognition, a rich set of reasoning principles and heuristics develop. Simonsen, 1995; Xu, 1999; Xu & Carey, 2001; see also Ullman, 2004; Wilson & Brannon, 2006; Xu & Xu, 2008 for a review). Nevertheless, our criteria for bona fide conceptual processes are straightforward (Hirsch, 1984; Macnamara, 1988; Xu, 1997; Xu, 2007) and represent a kind of methodological criterion that is not part of a system of rules that relies on path and motion information for individuation and identity. Perhaps a better term for the initial object concept is the "object sense," an antonym to the "number sense." Similarly, when we take a closer look at the evidence for the number concept, there is strong evidence for an approximate number system—the "number sense" (Dehaene, 1997; 2005; Gallistel, 1990). This system, as has been argued by many, shares many properties of well-studied perceptual systems for discriminating duration, lightness, weight, or length—for example, Weber's law applies (Bennett, 1981; Lipton & Spake, 2004; Xu & Spake, 2005; Xu, 2006; Xu & Spake, 2005; see Feigenson, Dehaene, & Spake, 2004, for a review and studies with preschoolers and adults). For example, 6-month-old infants, and even newborns, can still make the differences between 6 and 8, 8 and 16, but not 8 and 18 or 16 and 32. To acquire the concept of positive integer, learners would have to abandon every belief (implicitly represented in the system for numeracy reasoning) they had about how numbers work and construct a brand new set of beliefs, such as how the successor function works, and how sets relate to each other and to the counting routine (Carey, 2009). When it comes to learning the meaning of number words (2, 3, 4, 5, etc.) these must be symbolic and discrete. Even more importantly, the "number sense" is nothing like the number line in its formal properties. For the number sense, the difference between 8 and 16 is not equal to the difference between 2, 16, and 24, in fact, the difference between 8 and 16 (12) is the same as that between 16 and 35 (19). To acquire the concept of positive integer, learners would have to abandon every belief (implicitly represented in the system for numeracy reasoning) they had about how numbers work and construct a new set of beliefs, such as how the successor function works, and how sets relate to each other and to the counting routine (Carey, 2009). When it comes to learning the meaning of number words (2, 3, 4, 5, etc.), there is little dispute that genuine conceptual change is needed. While (1990), 1993, Saramaka (2008), Saramaka & Lee (2009), and Le Corre and Carey (2007), among others, documented the developmental time course of learning to meaning of positive integers. These researchers discovered that children engage in the counting routine quite early (a practice that is strongly encouraged in the average middle-class American households), but it takes children about a year and a half to figure out how counting works in terms of the cardinality principle—children begin by figuring out the meaning of 1, then the meaning of 2 after a few months, then the meaning of 3 after another few months, and so on. Eventually they make the inductive leap that the next number on the count list represents "one more" from the one before that, that is, the successor function. A number of researchers have provided theoretical accounts of this developmental process (Carey, 2009). It argues that the transition from prelinguistic representations of numerosity to using symbols to think about numbers is referring to discrete quantities requires genuine conceptual change. I agree—the prelinguistic systems of approximating...
number representations and parallel individuation are not up for the job of supporting the learning of verbal counting. New representational resources need to be constructed.

What about causality, agency, and space? Perhaps parallel arguments can be made, there exist arguments in the literature on how causality changes from perceptual (i.e., defined by spatiotemporal parameters. Michotte) to conceptual (see the interventionist accounts of causality, e.g., Gopnik et al., 2004; Gopnik & Wellman, 2001; Woodward, 2003), and causal language may play an important role (e.g., Bonatti et al., 2012). There also exist arguments on how our representations of space change dramatically when spatial language is acquired (e.g., Hennert-Vaquez, Mothes & Mortelholth, 1991; Hennert & Spelke, 1996; Spelke, 2003) and the initial representations are modular and encapsulated in a perceptual input analyser, I am not aware of parallel arguments for the development of our concept of agency. My reading of the literature on the early understanding of agency is that it is rather messy and controversial. Some have argued for strong invariant views (e.g., Gergely & Csibra, 2003; Gergely, Nádasdy, Csibra, & Bolyai, 2005), while others favor a more piecemeal learning characterization (e.g., Meltzoff, 1998). One aspect of theory of mind development, false belief understanding, may give us another relevant example where infants may have some implicit understanding of other people’s beliefs, but these do not appear to be accessible to verbal predictions until later in development (for papers on this controversy see Prior, 2007; Ruffman & Farmer, 2005; Wimmer & Perner, 1983; and many others).

An alternative construal is that infants’ initial knowledge is embedded in a set of pre­conceptual analysers that are capable of very sophisticated inferences and computations. Perhaps these representational schemes are best characterized as proto-conceptual primitives because they do not deliver representations in the relevant format for the learning of language, and children seem to have a great deal of difficulty building other, more complex concepts out of these primitives. In other words, it is not a simple mapping process that is genuinely symbolic and detailed understanding of these changes may require new primitives, new hypotheses, and new ways of learning and thinking.

THREE TYPES OF MECHANISMS OF LEARNING AND DEVELOPMENTAL CHANGE

What are the learning mechanisms that bring about developmental change? In other words, what is “rational” and “constructivist” about the young human learner’s mind and the courses of development? I suggest that there are at least three types of mechanisms for learning and developmental change: (1) language as a medium for providing placeholders for knowledge and pri­or knowledge for the purpose of peripheral concepts may become core concepts, and vice versa; and (2) exploration, analysis, thought experiments, and other internal processes of “learning by thinking” as tools for going beyond data and evidence to build larger conceptual structures. To date, we have some evidence for each of these three types of learning in the domains that I have considered in the first part of this paper.

Language Learning as a Vehicle for Conceptual Change

If infants’ initial knowledge is not in the right format for language and infants begin to learn the meaning bearing parts of language in earnest toward the end of the first year, some developmental changes need to take place. One hypothesis that has been argued for in the literature is the idea that various parts of language provide “placeholders” for conceptual development (e.g., Carey, 2009; Gelman, 2011; Xu, 2002, 2007; among others). Two case studies come to mind.

In research on whether infants’ representations of objects include the crucial ingredients of (a) understanding of how words refer to objects, and (b) understanding of the learning of language: (c.g., Fodor, 1998), the potential for symbolic representation and thinking” refers to a domain (e.g., Spelke et al., 1993; Xu & Carey, 1996; Wynn, 1993, among others). Later on, we have suggested, infants begin to learn words for object kinds such as BALL, CUP, and DOG, and it is through word learning that infants build new concepts that allow them to individuate and track identity under sortal-kinds BALL, CUP, and DOG (see Xu, 2001, for a review). The words—count nouns that refer to sortal-kinds—impose certain constraints on the format of the representations: Symbolic (often) refers to mutually exclusive categories that reflect distinct underlying essences or causal structures, and support inferences at the level of kinds. The case of number is similar, as sketched out earlier in the chapter. The pre­linguistic representations, especially the part that is genuinely a system capable of numerical computations, the approximate number system, does not seem to play a role in acquiring the meanings of number words. Instead, the counting list provides a placeholder structure (Carey, 2009; Sarnacka & Carey, 2008) that imposes a set of constraints on the format of the representations: symbolic, discrete, generated by the successor function, goes on to infinity, among others.

Core knowledge systems may rely on language to become truly symbolic and discrete, and this is a deep conceptual change that requires new representational resources to be constructed. The “language of thought” may not have existed before these changes have taken place, since by definition, the LOT is supposed to respect the syntactic constraints of a natural language (e.g., Fodor, 1975). Note that this is not a Whorfian idea—all languages have a set of syntactic tools and these are, as far as I know, universal. Even if different languages express some ideas differently using different syntactic devices, the totality of the thoughts we can entertain would remain the same for all human learners. A child learning English and a child learning Chinese may use different syntactic tools to transform core knowledge representations into symbolic
Processes of Conceptual Change and Conceptual Representations, but at the end of the day, the thoughts that can be expressed in each language would be the same.

Bayesian Inductive Learning as a Tool for Belief Revision

Much recent research on cognitive development focuses on understanding statistical learning and probabilistic inference mechanisms (see Gopnik & Wellman, 2002; Xu & Khusnutdinov, 2011, 2015 for reviews). This line of inquiry is largely inspired by the surge of Bayesian computational models in cognitive science (see Tenenbaum, Kemp, Griffiths, 2006, and Griffiths, Chater, Kemp, Perfors, & Tenenbaum, 2010 for reviews).

Much of this work was motivated by the idea that we need an approach to cognitive development that is neither extreme empiricism nor extreme nativism. Nativist theories have focused primarily on specifying innate concepts and core knowledge systems, and but also that of infants and young children (e.g., Chomsky, 1987; Fodor, 1975; Pinker, 1994), whereas empiricist theories have focused on specifying associative learning mechanisms and the graded nature of our learning and representations (e.g., Spelke, 1988; Spelke & Breinlinger, 1994). Neither view appears to provide an adequate account of all the empirical findings on cognitive and language development. The inadequacy of both extreme nativist and extreme empiricist views has led researchers to try to find a substantive middle ground (e.g., Johnson, 2000; Newcombe, Ratliff, Shaller, & Tsvyash, 2005).

The rational constructive approach blends elements of a constructive account of conceptual change with the account of learning as rational statistical inference that underlies probabilistic models of cognition (Chater & Oaksford, 2008; Griffiths et al., 2005; Tenenbaum et al., 2011). Several basic tenets have been laid out elsewhere (see Xu & Khusnutdinov, 2011):

- Human learning is best described as a form of rational Bayesian inference. The learner starts with some prior probability distributions over a set of hypotheses and computers the posterior probabilities of these hypotheses given the strength of the evidence as given by Bayes rule. This is a computational level characterization; that is, it describes the inferential process without making a priori commitments to how that process is instantiated at the algorithmic level (what steps to follow when a learner wants to solve a particular task) (Grunberg, 1982).
- Hypotheses can be represented as probability distributions. Inferences are probabilistic and gradient, so hypotheses are not simply ranked in or out. Instead, learners may be more or less confident about the various hypotheses.
- Learners represent the world not just by forming associations and correlations but by constructing abstract, causal, generative models.

A Rational Constructivist Approach to Cognitive Development

• Learners acquire new concepts and biases in the course of development; the newly acquired knowledge becomes part of the prior and thus constrains subsequent learning.
• Domain-general learning mechanisms may give rise to domain-specific knowledge.
• Representations may differ in their strengths: some support predictions, actions, and explanations, while others may not.
• Learners are actively engaged in the learning process, from infancy to adulthood.

Empirical evidence for this view has been rapidly accumulating in the last decade, starting with the work on causal learning in young children (Gopnik & Sobel, 2000; Gopnik et al., 2004; Gopnik & Wellman, 2011). Many studies have demonstrated that young children are sensitive to probabilistic evidence when inferring causal structure, and children and adults’ behavior can be well explained by Bayesian representations and Bayesian learning models (e.g., Schulz, Bonawitz, & Griffiths, 2007; Sobel, Tenenbaum, & Gopnik, 2012).

Many domains have been investigated using this set of analytic and computational tools. In word learning, studies have shown that children are sensitive to “suspicious coincidences” and their inferences in learning the meaning of new words can be modeled by a Bayesian model that rationally combines the input data with a structurally impoverished space (e.g., Frank & Goodman, 2010; Frank, Goldwater, Griffiths, & Tenenbaum, 2010; Xu & Tenenbaum, 2007, 2009, 2012). In social cognition, various researchers have demonstrated that learning about preferences and theory of mind can be construed as a rational inferential process (e.g., Baker, Saxe, & Tenenbaum, 2009; Kusnitch, Xu, & Wellman, 2010; Lucas, Griffiths, Xu, & Fawcett, 2009; Xu & Xu, 2011). In physical reasoning, several studies have shown that even pre-linguistic infants can carry out simple probabilistic inference tasks (e.g., Denison & Xu, 2008, 2010b; Tegla, Giorno, Gontra, & Borrani, 2009; Treagust et al., 2001; Xu & Denison, 2009; Xu & Garcia, 2008; see Denison & Xu, 2011, for a review). In pedagogical learning, formal Bayesian models have led to some innovative empirical studies with children and adults (e.g., Bonawitz et al., 2011; Shafro & Goodman, 2008; Shafro, Kemp, Manninghala, & Tenenbaum, 2011; see Shafro, Goodman, & Frank, 2013 for a review). In speech segmentation and rule learning in language, many have provided strong empirical evidence for rational inferential processes in infants (e.g., Dawson & Gerken, 2009, 2011; Frank & Tenenbaum, 2009; Gerken, 2006, 2009; Goldwater, Griffiths & Johnson, 2007).

Research has also uncovered mechanisms for making inferences at multiple levels that give rise to learned inductive biases, in domains such as word learning and causal reasoning (Dewar & Xu, 2010; Lucas & Griffiths, 2010; Samuelson & Smith, 2007; Sim, Yuan & Xu, 2011; Sim & Xu, in press; Smith, Jones, Landau, Gerishoff-Stowe, & Samuelson, 2013).

Explanation, Analogy, and Thought Experiments

The rapidly growing research on statistical learning and statistical inference paints a picture of a child as a superdata-crunching machine, but not all learning is data-driven.
Process of Conceptual Change

Some psychologists have used the phrase "learning by thinking" to refer to a set of cognitive activities that (seemingly) generate "knowledge from nowhere" (T. Lombrano, personal communication). The basic idea is that our native picture of learning—that we are exposed to various learnable facts in the world—is inadequate, and many celebrated examples in the history of science show that scientists such as Galileo or Einstein arrived at major scientific breakthroughs without the benefits of grants, graduate students, and laboratories. Their scientific insights come from "mere thinking"—ways to organize and extend what we already know by manipulating existing representations and data structures. Several such cognitive activities have been demonstrated in lay people and scientists: explanation, analogy, mental simulation, and thought experiments.

One well-studied case is the self-explanation effect (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi, de Leeuw, Chiu, & LaVancher, 1994). In a typical experiment, some participants are told to explain to themselves when given some math problems, while others are asked to "think aloud" with the same set of math problems. The main finding is that the explainers outperformed their nonexplaining counterparts, and the explainers did better on transfer problems that went beyond the studied examples. This appears to be a clear case of generating new insights by manipulating existing data, and the mental activity of explanation plays a crucial role. More recently, Williams and Lombrano (2010) found that in a category learning task, participants who were asked to explain were more likely to discover broad regularities that provide an account of category structure, compared to participants who were given free study time or were instructed to think aloud during the study. Their idea is that explanation may generate new knowledge by encouraging learners to find underlying rules and regularities.

Work on analogy (most notably Christie & Gentner, 2000; Gentner, 1983; Holyoak, 2001) has suggested that structural alignment is the mechanism by which a base domain is mapped onto a target domain. Such an alignment encourages the learner to see the structural similarities between two domains and allow them to use their existing knowledge in one content domain to understand the structure of a new content domain.

Other forms of "learning by thinking" include mental simulations and thought experiments. Some research suggests that we solve mental problems by mentally simulating the process—appealing to these visuospatial representations provides solutions that we do not seem to have access to by verbally reasoning through the problem (e.g., Hargreaves, 2004). Thought experiments have been mostly studied in philosophy of science (e.g., Gendler, 1996, 2000). One celebrated example is how Galileo worked out that all objects, regardless of their weight, would fall at the same speed (contrary to the then standard Aristotelian theory).

These learning processes are not driven by new data and new evidence. Instead, the learner possesses the ability to manipulate existing representations and data structures in the head, and new insights emerge. There exists some research on how children use explanations and analogy in learning (e.g., Christie & Gentner, 2000; Legare, 2002; Legare, Gelman, & Wellman, 2002), much more work is needed to truly understand how these processes play a role in cognitive development.
Processes of Conceptual Change

...—that the learner comes to a task with, we can proceed to study how that person combines prior knowledge with new evidence to choose among a set of hypotheses and to update his or her beliefs. I think this is an important point and a "bonus" for thinking about inductive learning as Bayesian updating. Second, most Bayesian models have focused on modeling the data-driven processes. It is yet to be seen if these formal models can capture the effects of language learning and "learning by thinking" (see Piantadosi, Goodman, & Tenenbaum, 2012, and Ullman, Goodman, & Tenenbaum, 2012, for some examples). The formal tools may become very useful when we want our assumptions and algorithms clearly specified. Here I discuss Bayesian inductive learning as one type of learning mechanism by focusing on the conceptual idea of hypothesis testing and belief-revision. Third, it may turn out to be the case that the mechanisms of learning discussed here would account for the development of some aspects of core knowledge, but this jury is still out. So far we have not shown that any particular aspect of core knowledge is learned, but the developmental changes we see during the first year of life suggest that a learning story (as opposed to mere maturational one) is possible. Fourth, many have asked the question, "Where do hypotheses come from?" When we talk about Bayesian learning—none of us has a clear answer at the moment. Here are two possibilities: Pomo-conceptual primitives may provide an initial hypothesis space, and non-data-driven cognitive activities may also generate new hypotheses for the learner along the way. Lastly, most developmental psychologists have focused on demonstrating that infants and children learn quickly—within the time limits of a single, 10- to 20-minute, lab visit. However, long-term development and conceptual change are surely more complicated. Not only do we keep track of statistical evidence over time and across subdomains, we may also need the non-data-driven cognitive activities such as exploration, analogy, mental simulation, and thought experiments in order to achieve genuine, qualitative conceptual change.

REFERENCES


