TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science 2 (2010) 751–770

ISSN: 1756-8757 print / 1756-8765 online
DOI: 10.1111/j.1756-8765.2010.01116.x

Emergence in Cognitive Science

James L. McClelland

Department of Psychology, Stanford University

Received 9 February 2010; received in revised form 21 June 2010; accepted 3 July 2010

Abstract

The study of human intelligence was once dominated by symbolic approaches, but over the last 30 years an alternative approach has arisen. Symbols and processes that operate on them are often seen today as approximate characterizations of the emergent consequences of sub- or nonsymbolic processes, and a wide range of constructs in cognitive science can be understood as emergents. These include representational constructs (units, structures, rules), architectural constructs (central executive, declarative memory), and developmental processes and outcomes (stages, sensitive periods, neurocognitive modules, developmental disorders). The greatest achievements of human cognition may be largely emergent phenomena. It remains a challenge for the future to learn more about how these greatest achievements arise and to emulate them in artificial systems.

Keywords: Development; Emergence; Explanation; History; Language; Modeling; Neural networks

1. Introduction

This article arose from an invitation to consider ''statistical models'' in cognitive science on the occasion of the 30th anniversary of the Cognitive Science Society, and it paralleled an invitation to John Anderson to consider ''symbolic models.'' In considering the invitation, it seemed to me that ''statistical models'' did not exactly capture what I take to be the alternative to symbolic approaches that has arisen in our field since the first Cognitive Science Society meeting—an approach represented in the work of connectionist modelers as well as other researchers who come from a wide range of different starting places. A better framing concept for this alternative seemed to me to be the concept of *emergence*.

The symbolic approach takes as its starting point the idea that human and artificial minds should be viewed as symbol processing machines. This approach was well represented at

---

Correspondence should be sent to James L. McClelland, Department of Psychology, Stanford University, Stanford, CA 94303. E-mail: mcclelland@stanford.edu

the first Cognitive Science Society meeting by Alan Newell, in his lecture on *Physical Symbol Systems* (Newell, 1980). The idea of emergence in cognitive science is the contrasting idea that there are more basic or elementary processes that are really the fundamental ones, and that physical symbol systems of the kind Newell described are sometimes useful approximate characterizations which, however, have difficulties in capturing in full the context-sensitive, flexible, graded, and adaptive nature of human cognitive abilities. While it might not have seemed so until recently, it has become clear that statistical and symbolic approaches can easily coexist with each other, as they do in the structured probabilistic models of Kemp and Tenenbaum (2009) and Goodman, Mansighka, Roy, Bonawitz, and Tenenbaum (2008). While these approaches advance the state of symbolic approaches, they are quite different from the emergentist approaches this article is considering (McClelland et al., 2010).

The concept of emergence may be relatively new to cognitive science, but it is not new to science as a whole or to philosophers of science. Lewes (1875/2005), an early protagonist of the concept, contrasted emergence with resultance: ‘‘In resultance, every result is either a sum or a difference of the cooperant forces. It is otherwise with emergence where there is a cooperation of things. The emergent is unlike his components insofar as these are incommensurable, and it cannot be reduced to their sum or their difference.’’

Emergent properties are often defined as properties that are not found in any component of a system but are still features of the system as a whole. Interestingly, the Wikipedia article on emergence (Wikipedia, 2010) uses intelligence as one of its examples: According to an emergentist perspective, it states, intelligence emerges from interactions among neurons. For Wikipedia, the mind may be the most interesting emergent system that we know about. From the perspective of emergence, the article states, it is not necessary to propose a soul or a central executive to account for the fact that brains can be intelligent, even though the individual neurons of which they are made are not. In a similar vein, adaptive computer programs can have emergent properties; some such programs have acquired skills in board games that exceed those of their designers—and they exhibit emergent patterns that can be described as units or processes, such as gambits or strategies (Holland, 1998).

Philosophers of science who have studied emergent phenomena (Bunge, 2003; Morowitz, 2002) have noted that they may be more or less complex than the elements of the substrate that give rise to them. The ideal gas laws seem very simple. The interactions among the molecules in those gases are complex, and yet there is a very simple regularity that emerges.

When it comes to the mind, we also often observe simple and robust regularities, like the power law of practice. These things are the emergent consequences of a very complex system indeed, even though the law itself is simple. It is certainly interesting to understand how people’s behavior can exhibit simple emergent regularities such as the power law of practice while still having a rich and complex internal structure in each particular case, and this is one legitimate form of research on emergent phenomena in cognitive science, as it is in physics and chemistry.

But, in fact, these simple regularities are not the essence of intelligence or the supreme achievements of nature. When it comes to intelligence, the real stuff consists of human success in everyday acts of perception, comprehension, inductive inference, and real-time behavior—areas where machines still fall short after nearly 60 years of effort in artificial intelligence—as well as the brilliant creative intellectual products of scientists and artists such as Newton, Darwin, Einstein, Shakespeare, Michaelangelo, and Beethoven. According to an emergentist perspective, all of these products of the mind are essentially emergents. I do not think anyone who emphasizes the importance of emergent processes would deny that planful, explicitly goal-directed thought plays a role in the greatest human intellectual achievements. However, such modes of thought themselves might be viewed as emergent consequences of a lifetime of thought-structuring practice supported by culture and education (Cole & Scribner, 1974). Furthermore, proponents of the essence of human thought as an emergent phenomenon might join with Hofstadter (1979) and others in suggesting that key flashes of insight and intuition may not have arisen from planful, explicit goal-directed thought alone, but instead might reflect a massive subsymbolic constraint-satisfaction process taking place outside of awareness. In the case of Darwin, for instance, biographers (e.g., Quammen, 2006) have written about the origins of his work on his theory of evolution. It appears that Darwin set his mind to this investigation knowing intuitively that there was something interesting to discover, while not knowing exactly what it was. This intuition, arguably the key factor in his discovery, might have arisen as an emergent consequence of a subconscious constraint-satisfaction process, which then led him to engage in more intentional (yet still perhaps intuition-guided) exploration. This sequence in discovery may be the rule even in formal domains such as mathematics and physics, where the intuition may come first, followed only later by formal specification and rigorous proof (Barwise & Etchemendy, 1991).

2. The generality of emergence in nature

Before turning to applications of the concept of emergence in the field of cognitive science in more detail, let us consider examples of emergent phenomena more broadly. Several examples are listed in Table 1. Many of these are taken from the popular book *Emergence* by Johnson (2001) and many are discussed in *The Emergence of Everything* (Morowitz, 2002). Emergent phenomena are pervasive in physics. Transitions between solid, liquid, and gaseous states are considered to be emergent phenomena. The properties

Table 1
Examples of emergent phenomena in nature

Transitions between solid, liquid, and gaseous states
Properties of molecules, proteins, organelles, cells, organs, organisms
Bubbles, honeycombs
Mountains, oceans, rivers, continents, planets, solar systems, galaxies, universes
Life forms of all types; evolution and development of organisms
Ant colonies
Properties of individuals in collections
Markets, economies
Cities

of atoms and molecules are seen as emergents. The properties of water are not contained either in oxygen or hydrogen or in any sort of resultant from their additive or subtractive effects on each other, but rather from the consequences of the particular configuration that they enter into when combined into molecules and from the ways in which these molecules interact with each other and with many other kinds of molecules. So it is with proteins, organelles, cells, organs, organisms, and many other things found in nature. Liz Bates, an eloquent proponent of emergentist approaches in language and development, liked to talk about the regular hexagonal structure of the honeycomb to illustrate her point that there are many structures in nature that are not produced by design, but that emerge as a result of simple forces operating on the individual (spherical) honey-bubbles produced by bees (Bates, 1976). Bubbles themselves are emergents—as are mountains, oceans, rivers, continents, planets, solar systems, and so on.

In biology, the most important single development in all of biological science may have been the articulation of the proposal that organisms—and all of their highly complex structures and functions—were not designed but evolved through the effects of random mutation and recombination, together with a greater likelihood of survival for some of the resulting variants. The revolution Darwin produced by introducing this idea is certainly one of the most profound revolutions in scientific thinking that has ever occurred. It replaces an idea about our origins and our nature in which some external agent placed us here with a process that arose completely without design or plan, from things far simpler than those that resulted from it.

Many, many other examples could be cited. Johnson's *Emergence* discusses ant colonies, properties of individuals in collections, markets, economies, and cities as things that arise naturally from the interactions of large numbers of constituents unlike themselves, without central coordination or design. Of course, institutions subsequently come into existence that attempt to impose design and regulation on these emergent entities, much as explicit thought processes subsequently come along and lead us to try to bring clarity to our intuitions.

Consider the entities that we find in nature. The galaxies, solar systems, stars, planets, continents, mountain ranges, and river systems that we find are the consequences of processes many of which are well understood in principle; in some cases it may even be well understood why there should be structures of all of these types, and what the forces are that cause such structures to emerge, and even what some of their statistical properties are considered in aggregate. But when we consider the individual cases, we find that every one is different, and that the forces that shaped each one are highly nonlinear and context dependent, to the extent that a full understanding of how it came out the way it did may not be possible. The same is true in biology, but here it applies not only to individual organisms but also to species. Morowitz (2002) notes how we may know something of the context that led to the extinction of dinosaurs and the emergence of large mammals, or to the emergence of bipedal hominids from tree-dwelling primates (an ancient change in climate promoting open savanna over forest may have been involved), but the particular details of which hominid line gained the ascendency may never be fully clear; nor should we think (he and many biologists have argued) that all the features of *homo sapiens* are in any global sense optimal. Rather, these are the features we inherited from ancestors that had a relative advantage over other competing species, in the particular context that existed at that time. On this view, we may never understand how we came to be as we are, nor should we be deluded into thinking that we are truly optimal in any of our properties.

3. Emergence of emergence in cognitive science

The notion that many of our mental abilities may be emergent phenomena was not prominent at the first meeting of the Cognitive Science Society in 1979, but it began to emerge around that time. Concepts related to emergence already existed in the field of development (Waddington, 1942) and began to show up in writings by systems neuroscientists (Braitenberg, 1984), artificial intelligence researchers (Minsky, 1980, 1986), and cognitive scientists (Hofstadter, 1979) around this time. John Holland's work in computer science, introducing genetic algorithms that allow computational agents to evolve through a stochastic, competitive, adaptive process is another early example of an emergentist approach to intelligence (Holland, 1975).

A difficulty with Hofstadter's and Minsky's ideas was the vagueness that accompanied their breadth. It seems fair to say that these ideas remain more sources of inspiration than actual tools for working toward an explicit understanding of human mental abilities (but see Hofstadter and the Fluid Analogies Research Group, 1995). As one example of such inspiration, Smolensky's (1986) Harmony Theory, which he applied to intuitive reasoning about electrical circuits, and later language, drew inspiration from Hofstadter's ideas.

Dynamical systems approaches to cognition, motor control, and development provide an important source of emergentist ideas in our field. A series of papers in the late 1980s (e.g., Schöner & Kelso, 1988) introduced ideas originating in physics on self-organization in complex systems by Haken (1977) into the investigation of motor control (see Turvey, 2004 and other papers in Jirsa & Kelso, 2004 for more recent consideration of these ideas). This seminal work from the 1980s serves as the foundation for the dynamical systems approach in psychology and development, to be considered further below.

3.1. *Proto-emergence: Conspiracies of mental agents*

Although not featured at the first cognitive science meeting, what one might call a proto-emergentist perspective lies at the heart of the interactive activation model and related models developed around that time (McClelland, 1981; McClelland & Rumelhart, 1981). In experimental psychology, there was also a related body of work on exemplar models (e.g., Medin & Shaffer, 1978), in which categories and rule-like behavior are emergent consequences of the combined activity of many elements (the ''mental agents'' of the title of this section) corresponding to individual items or experiences.

In these models, there is a conspiracy of memory representations of items, to produce emergent category representations. Indeed, the representation of a given item is a result of a similar conspiracy, depending in a graded way on an ensemble of stored representations of items previously experienced (including representations of previously encountered instances of the item itself if it is a familiar item). Experimental work by Glushko (1979) applied this idea to the construction of pronunciations of both familiar and unfamiliar letter strings; he suggested that a conspiracy, rather than a race between a lexical lookup process and a system of rules, might be the basis of constructing a pronunciation of all kinds of letter strings. This idea was then incorporated into the interactive activation model of the perception of letters in words and in novel but pronounceable pseudowords (McClelland & Rumelhart, 1981): For pseudowords, activation of representations of many words partially matching an input letter string resulted in the facilitation of the perception of all of the letters in the string. This process gives rise to an emergent tendency to facilitate perception of items consistent with the patterns of English orthography, without explicitly representing this knowledge in a system of rules, as in other approaches (e.g., Spoehr & Smith, 1975). This not only challenges the need to have explicit knowledge of the rules but also allows us to ask: Does it make sense to think that a consistent set of rules actually exists? In fact English orthography reflects factors such as the need to distinguish between letters, such as *n* and *v*, which were not distinguished in the script used by scribes in an earlier day; the letter *e* was thereby added to several words, creating irregular forms like *live*, *give*, and *have* (Venesky, 1970). Perhaps any rules we might construct really should be viewed simply as approximate and sometimes useful descriptive characterizations.

3.2. Emergence in distributed connectionist models

I describe these localist connectionist and instance-based models as proto-emergent approaches because they still contain individual units corresponding to familiar items, and in that sense are not completely subsymbolic. Many researchers cling to such models, taking it as a crucial property of our cognitive system that it contains units that stand for individual things (Bowers, 2009; Page, 2000). However, for others these models were stepping stones to a more thorough-going emergentist approach of the kind embodied in distributed connectionist models (Hinton, McClelland, & Rumelhart, 1986), though they remain useful for many purposes. The approach builds on ideas traceable back to Hebb (1949) and Lashley (1950). Around the time of the first cognitive science meeting, explicit computational models embodying these ideas (Anderson, Silverstein, Ritz, & Jones, 1977; Hinton, 1981; Hopfield, 1982) began to influence the field of cognitive science.

Hopfield (1982) considered a ''memory'' to be a pattern of 1's and −1's over the units in such a network; such a memory was stored by adding a (positive or negative) increment to the weight from one unit to another, and he considered memory retrieval to be the process whereby a distorted version of a memory would be input to the units in the network, which would then be allowed to adjust their activations one by one in random succession. Applying this process, Hopfield observed that his network would always settle to a state that was in a sense better than the state in which it started—a state of lower ''energy''—and that, if only a few memories were stored in a set of neurons, the state reached would usually correspond to one of the memories.

The idea that the global state toward which a network settles, rather than a single unit within the network, corresponds to a memory represents the needed conceptual advance to go beyond the proto-emergence of the localist models to a full-blown emergentist approach, in which the representation for a familiar item cannot be found stored separately in any part of the system, and in which its retrieval as a memory emerges from the interactions among simpler units each doing something subcognitive, and very simple.

Rumelhart built on Hopfield's idea in a paper he first described at Cognitive Science in 1983 (Rumelhart et al., 1986b). He used a slight variant of the Hopfield model to instantiate his vision of the *schema*—a concept that, he argued, was not well instantiated in models which specified that schemata were directly represented as such (Rumelhart & Ortony, 1977). By representing schemata as emergent states of neural networks, he suggested, they could capture the flexability, blendability, and embedability of human knowledge. In the particular model Rumelhart proposed, the connection weights represented the aggregate co-occurrence statistics of properties of many different rooms (rooms Rumelhart and his secretary were personally familiar with). By activating one or two properties of a room (''has an oven and a sink,'' e.g., or ''has a bed and a dresser''), one could guide the network to settle into one of several different attractor states (Fig. 1). Some atypical combinations could be accommodated more easily than others—for example, sofa, most at home in living-room-like states, could be better accommodated in bedroom-like states than in kitchen- or bathroom-like states. Thus, the states captured graded goodness of particular combinations of elements, even in cases where particular specific combinations had never been encountered before. Instantiated schemata (specific activated patterns corresponding to particular rooms) also could contain embedded schemata, such as a schema for the combination of properties *has a window* and *has a curtain*. Without a window, curtains were not good, but with a window, curtains tended (again, in many but not all contexts) to be better than no curtains. These emergent schemata provided the flexibility and context sensitivity that Rumelhart was looking for, and they contributed tremendously to the enthusiasm he

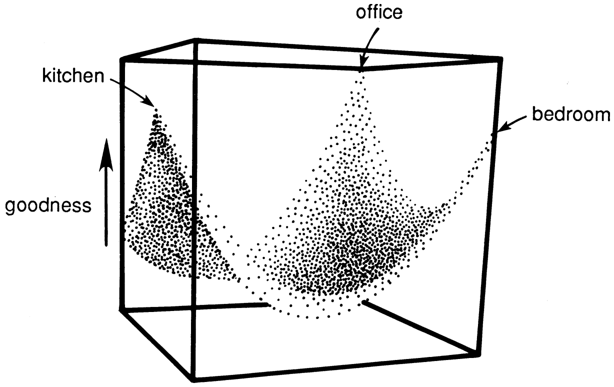

Fig. 1. Emergent room schemata and their associated goodness values. Goodness values shown for kitchen, bedroom, and office are attractor states in a Hopfield-like connectionist network. Goodness values are also shown for intermediate states on the plane through these three attractor states. From Figure 7, p. 28 of Rumelhart, Smolensky, McClelland, & Hinton, 1986b. Copyright 1986 by MIT Press.

brought to the effort to characterize human cognition as an emergent consequence of the interactions among simple neuron-like processing units.

4. Cognitive science constructs as emergents

Since this early work, the investigation of emergence in cognitive science has continued. In Table 2, I list a number of constructs in use within our field, all of which may be viewed as emergents of one form or another, as I will now discuss. I have chosen to list constructs of several different types, and because of space constraints I have been selective—similar issues arise with many other constructs.

4.1. Putative representational constructs

The first type of construct I consider includes cases related to the schema example. One such case—the treatment of linguistic rules as emergents—arose in parallel with the development of the emergent approach to schemata, in the PDP model of past-tense formation (Rumelhart & McClelland, 1986). This and related work touched off a firestorm of critical reaction in the late 1980s (Fodor & Pylyshyn, 1988; Lachter & Bever, 1988; Pinker & Prince, 1988), as well as an exciting body of further work that extended these ideas in several important directions. Shortly afterward, related approaches to single-word reading (Sejnowski & Rosenberg, 1987), sentence processing (Elman, 1990), and to the organization

Table 2
Some objects of investigation in cognitive science: Are they all emergents?

- Putative representational entities
  - Categories, prototypes, schemata
  - Rules of language and thought
  - Lexical entries
  - Grammatical and semantic representations
- Architectural constructs from cognitive psychology and cognitive neuroscience
  - Working memory
  - Attention
  - Central executive
  - Declarative memory
- Cognitive processes and their outcomes
  - Choices
  - Decisions
  - Inferences
  - Beliefs
- Developmental processes and outcomes
  - Developmental stages
  - Sensitive periods
  - Cognitive modules
  - Patterns of deficit seen in developmental disorders
  - The structure of natural language

of conceptual knowledge arose (Rumelhart & Todd, 1993), spurred on by the development of powerful learning algorithms for training multilayer networks of simple processing units (Rumelhart, Hinton, & Williams, 1986a).

This work has raised a host of still unresolved questions. Can we understand people's knowledge of words, including their knowledge of whether an item is in fact a word in their language, and their knowledge of all things about the word, such as what the word means, how it is spelled, how its past tense is formed, and how it is used in context, without postulating that people have explicit ''lexical entries'' in their heads corresponding to the entries one finds in the dictionary? Is knowledge of the structure of sentences, and the rules needed to create them, represented as such, and perhaps innately predetermined, or could this knowledge too arise from interactions of simple processing units, whose connection weights are affected by experience, when exposed to spoken language?

The emergentist perspective articulated in distributed connectionist models answers ''yes'' to these and many related questions. There are still many protagonists of alternative approaches in which separate units (or small collections of dedicated neurons) are allocated to familiar items (Bowers, 2009) and/or in which there is a separate and specialized mechanism good for use in implementing the ''algebraic rules'' that some still suppose are best suited to capturing knowledge of linguistic regularities and certain types of inference (Marcus, 2001). Elsewhere, my colleagues and I have vigorously defended the emergentist approach (McClelland & Bybee, 2007; McClelland & Patterson, 2002a). One cornerstone of our reply is that the shortcomings critics found in the earliest models lay not in their essential features but in specific ancillary implementation details; improved versions have addressed most of the criticisms (MacWhinney & Leinbach, 1991; Plaut, McClelland, Seidenberg, & Patterson, 1996).

More important, we have argued, the behavioral and linguistic phenomena do not support a separation into dissociable categories of regular and exceptional items subject to dissociable mechanisms and processes, as some continue to claim (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Jackendoff, 2007; Pinker & Ullman, 2002). Many items others are forced to treat as exceptions exhibit characteristics of regular forms (e.g., ''said'' or ''kept,'' treated by some as exceptions but really just slight reductions of the regular past tenses of ''say'' and ''keep''), and behavioral and linguistic evidence suggests that these items benefit from their similarity, not only to other similar exceptions (''wept,'' ''slept,'' etc.) but also from their near-regularity. We argue that approaches in which sensitivity to regularities as well as to the properties of individual items arise within a single system are better suited to capturing the regularity in both fully and partially regular items (McClelland & Bybee, 2007; McClelland & Patterson, 2002b). Further development of such approaches is clearly warranted, in my view.

4.2. Architectural constructs from cognitive science and cognitive neuroscience

Cognitive scientists and cognitive neuroscientists have introduced or maintained a set of constructs that sometimes show up as explicit separate parts of an overall architecture, but that, in many cases, can be seen instead as functions that emerge from the interdependent activity of a number of contributing mechanisms or processes. Consider, for example, the ‘‘central executive.’’ The central executive might be viewed as a separate mechanism that plays a special role in controlling processes elsewhere in the system, and it has often been associated with regions in the prefrontal cortex—regions that, when damaged, lead to reduced ability to flexibly control behavior according to current goals and an increased tendency to make stereotypical or habitual responses. If fact, however, many researchers propose that the control of processing is actually distributed across many contributing brain areas; in some versions of these ideas, the ability to maintain control depends on coherent engagement of several brain areas, including but not limited to regions of prefrontal cortex (Fuster, 1997). Others have noted that patterns of behavior others treat as reflecting the loss of a central executive can arise from diffuse damage and capture such patterns by diffuse damage to simple recurrent networks in which there is no separate component corresponding to the central executive (Botvinick & Plaut, 2004).

Several putative features of the architecture of memory can also be approached in similar ways. But perhaps there still may be a tendency to see declarative memory as a separate memory system and within declarative memory, to see semantic and episodic memory as further subdivisions (see, e.g., the widely reprinted Figure 1 in Squire, 1992). It seems increasingly clear, however, that declarative memory involves the collaborative engagement of complementary learning systems, including the medial temporal lobes and many areas of the anterior and lateral temporal neocortex (McClelland, 2010; McClelland, McNaughton, & O’Reilly, 1995), as well as many regions of the prefrontal cortex (Badre & Wagner, 2007). Likewise, ‘‘semantic memory’’ is now often seen as an emergent property of mutually interconnected brain areas (Martin, 2007).

4.3. Cognitive processes and their outcomes: Decision making

In both popular and scientific writing about choice and decision making, there is a sense that a choice is being made by a deliberative agent, and in common parlance we speak of people as making choices, and treat choices and decisions as discrete outcomes. But how are these choices and decisions made, and are they necessarily discrete? Some approaches have tended to ignore the process, instead focusing on the principles of rational choice or on specific violations of rationality, which are then described in terms of ‘‘heuristics’’ and ‘‘biases’’ that affect the decisions made. As alternatives to these approaches, dynamical process models that treat decision making as arising from a dynamic, competitive process now account in detail for the pattern of data from a large number of choice and decision-making experiments, including details of response time distributions and dependence of responses on stimulus and payoff information (Roe, Busemeyer, & Townsend, 2001; Usher & McClelland, 2001; Wang, 2002). These models rely on competition between pools of units representing choice and response alternatives to create a decision-like state, in which the population of units associated with one outcome reaches a sustained activity state while the activation of units in populations associated with alternative outcomes are relatively suppressed. Decision states can be attractor states in such systems, emergent consequences of the interplay of excitatory and inhibitory interactions among neuronal processing elements.

One interesting feature of such states is that they can have characteristics that lie between completely discrete and completely continuous states. For example, because of competition, one alternative can win out over other alternatives and suppress their activity; yet the degree of activation of the winning alternative can still reflect the strength of the input supporting that state, and this can have consequences for the speed, confidence, and malleability of the ''decision'' the system has ''made.''

4.4. Developmental processes and outcomes

Cognitive development affords a very rich domain for emergentist approaches, including both connectionist (Elman et al., 1996) and dynamical systems approaches (Thelen & Smith, 1996). The idea that the mind emerges gradually from interactions between the child and its physical and cultural milieu has parallels in ideas about the physical development of the human body. Until surprisingly recently, it was widely believed that human adults (and other adult animals) developed from preformed miniature versions—an idea known as preformation (Correia, 1997). Early microscopists fancied that they could faintly make out these miniature forms within individual sperm cells. Of course, preformation presented a problem—if humans existed in miniature form before conception, when and how were these miniature forms themselves created? The problem leads to a regress, in which each sperm cell contains within it all possible future sperm cells, in an infinite series; in accordance with this, a 17th-century philosopher named Nicolas Malebrance proposed that Adam must have had in his testicles the preformed versions of every single human being that subsequently ever was born or would be born. This problem of infinite regress no longer exists, of course, now that we understand that the forms of humans and other animals emerge through merger of parent cells, followed by cell division and differentiation that occurs in a highly context-dependent fashion over the course of embryological development (Waddington, 1942).

Similar issues arise in the history of thinking about the origins of conceptual knowledge. Keil (1981) and others argued that conceptual knowledge consisted of theory-like structures. These theory-like structures provided the crucial scaffolding for the acquisition and elaboration of domain knowledge, on this view; without them, learning would be under-constrained; and therefore, Keil proposed that an initial set of proto-theories would have to be available from the outset—that is, innate. There are, of course, emergentist alternatives. As one example, Tim Rogers and I (Rogers & McClelland, 2004) have explored the possibility that some of the constraints on learning can themselves be learned; relatedly, it was our view that the constraints were not explicitly represented as they would be in an actual scientific theory but where instead embedded in the learned connections present in a general-purpose distributed connectionist network.

Four other topics in developmental cognitive science and developmental cognitive neuroscience also deserve at least brief consideration; I have grouped them into two subsets.

4.4.1. Stages and sensitive periods

Development is clearly not a completely continuous process. Piaget, of course, was famous for identifying developmental stages (Flavell, 1963), and although the broad stages that he envisioned have not held up, there remain good reasons to believe that children's cognitive abilities do not advance in a completely gradual and continuous fashion. Relatedly, there appear to be sensitive periods in a wide range of domains, including vision (ocular dominance) and language (especially for syntactic and phonological aspects of language if not for other aspects). Just what are the factors that are responsible for these effects? Recent approaches based on connectionist models have provided a way of seeing stage-like progressions as possible emergent consequences of a gradual learning process. In the early days of distributed connectionist models, I considered the developmental progress children make on a Piagetian task called the balance scale task (Siegler, 1976). In this work, I found that multilayer networks undergo accelerations and decelerations, exhibiting stage-like effects (McClelland, 1989). This work remains controversial; as a model that shares features with many other emergentist models, the transitions in it are not in fact completely abrupt; furthermore, around transitions in particular, performance in the model is graded and only approximately characterizable as characteristic of the stages others have seen in children's behavior. In recently revisiting these issues (Schapiro & McClelland, 2009), we found renewed support for the view I have held from the outset, namely that on close inspection of the data, there is evidence in children of exactly the kinds of graded effects that are seen in the model. Indeed, even stage theorists now speak in terms of ''overlapping waves'' instead of discrete transitions between stages (Siegler & Chen, 2002).

Critical periods in development (and subtler phenomena, including age of acquisition effects) are another area where emergence-based approaches have received considerable attention. A wide range of different ways of thinking about the basis of sensitive periods has been considered. Many modelers have proposed that reduced plasticity might not reflect a biological switch, but might be an emergent consequence of the accumulated effects of earlier experience (Flege, 1995; Munro, 1986; Vallabha & McClelland, 2007; Zevin & Seidenberg, 2004). Similarly, McMurray (2007) has shown how the vocabulary spurt in child development could reflect the simple cumulative consequences of experience. Though not quite a critical period phenomenon, it is also worth noting the work of Thelen, Fisher, and Ridley-Johnson (1984) on the disappearance of stepping behavior in infancy, which offers an emergentist alternative to the standard notion that this behavior disappears because of maturation of top-down inhibitory circuits. This work played a seminal role in the further development of dynamical systems approaches to development (Smith & Thelen, 2003; Thelen & Smith, 1996).

4.4.2. Cognitive modules and developmental disorders

The notion that mental (and corresponding neural) modules are not intrinsic or biologically preprogrammed as such, but emerge from a complex competitive and interactive process shaped by a wide range of forces, has become increasingly recognized as an important possibility. What exactly these forces are, and how strong any initial constraints may be, may vary from case to case. Crucial support for such a view comes from studies like those of Sur, Angelucci, and Sharma (1999) demonstrating that cortical areas that usually specialize in auditory processing can take on many of the properties of visual cortex if incoming connections are redirected so that visual instead of auditory input comes to these areas.

Two interesting cases that have been the focus of considerable interest are the so-called visual wordform area and the fusiform face area. The former surely must be the product of a convergence of forces rather than a prespecified module as such, given the recency (in evolutionary terms) of the introduction of visual word forms. Whether the latter is also an emergent structure (as Plaut, 2008, has proposed) or whether it and a few other evolutionarily privileged modules are strongly predetermined (Kanwisher & Yovel, 2009) remains a heated topic of debate. Whatever the details, surely *some* of the apparent physiological modularity that is seen in the brain is the emergent consequence of a complex interplay of forces.

In recent years, there has also been a very important shift in thinking about the basis of a wide range of developmental disorders, including specific language impairment (SLI), Williams syndrome, and others. In the early 1990s, it was possible for Pinker (1994) to claim that there was a specific gene that targeted the ability to learn regular but not exceptional morphology, popularly described as ''the grammar gene.'' The evidence Pinker pointed to in support of this particular claim (Gopnik & Crago, 1991) has not held up (Vargha-Khadem et al., 1998). Other work has found that individuals diagnosed with SLI may have greater problems with regular inflectional morphology, but this can occur either because of a system-wide change in an underlying parameter of the learning system (Thomas & Karmiloff-Smith, 2003) or because of difficulty processing complex and/or acoustically weak phonological material. Karmiloff-Smith et al. (2004) have made a similar argument in their analysis of the pattern of spared and impaired performance in children with Williams syndrome. They argue that the disorder should not be viewed as one of spared modules for one set of skills and impaired modules for other skills, but of a developmental trajectory resulting in different parameterizations of neural circuits, leading to a broad pattern of performance differences that belies any specificity with respect to putative modules such as ones for face processing or theory of mind.

4.5. Language structure and language change

Another very fertile domain for emergentist approaches is the evolution of language. Instead of thinking, as Chomsky (1965) proposed, that there exists a set of innate constraints on language in the form of Universal Grammar, many researchers have begun to explore instead the idea of language as an emergent process (Christiansen & Kirby, 2003). From an emergentist starting point, it is easy to envision how the structure of language might be shaped by a wide range of factors, including the sequential nature of speech, a tendency for temporal proximity to accompany relatedness, a pressure to keep messages simple, and a tendency toward faster and more fluent processing of items that are used more frequently. As one example, Bybee (2003) has proposed that the units we find in natural languages—phonemes, morphemes, words—and the changes that we see over time in these units, often resulting in the grammaticization of elements of meaning into such things as inflectional morphemes—are consequences of effects of usage on accessibility and fluency on articulation. Many other researchers have explored how characteristics of grammatical and phonological structure might arise as consequences of simple and very general constraints. This is a large and growing domain in which emergentist approaches are burgeoning.

As with the characteristics of the human species and the characteristics of English orthography, the characteristics of natural language may not in fact conform perfectly to any specific grammar. It was exciting in an earlier day to see just how much of the structure of English sentences Chomsky (1957) was able to capture with the small set of rules he introduced in *Syntactic Structures*. But it became clear very soon that the original approach was not fully adequate. The exploration of alternatives began, and of course such investigations continue. While continued progress along these lines may indeed occur, an emergentist approach may lead us to question whether in fact it makes sense to think that there really is a specific grammar or type of grammar that underlies real natural languages. Perhaps these too are best construed as emergent consequences of an interplay of forces that will render any specific characterization of structure only approximate. Culicover and Nowak (2003) is one example of a step in this direction.

4.6. Consciousness

Our discussion of emergence in cognitive science would not be complete without consideration of the emergence of consciousness, as consciousness seems clearly to be an emergent property. Just where does consciousness emerge? Is it a property of all matter? Of all biological systems? Of all multicellular organisms? Of all organisms with multicellular nervous systems, including molluscs? Of organisms with remote sensing systems (vision and/or audition), such as flies and bees? Of all vertebrates, but not invertebrates? Of only primates? Of only humans? It is interesting how little agreement there is on these questions. A related question: Where in the biological development of those organisms that possess it does consciousness arise? Is consciousness something that admits of matters of degree, or only of differences in content? Many authors, scientific and nonscientific, have considered these questions. One notion is that consciousness is a special kind of emergent property—a property that does arise somehow from merely physical and/or biological processes but is so complex in the way in which it arises that it will never be possible for science to understand exactly how it arises (Chalmers, 2006). Others view consciousness as little different in character from other emergent phenomena and pursue the possibility that the properties of consciousness are subject to scientific explanation in ways not fundamentally different from the investigation of other macro-properties of physical systems. While I tend to side with the latter perspective, it is worth reminding ourselves that there are other instances of emergent phenomena that may never be fully explained. Perhaps we will someday have a fuller understanding of the nature of conscious experience in general and of how such experience can in principle arise, without yet having a full understanding of its detailed properties in the case of human consciousness. This parallels the idea that we may someday achieve an understanding of the general nature of the processes through which physical, biological, and linguistic structures arise, yet this may not yield a full understanding of the basis for all the specific properties of particular structures.

5. The future directions and challenges for emergentist approaches in cognitive science

Throughout the previous section, I have suggested how various constructs in cognitive science, cognitive development, and cognitive neuroscience might best be seen, not as fundamental entities that serve as the basis of explanation, but as the emergent consequences of simpler, more general, or more basic processes. It seems fair to say that over the last 30 years or so, the tendency to see such entities as emergents has certainly increased. However, this does not mean that we now understand these emergent phenomena. I believe there has been real progress in some domains, at least in developing models that account for quite a bit of experimental data. Yet even in such cases, this does not necessarily mean that we really understand how the phenomenon occurs; as Holland (1975) explores at length, models governed by simple laws, especially models that learn, can give rise to emergent patterns that were not predicted in advance and that, even after they have been observed, may not be fully comprehended by the modeler. To what extent such patterns can or ever will be fully understood remains an open question.

I think the greater challenge will be to extend emergentist approaches so that they address more fully the achievements of human intelligence of the sort I mentioned in the introduction. Most of the models I have described address relatively simple tasks (single-word reading) or else rely on highly simplified versions of the cognitive tasks humans are capable of performing. Neither symbolic nor emergentist approaches have yet succeeded in capturing everyday cognitive abilities in real-world context, much less emulating great insights and artistic accomplishments.

New fields and new approaches have sprung up to address some of these challenges. As one example, there is now a branch of robotics that focuses on the problem of autonomous mental development. This emerging discipline arises from the belief that artificially intelligent systems cannot be programmed in advance, but must instead be created out of an epigenetic process, one in which, through their emerging abilities, such systems largely organize their own development (Weng et al., 2001). It is a dream that work of this kind may someday lead to synthetic systems with emergent cognitive functions equaling our everyday abilities to perceive, remember, and act appropriately in our environments, and may even someday lead to abilities that rival the accomplishments of great scientific and artistic minds. Up to now such approaches have not been strongly represented in the Cognitive Science Society. Perhaps a greater degree of engagement with these approaches will arise over the course of the next 30 years of the Society's history.

Even if such approaches lead us toward synthetic systems with ever-greater intelligence, we should not lose sight of the very real possibility that a full understanding of the emergent properties of both real and synthetic adaptive intelligent systems will still remain elusive. Holland (1975) has considered neural network models and models based on genetic algorithms and their emergent properties. The properties of such systems can be fully specified by the modeler, up to, say, the random initial values of connection weights or other parameters of the model system; likewise, the rules by which these systems adapt to experience may also be fully specified (either completely deterministically or subject to some degree of randomness). The modeler may even completely specify the regime of experiences to which such a system is subjected. Even so, a full understanding of the properties these specified conditions give rise to often remains elusive. Holland exhibits optimism that a fuller understanding may be possible. Yet there remains the possibility that it may never be possible to succinctly characterize all aspects of the evolved computational properties that arise in such systems. Might it be likewise for our efforts to understand the nature of human insight and other cognitive abilities? Perhaps the next 30 years of research in cognitive science will provide the answer to this question.

Acknowledgments

I thank Larry Barsalou for encouragement and constructive comments on an earlier version of this manuscript.

References

Anderson, J. A., Silverstein, J. W., Ritz, S. A., & Jones, R. S. (1977). Distinctive features, categorical perception, and probability learning: Some applications of a neural model. *Psychological Review*, *4*, 413–451.

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*, *45*, 2883–2901.

Barwise, J., & Etchemendy, J. (1991). Visual information and valid reasoning. In W. Zimmerman & S. Cunningham (Eds.), *Visualization in mathematics* (pp. 9–24). Washington, DC: Mathematical Association of America.

Bates, E. (1976). *Language and context: Studies in the acquisition of pragmatics*. New York: Academic Press.

Botvinick, M., & Plaut, D. C. (2004). Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review*, *111*, 395–429.

Bowers, J. S. (2009). On the biological plausibility of grandmother cells: Implications for neural network theories in psychology and neuroscience. *Psychological Review*, *116*, 220–251.

Braitenberg, V. (1984). *Vehicles: Experiments in synthetic psychology*. Cambridge, MA: MIT Press.

Bunge, M. (2003). *Emergence and convergence*. Toronto: University of Toronto Press.

Bybee, J. (2003). Mechanisms of change in grammaticization: The role of frequency. In R. D. Janda & B. D. Joseph (Eds.), *Handbook of historical linguistics* (pp. 602–623). Oxford, England: Blackwell Publishers.

Chalmers, D. (2006). Strong and weak emergence. In P. Clayton & P. Davies (Eds.), *The re-emergence of emergence* (pp. 244–254). New York: Oxford University Press.

Chomsky, N. (1957). *Syntactic structures*. The Hague, The Netherlands: Mouton.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Christiansen, M., & Kirby, S. (Eds.) (2003). *Language evolution*. New York: Oxford University Press.

Cole, M., & Scribner, S. (1974). *Culture and thought: A psychological introduction*. New York: Wiley.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256.

Correia, C. P. (1997). *The ovary of eve: Egg and sperm preformation*. Chicago: University of Chicago Press.

Culicover, P. W., & Nowak, A. (2003). *Dynamical grammar*. New York: Oxford University Press.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.

Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Flavell, J. (1963). *The developmental psychology of Jean Piaget*. Oxford, England: D. Van Nostrand.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 233–277). Timonium, MD: York Press.

Fodor, J., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*, 3–71.

Fuster, J. M. (1997). *The prefrontal cortex-anatomy physiology, and of neuropsychology of the frontal lobe* (3rd ed.). Philadelphia: Lippincott-Raven.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 674–691.

Goodman, N. D., Mansighka, V. K., Roy, D., Bonawitz, K., & Tenenbaum, J. B. (2008). Church: A language for generative models. *Uncertainty in Artificial Intelligence*, *24*, 222–229.

Gopnik, M., & Crago, M. (1991). Familial aggregation of a developmental language disorder. *Cognition*, *39*, 1–50.

Haken, H. (1977). Synergetics, an introduction. Nonequilibrium phase-transitions and self-organization in physics, chemistry and biology. Berlin: Springer.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychologicial theory*. New York: Wiley.

Hinton, G. E. (1981). Implementing semantic networks in parallel hardware. In G. E. Hinton & J. A. Anderson (Eds.), *Parallel models of associative memory* (pp. 161–187). Hillsdale, NJ: Erlbaum.

Hinton, G. E., McClelland, J. L., & Rumelhart, D. E. (1986). Distributed representations. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Volume I, Chapter 3, pp. 77–109). Cambridge, MA: MIT Press.

Hofstadter, D. R. (1979). *Gödel, Escher, Bach: An eternal golden braid*. New York: Basic.

Hofstadter, D. R., & the Fluid Analogies Research Group (1995). *Fluid concepts and creative analogies. Computer models of the fundamental mechanisms of thought*. New York: Basic Books.

Holland, J. (1975). *Adaptation in natural and artificial systems*. Cambridge, MA: MIT Press.

Holland, J. H. (1998). *Emergence: From chaos to order*. Reading, MA: Addison-Wesley.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences (USA)*, *79*, 2554–2558.

Jackendoff, R. (2007). Linguistics in cognitive science: The state of the art. *The Linguistic Review*, *24*, 347–401.

Jirsa, V. K., & Kelso, J. A. S. (Eds.) (2004). *Coordination dynamics: Issues and trends*. Berlin: Springer.

Johnson, S. (2001). *Emergence*. New York: Scribner.

Kanwisher, N., & Yovel, G. (2009). Face perception. In G. G. Bernston & J. T. Cacioppo (Eds.), *Handbook of neuroscience for the behavioral sciences*. Volume 2 (pp. 841–858). Hoboken, NJ: J. Wiley and Sons.

Karmiloff-Smith, A., Thomas, M., Annaz, D., Humphreys, K., Ewing, S., Brace, N., Duuren, M., Pike, G., Grice, S., & Campbell, R. (2004). Exploring the Williams syndrome face-processing debate: The importance of building developmental trajectories. *Journal of Child Psychology and Psychiatry*, *45*, 1258–1274.

Keil, F. C. (1981). Constraints on knowledge and cognitive development. *Psychological Review*, *88*, 197–227.

Kemp, C., & Tenenbaum, J. B. (2009). Structured statistical models of inductive reasoning. *Psychological Review*, *116*, 20–58.

Lachter, J., & Bever, T. G. (1988). The relation between linguistic structure and associative theories of language learning. *Cognition*, *28*, 195–247.

Lashley, K. (1950). In search of the engram. *Symposian of the Society of Experimental Biology*, *4*, 2–31.

Lewes, G. H. (1875/2005). *Problems of life and mind*. First Series. Michigan Series: Scholarly Publishing Office, University of Michigan Library.

MacWhinney, B., & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the Verb Learning Model. *Cognition*, *40*, 121–157.

Marcus, G. F. (2001). *The algebraic mind: Integrating connectionism and cognitive science*. Cambridge, MA: MIT Press.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, *58*, 25–45.

McClelland, J. L. (1981). Retrieving general and specific information from stored knowledge of specifics. *Proceedings of the third annual meeting of the cognitive science society* (pp. 170–172). Berkeley, CA.

McClelland, J. L. (1989). Parallel distributed processing: Implications for cognition and development. In R. G. M. Morris (Ed.), *Parallel distributed processing: Implications for psychology and neurobiology* (pp. 8–45). Oxford, England: Clarendon Press.

McClelland, J. L. (2010). Memory as a constructive process: The parallel-distributed processing approach. In S. Nalbantian, P. Matthews, & J. L. McClelland (Eds.), *The memory process: Neuroscientific and humanistic perspectives* (pp. 129–151). Cambridge, MA: MIT Press.

McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C., Rogers, T. T., Seidenberg, M. S., & Smith, L. B. (2010). Letting structure emerge: Connectionist and dynamical systems approaches to understanding cognition. *Trends in Cognitive Sciences*, *14*, 348–356.

McClelland, J. L., & Bybee, J. (2007). Gradience of gradience: A reply to Jackendoff. *The Linguistic Review*, *24*, 437–455.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*, 419–457.

McClelland, J. L., & Patterson, K. (2002a). Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences*, *11*, 465–472.

McClelland, J. L., & Patterson, K. (2002b). 'Words or Rules' cannot exploit the regularity in exceptions. *Trends in Cognitive Sciences*, *11*, 464–465.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of the effect of context in perception, Part I. An account of basic findings. *Psychological Review*, *88*, 375–407.

McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science*, *317*, 631.

Medin, D. L., & Shaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*, 207–238.

Minsky, M. (1980). K-Lines: A theory of memory. *Coginitive Science*, *4*, 117–133.

Minsky, M. (1986). *The society of mind*. New York: Simon and Schuster.

Morowitz, H. J. (2002). *The emergence of everything: How the world became complex*. New York: Oxford University Press.

Munro, P. (1986). State-dependent factors influencing neural plasticity: A partial account of the critical period. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Volume II, Chapter 24, pp. 471–502). Cambridge, MA: MIT Press.

Newell, A. (1980). Physical symbol systems. *Coginitive Science*, *4*, 133–183.

Page, M. (2000). Connectionist modelling in psychology: A localist manifesto. *Behavioral and Brain Sciences*, *23*, 443–467.

Pinker, S. (1994). *The language instinct*. New York: William Morrow.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, *28*, 73–193.

Pinker, S., & Ullman, M. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, *6*, 456–463.

Plaut, D. C. (2008). Face and word processing: Two sides of the same brain. Plenary address presented at the 30th annual meeting of the Cognitive Science Society, Washington, DC, July 25, 2008.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56–115.

Quammen, D. (2006). *The reluctant Mr. Darwin*. New York: Norton.

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review*, *108*, 370–392.

Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986a). Learning representations by back-propagating errors. *Nature*, *323*, 533–536.

Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tenses of English verbs. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Volume II, Chapter 18, pp. 216–271). Cambridge, MA: MIT Press.

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 99–135). Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986b). Parallel distributed processing models of schemata and sequential thought processes. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Volume II, Chapter 14, pp. 7–57). Cambridge, MA: MIT Press.

Rumelhart, D. E., & Todd, P. M. (1993). Learning and connectionist representations. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 3–30). Cambridge, MA: MIT Press.

Schapiro, A. C., & McClelland, J. L. (2009). A connectionist model of a continuous developmental transition in the balance scale task. *Cognition*, *110*(1), 395–411.

Schöner, G., & Kelso, J. A. S. (1988). Dynamic pattern generation in behavioral and neural systems. *Science*, *239*, 1513–1520.

Sejnowski, T. J., & Rosenberg, C. R. (1987). Parallel networks that learn to pronounce English text. *Complex Systems*, *1*, 145–168.

Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology*, *8*, 481–520.

Siegler, R. S., & Chen, Z. (2002). Development of rules and strategies: Balancing the old and the new. *Journal of Experimental Child Psychology*, *81*, 446–457.

Smith, L. B., & Thelen, E. (2003). Development as a dynamical system. *Trends in Cognitive Sciences*, *7*, 343–348.

Smolensky, P. (1986). Information processing in dynamical systems: Foundations of Harmony Theory. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Volume I, Chapter 6, pp. 194–279). Cambridge, MA: MIT Press.

Spoehr, K. T., & Smith, E. E. (1975). The role of orthographic and phonotactic rules in perceiving letter patterns. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 21–34.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, *99*, 195–231.

Sur, M., Angelucci, A., & Sharma, J. (1999). Re wiring cortex: The role of patterned activity in development and plasticity of neocortical circuits. *Journal of Neurobiology*, *41*, 33–43.

Thelen, E., Fisher, D. M., & Ridley-Johnson, R. (1984). The relationship between physical growth and a newborn reflex. *Infant Behavior & Development*, *7*, 479–493.

Thelen, E., & Smith, L. (1996). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Thomas, M. S. C., & Karmiloff-Smith, A. (2003). Modeling language acquisition in atypical phenotypes. *Psychological Review*, *110*, 647–682.

Turvey, M. T. (2004). Impredicability, dynamics, and the perception-action divide. In V. K. Jirsa & J. A. S. Kelso (Eds.), *Coordination dynamics: Issues and trends*. Berlin: Springer.

Usher, M., & McClelland, J. L. (2001). On the time course of perceptual choice: The leaky competing accumulator model. *Psychological Review*, *108*, 550–592.

Vallabha, G. K., & McClelland, J. L. (2007). Success and failure of new speech category learning in adulthood: Consequences of learned Hebbian attractors in topographic maps. *Cognitive, Affective and Behavioral Neuroscience*, *7*, 53–73.

Vargha-Khadem, F., Watkins, K. E., Price, C. J., Ashburner, J., Alcock, K. J., Connelly, A., Frackowiak, R. S. J., Friston, K. J., Pembrey, M. E., Mishkin, M., Gadian, D. G., & Passingham, R. E. (1998). Neural basis of an inherited speech and language disorder. *Proceedings of the National Academy of Sciences (USA)*, *95*, 12695–12700.

Venesky, R. (1970). *The structure of English orthography*. The Hague, The Netherlands: Mouton.

Waddington, C. H. (1942). The epigenotype. *Endeavour*, *1*, 18–20.

Wang, X.-J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*, *36*, 955–968.

Weng, J., McClelland, J. L., Pentland, A., Sporns, O., Stockman, I., Sur, M., & Thelen, E. (2001). Autonomous mental development by robots and animals. *Science*, *291*, 599–600.

Wikipedia, (2010). *Emergence*. Available at: http://en.wikipedia.org/wiki/Emergence. Accessed January 9, 2010.

Zevin, J. D., & Seidenberg, M. S. (2004). Age-of-acquisition effects in reading aloud: Tests of cumulative frequency and frequency trajectory. *Memory & Cognition*, *32*, 31–38.