Origins of Order in Cognitive Activity

Geoff Hollis

Heidi Kloos

Guy C. Van Orden

University of Cincinnati

Author note: We gratefully acknowledge funding from the National Science Foundation to Heidi Kloos and Guy Van Orden. We thank Patricia Carpenter, Ray Gibbs, Larry Gottlob, and Damian Stephens for comments and other help. Editorial suggestions by David Pincus and Stephen Guastello consistently improved the chapter.

Correspondence:

Guy Van Orden
Department of Psychology
ML 0376, 429 Dyer Hall
University of Cincinnati
Cincinnati, OH 45221-0376

Origins of Order in Cognitive Activity

Most cognitive scientists have run across *The War of the Ghosts*, a Native American story used by Bartlett (1932) in his classic studies of remembering. British college students read the story twice, and recalled it in detail after 15 minutes, hours, days, months or years, “as opportunity offered” (p. 65). The compelling finding was that participants re-interpreted parts of the story, in addition to omitting details. The mystical story was reorganized and changed in the retelling to fit cultural norms of the British participants. In other words, errors in retelling the story were neither random nor arbitrary but fit together within a larger created narrative. The memory errors illustrate the ordinary constructive performance of cognition, and the creation of orderly and sensible thought. Despite perpetually moving eyes, swaying body, and ambiguous stimuli, people perceive coherent and orderly objects. Despite the lack of explicit links between events, higher-order cognition fits thought and behavior within larger coherent narratives. However, the origin of such order remains a mystery. What is the basis of orderly thought, memory, speech, and other cognitive abilities?

The origin of order in cognition is the topic of this chapter. We begin with a discussion of how order is explained within a traditional approach of information processing. Taking the shortcomings of this account seriously, we then turn to other disciplines – those that have framed the question of order more successfully. The answers have relied upon the concept of self-organization, the idea that order can emerge spontaneously from the non-linear interaction of a system’s components. In the remainder of the chapter, we discuss empirical evidence for self-organization in cognition. The accumulated evidence in reasoning, speaking, listening, reading, and remembering motivates a complex system approach to cognition.

Order and Information Processing

The development of the computer promised a workable metaphor for human cognition. A computer's flow of information processing appears similar to cognition: information comes in (it is transduced into neural signals), information is manipulated (it is perceived, remembered, reasoned about), and an output is produced (overt behavior takes place). Conveniently, the orderly workings of a computer are transparent. Programs, memory storage, and peripheries are arranged in a systematic way. Thus a reasonable abduction is to equate the workings of cognition to the workings of a computer's software. As Neisser (1968) suggested, "The task of a psychologist trying to understand human cognition is analogous to that of a man trying to discover how a computer has been programmed." (p. 6).

The computer perspective has flourished and still persists. Cognitive activity is often seen as a step-by-step process of detecting, combining, storing, retrieving, and outputting information. In line with this framework, questions pertain to the duration of information processing (e.g., Sternberg, 1969**),** the capacity of memory systems (e.g., Miller, 1956), low-level vs. high-level detection of information (e.g., Marr, 1982), the properties of the central executive (e.g., Allport, 1989), and the specific format of symbolic representations (e.g., Fodor & Pylyshyn, 1981). Accordingly, order is generated through symbol manipulations, "recipes for selecting, storing, recovering, combining, outputting, and generally manipulating [symbols]" (Neisser, 1968, p. 7).

An information processing account envisions the complexities of perceiving, remembering, and thinking as hierarchies of functional components, wherein each component solves a simpler part of cognition. Depending on the requirements of the task, components include instructions to manipulate internal symbols and communicate solutions to other components. Internal symbols such as schemas and representations are themselves the outputs of lower level components such as edge detectors, acoustic feature detectors, or detectors of other primitive elements. Higher-level executive functions also include instructions that determine planning, selective attention, and deciding among alternatives. Executive functions organize cognition to follow explicit task instructions, for example, by retrieving the appropriate representations and planning the action.

This somewhat compressed summary of the information-processing view makes one point clear on the origins of ordered cognition: Overt order in cognitive activities stems from internal components that are themselves ordered. We perceive a stable world because we have stable feature detectors. We perform systematically in a categorization task, because we have a stable representation of the category, and we have stable detectors for recognizing category members. And we make errors, for example, when asked to recall an unusual story, because there is a mismatch between the presented stimulus and stable schemas. Behavioral order is equated with intrinsic sources of order, a priori order that exists in the components of cognition. Similar to the functioning of a clock, ordered behavior comes from the interaction of ordered parts.

The obvious advantage of holding this view is that internal components, the ordered parts of the mind, can be described in detail. We can ask such questions as: How do children represent numbers? What is the format of the mental lexicon? Or how do adults represent cause and effect? This view makes it fairly easy to control behavior. If stable components cause behaviors, particular components can be changed to yield the desired behavior. Merely changing the appropriate component could correct erroneous behavior. Why then are we focused on the shortcomings of this account?

Limits of the Information Processing Account

Early on, Neisser (1968) noted puzzling discrepancies between peoples' behavior and a computer's output. For instance, some computations (e.g., many-digit multiplication) are trivially simple for a computer, but exceedingly difficult or impossible for people. Other tasks are immediate and simple for a person (e.g., self-guided locomotion), but exceedingly difficult or impossible for computers. These differences cannot be reconciled easily. Claiming limited capacity in humans could explain poor performance of many-digit multiplications, but it could not explain the relative ease with which even a crawling infant navigates a terrain. Claiming limited complexity in computers could explain a computer's inability to perceive and act in a naturally complex environment, but it could not explain the computer's ease with complex calculations.

Such obvious discrepancies supply a first indication that the computer metaphor has serious limitations for explaining the order in cognition. These and other discrepancies between computations and common sense subsequently were called the *frame problem* (e.g., see Haselager, Bongers, & van Rooij, 2003). However, even if we could set aside the frame problem, important questions remain: How can relatively static components explain the uniqueness of thought? How can ordered components explain context sensitivity? And how does order get into the components in the first place? Each of these concerns is addressed in turn.

Uniqueness of thought. Information processing takes for granted that the same conditions produce the same output time and time again. But a central finding of empirical cognition is that cognition is always under construction and rarely repeats itself exactly. As Williams James noted: "A permanently existing idea … is as mythological an entity as the Jack of Spades" (1890, Vol. 1, p. 236). An on-coming thought is almost never a mere repetition of a previous thought; and the same idea communicated more than once by the same person will be conveyed in different sentences each time. To do otherwise creates added idiomatic meaning; repetitious speech acts are seen as stubborn, pedantic, sarcastic or ironic (Gibbs, 1999). Even a repetitive motion, such as swinging a hammer or scratching an itch, will express unique kinematics within each repetition (Berkinblit, Feldman, & Fukson, 1986; Bernstein, 1967). As Neisser explains, "exact repetitions of earlier acts or thoughts are the exception, extremely difficult to achieve, and ascribed to long practice or neurotic defensiveness" (Neisser, 1968, p. 282).

One strategy to address the problem of uniqueness is to add random variation among component outputs, such as adding white noise. Some modern information processing models have sources of randomness built in (e.g., ACT-R; Anderson, Lebiere, Lovett, & Reder, 1998). Indeed, successive outputs of such models do not just repeat the same solution; they vary from iteration to iteration. However there is no evidence to support this functional role for random cognitive noise. In fact, no matter how unconstrained the task, variability in repeated measures of cognitive performance is not simply random (Gilden, 2001). Take the perception of a Necker cube as an example (Necker, 1932; see Figure 7.0). If the display is sufficiently ambiguous, a person's percepts will switch among possible percepts, from one to another and back again, as time passes. If random processes govern switching among percepts, then the time between switches should exhibit white noise. This is not what is found, however. The time series of time between switches displays *pink noise* (Aks & Sprott, 2003), a kind of noise that aliases a deterministic, interdependent system, not random perturbations.

Insert Figure 7.0 about here

Other information processing models use learning instead of noise to account for the uniqueness of behavior (e.g., Sutton & Barto, 1998). A system that iteratively changes itself based on its interactions with an environment has the potential to display similar, but nonrepeating behavior over time. Learning as the basis for uniqueness is plausible, and it is consistent with the idea of cognition being constructive. However, examples of non-iterative, single trial learning (e.g., Rock, 1957) suggest that there are qualitative differences between human learning and current machine learning. Uniqueness is not always due to incomplete, ongoing, incremental learning. Moreover, although some learning is intuitively asymptotic, massed practice of a precision aiming motor task (Wijnants, Bosman, Hasselman, Cox, & Van Orden, 2008), for example, and development of human gait (Hausdorff, Zemany, Peng, & Goldberger, 1999) both seem to asymptote on the pink noise mentioned above. Thus although learning might contribute to uniqueness, it does not exhaust the facts of the matter.

Finally, another conventional way to address uniqueness is to supply a unique mental structure for each cognitive act that qualifies as sufficiently unique. Such an approach has lead to a variety of distinctions, including implicit vs. explicit processes, declarative vs. procedural knowledge, and long term vs. short term memory. But an inherent problem comes with this solution as well. The general basis for any particular distinction is indeterminate (Shallice, 1988; Van Orden, Pennington, & Stone, 2001). How large, or how reliable, or what kind of a difference must exist in behavior before separate mental structures are justified? One can find distinct and uniquely ordered behavior on a long time scale such as planning a vacation or a retirement. And one can find distinct and uniquely ordered behavior on a short time scale such as hammering a nail or scratching an itch. Do all of these distinctions necessitate distinct mental structures? Clearly, accounting for uniqueness by resorting to a-priori ordered components raises more questions than it provides answers.

Context dependence of cognition. Information processing accounts assume that ordinary perception and cognition originate in atomic components, feature detectors of some sort that combine through internalized rules to create cognition. Small or superficial changes in context should therefore not change the cognition. However, cognition is not only unique but also strongly affected by small changes in context. Adults' performance of strictly mathematical problems changes as a function of their superficial spatial arrangements (Landy & Goldstone, 2007). The well-documented A-not-B search error disappears in infants who are briefly lifted from their sitting position before the toy is hidden at position B (Smith, Thelen, Titzer, & McLin, 1999).

Likewise, children's well-documented difficulty to distinguish between sinking and floating objects disappears when the distribution of experimental objects is altered (Kloos, Silverman, & Fenstermaker, 2007). And well-documented syndromes of brain damage or developmental disabilities change, disappear, or even "reverse" to become an opposite syndrome after changes in the method of observation (Van Orden et al., 2001). For example, the same patient exhibits telegraphic speech under some task conditions, a symptom of agrammatic or Broca's aphasia, and morphological substitutions under different task conditions, a symptom of paraphasia (Hofstede & Kolk, 1994; Kolk & Heeschen, 1992) – but these aphasias (symptoms) are sometimes put forward as opposite syndromes, composing a double dissociation.

One way to explain context-dependent cognition is to assume that the stable components of information processing combine in context-unique arrangements for each instance of cognitive activity, as situated expectations for instance (Mandler, 1984). If so, however, there is no reliable basis by which to distinguish between cognitive components and methodological contexts. Where atomic components have been most closely examined, for example, they are reliably different in different task contexts. Atomic components are sufficiently context sensitive that they are impossible to dissociate from the contexts of their discovery (Goldinger & Azuma, 2003; Van Orden & Kloos, 2005). In other words, no data exist that reliably pick out cognitive components separate or independent of the contexts in which they are observed.

What is the origin of order in components? Our discussion on the uniqueness of cognition and its context-dependence already hints at problems in reducing the global order of behavior to underlying ordered components. A final problem then pertains to the question about how order gets into underlying components in the first place. For example, what is the source of order in participants' schemas that affects their recall of *The War of the Ghosts*? What is the source of order in representations about a category or a concept? And what is the source of order in intentional thought and purposeful behavior? Reducing the order of cognitive activity to ordered information processing components leads to a sort of dead-end, making it impossible to explain the ultimate origin of order without homunculi or empirically opaque automatic components. Such an approach leaves psychology with no eventual answer for the origins of order (Juarrero, 1999; Kugler & Turvey, 1987; van Gelder & Port, 1995). Consequently psychology inevitably sees either a stimulus or a homunculus-like executive process as the *prime mover* of behavior (Oyama, 2000; Shaw, 2001).

In the remainder of this chapter, we describe how the complex systems approach has reframed the question of the origin of order. The pivot point of complex systems is self-organization, the emergence of order from unordered parts. We expand on this point to supply answers to the important questions identified by Neisser: "How is the raw material [of constructions] organized? How is the process of construction organized? What determines what is constructed? And what purpose does it serve?" (Neisser, 1968, p. 280).

Emergence of Order in Complex Systems

Questions about order and its origins are asked in a larger sphere than just psychology. Other disciplines have been haunted by similar problems of trying to explain order without having to stipulate its existence a priori in the parts of the system. In contemporary physics and biology, for instance, order has become something to be understood (e.g., see Depew & Weber, 1995). Its origins no longer appear so mysterious. The synthesis of shared questions of order has yielded the contemporary science of complexity. It concerns a meta-disciplinary nonlinear dynamical systems view in which the same principled origins of order are expressed in very different material systems. A number of conceptual building blocks lie within this approach, laws that govern the emergence of stable behavior. Of course, additional necessary laws of order in living systems remain to be discovered (Kauffman, 2000).

Cognitive scientists have availed themselves of the complexity view and imported theoretical concepts that come out of it. One of these concepts is self-organization. Applied to cognitive activity, self-organization concerns order in the larger system of organism and environment, order that reduces neither to homunculi nor to stimuli, nor to pieces and bits of a cognitive architecture. Instead, it is based on self-sustaining processes that exploit gradients of uncertainty. In this section we explain self-ordering and self-sustaining processes in more detail, discussing in particular their relation to traveling waves, constraints, and critical states.

Traveling waves. Self-perpetuation is a prominent characteristic of physical and biological structures. Convection perpetuates a tornado through its relatively brief lifespan, for instance. The heart supplies blood and nutrients to the body and itself, allowing its structure to exist through time. Neighboring brain cells supply mutual life support (which is why more cells die in brain trauma than are killed directly). In effect, metabolism is the primary function of the brain and body.

If metabolism is the primary function of a nervous system, then an elegant theory would be one in which cognitive activity emerges out of metabolism. Such a theory would begin to bridge the chasm between laws of physical processes and cognition. Davia (2005) outlines such a theory based on *autocatalytic reactions,* which are fundamental metabolic processes. Autocatalytic reactions are chemical reactions in which a catalyst, an enzyme for example, accelerates a chemical reaction while remaining unchanged. Enzymes catalyze the biochemical reactions in metabolism that are necessary to sustain life, develop, and reproduce.

Davia (2005) equates enzymes with self-perpetuating structures called traveling waves or solitons. Traveling waves are the basis in the nervous system for perception and action. To relate metabolism to perception and action, Davia argues that the nervous system functions as an excitable medium. An excitable medium is a landscape containing energy that can be consumed and replenished. A field of grass may be conceived as an excitable medium. It can be grazed but also replenished with nutrients and sunlight. The brain and body too may be conceived as excitable media, where glucose is consumed and re-supplied for instance. Traveling waves are temporally invariant structures that exist as coherent, ordered entities within the space and energy of an excitable medium.

Consider traveling waves in olfactory perception, the dominant mode of perception in most animals. A rabbit, conditioned to respond to banana oil, will exhibit a traveling wave across its olfactory bulb as it inhales banana oil. The wave itself comprises a complex pattern of amplitude-modulated activity; neurons firing at particular amplitudes compose the wave structure. The wave is context sensitive to the extent that learning a new unrelated association to sawdust will change the amplitude wave profile to banana oil. The pattern of the amplitude wave pertains to what the odor affords for the animal; it is not a representation of smell. For example, if banana oil is a conditioned stimulus that subsequently serves as an unconditioned stimulus, as the same animal acquires the conditioned stimulus sawdust, then the banana-oil traveling wave pattern will stay with sawdust, the new conditioned stimulus. Finally, all else equal the wave is subject to long term drift in structure at the pace of hair or fingernail growth. The same findings have been corroborated for other senses; "all the central sensory systems use essentially the same dynamics" (Freeman, 2000; p. 88).

Traveling waves are seen in many forms. Tsunamis are prime examples. They propagate a fixed structure and its associated energy across long distances, even when faced with obstructions. Forest fires are traveling waves that perpetuate their shape outward – at least until confronted with non-flammable terrain or asymmetrically dense, flammable terrain. So is the unwinding of DNA before transcription, as are the movements of a millipede's legs. In the case of the millipede, the traveling wave allows it to move, while also unifying perception and action. This unity poises the millipede to react seamlessly with new movements that may be required.

Traveling waves unify energy and structure in that they perpetuate a fixed form and energy, and do not dissipate easily. In classical mechanics, when a force is applied to a physical structure, its energy is quickly dissipated. For traveling waves this is not true. Consequently, for the millipede, unification of perception and action explains how a reaction can be immediate. Human perception and action can also be immediate. In particular, ultra-fast cognition has been observed in which perception or action occur so quickly, as to leave little or no time for information processing (e.g., see Fabre-Thorpe, Delorme, Marlot, & Thorpe, 2001; Grill-Spector & Kanwisher, 2006; Petterson, 1994; Rauschenberger, Peterson, Mosca, & Bruno, 2003).

Traveling waves maintain their own structure across time and provide a continuous capacity for action (for they are both perpetuating structures of physical material *and* energy). As such, they may provide a basis for cognition without having to refer to fixed static cognitive components (Davia, 2005). Furthermore, traveling waves are demonstrable entities, not hypothetical constructs such as component functions. In biological entities specifically, traveling waves have been observed in enzyme catalysis (Sataric, Zakula, Ivic, & Tuszynski, 1991), DNA (Englander, Kallenbach, Heeger, Krumhansl, & Litwin, 1980; Yakushevich, 2001), heart functioning (Beaumont, Davidenko, Davidenko, & Jalife, 1998), nerve action potential (Aslanidi & Mornev, 1996), the basilar membrane (Duke & Julicher, 2003), the brain (Koroleva & Bures, 1979), muscle contraction (Davydov, 1979), population dynamics (Odell, 1980), and spontaneous traveling waves in early development organize auditory cortex, neocortex, the hippocampus, spinal cord networks, brainstem nuclei, and the retina (Godfrey & Swindale, 2007).

Constraints. Traveling waves do not reduce to components of the nervous system; they are emergent phenomena that depend instead on the existing balance of constraints. Constraints are accrued through idiosyncratic experience, and constraints are implicit in the immediate context. Constraints are aspects of biology, culture, history, context, or current states that narrow down the possibilities for cognitive activity, prior to its occurrence (see also Mandler, 1997). They are necessary to insure ordered task performance because there are too many degrees of freedom otherwise (Bernstein, 1967; Turvey, 1990).

Consider adults' memory performance in Bartlett's (1932) *The War of the Ghosts* experiment. The mystical content of the story and its unfamiliar cultural norms provided fewer constraints than what a British story might have supplied to a British reader. The British reader whose experiences did not match those of Native Americans had available many degrees of freedom for interpretation and subsequent recall of the story. Under these circumstances of unfamiliar norms and loosened cultural constraints, participant idiosyncrasies were more fully expressed. While every participant recalled a coherent story, the way in which this was done varied greatly from one participant to the next. In Bartlett's words, "the particular form adopted [was] due directly to the functioning of individual special interests … or to some fact of personal experience, or to some peculiarity of individual attitude which determines the salience or potency of the details in the whole material dealt with" (Bartlett, 1932, p. 71).

It is also possible to tighten contextual constraints. A scenario that supplies sufficiently tight constraints should yield performance that is identical across many participants. For example, when presented with a list of words such as "bed," "rest," and "awake," many adults falsely recalled the term "sleep," a related term that was not presented in the list. This memory paradigm is so reliably constraining that participants falsely produced predicted terms on about 50% of recall opportunities and falsely recognized such terms at almost the same rate as hit-rates for presented terms (Roediger & McDermott, 1995). In the face of this probabilistic outcome, it goes well beyond the facts to invoke causal properties of 50% of participants' sleep schemas. Instead, the methodologically constrained context (providing many words related to the concept sleep) creates the potential for false positive sleep memories contingent on idiosyncratic states of mind of participants.

Constraints combine to reduce the degrees of freedom for behavior and thereby increase the likelihood of the behaviors that remain. Note, however, that constraints do not combine additively. Take for example the proportion of false positives to target items such as "sleep" after studying a list of related words. Augmenting the list of semantically related words (e.g.," bed," "rest," "awake") with up to three additional semantic relatives does not change the proportion of false positives. But augmenting the same list with only one word similar in phonology (e.g., “keep”) sharply increases the proportion of false positives. And adding three phonologic relatives doubles the proportion of false positives (Watson, Balota, & Roediger, 2003).

In another example of non-additive and nonlinear effects, participants were presented with a list of homophones such as “paws,” after which they were given a surprise recognition task of old and new homophones (e.g., “paws” and “pause,” respectively). The proportion of false positives (judging “pause” to be old) was measured. This condition was contrasted with *memory-load* conditions, containing words similar in phonology and spelling to either the old homophone (e.g., “jaws”) or to the new homophone (e.g., “cause”). As expected, false positives were highest in the latter condition, when memory-load words shared the phonology of the new homophones (Azuma, Williams & Davie, 2004). But memory-load words by themselves (e.g., “cause”) produced no false positive effect (e.g., of “pause”) above chance. “Cause” on its own did not change the degrees of freedom for responding; but in combination with “paws” it did making false positives to “pause” more likely (cf. Humphreys, Burt, & Lawrence, 2001).

Importantly, constraints do not cause false positives or behavior in general. Behavior does not directly originate in constraints and is not directly determined by constraints. Instead constraints change the likelihood of behaviors, including false positives, as they change the potential set for responding. Constraints merely narrow this potential set. By contrast, as new capacities become available, constraints are relaxed and the degrees of freedom for behavior are increased. Braitenberg (1984) illustrated this fact when he imagined robots or vehicles that differed in their capacities for movement, their sensory capacities, their wiring complexity, and their learning capacities. Adding capacities incrementally allowed him to deduce the new behaviors that were brought on line.

New capacities reduce constraints and increase the degrees of freedom for behavior. More contexts may become available in the process. Adding constraints reduces the degrees of freedom and the available contexts of constraint. Thus constraints change the possibilities for behavior in general and in particular. When constraints conflict, in fact, they can add new capacities for behavior. Consider Walter's (1953) experiments with robot tortoises in this regard. The robots were programmed with simple fundamental capacities for behavior: to approach a source of light but to avoid any light that is too intense. Situated in sufficiently complex lighted environments, however, tortoises produced rich trajectories of movement over and above and different than merely approaching a source of light. For example, when a light was fixed to the front of a tortoise placed in front of a mirror, the tortoise began "twittering, and jigging like a clumsy Narcissus" (1953, p.128). The robot experiments of Braitenberg (1984) and Walter (1953) when taken together, illustrate how intrinsic and extrinsic constraints combine to define potentials for behavior.

So far, we have argued that cognition originates in capacities and constraints in the agent and the environment (see also Gibbs, 2006; Clark, 1997). However, it is the relation between agent and environment that defines the potential for cognitive activity. Agent and environment are causally intertwined in the potential for activity (Turvey, 2004), and unfolding contingencies enact and unfold action trajectories (Van Orden, Kello, & Holden, in press). Capacities and constraints make different behaviors more or less likely, but incidentals realize behavior. In the case of equally likely behaviors, for instance, it is attendant contingencies that decide which behavior will occur.

Gradients of uncertainty. A useful analogy can be made to a connectionist model. Connectionist models internalize constraints as changing weights among nodes, in a crude analogy to synapses and neurons. Behavior then originates in constraint satisfaction (e.g., Grossberg, 1980; McClelland & Rumelhart, 1981). Now imagine an unresolved cloud of active states in a connectionist model. These states could include all the possible ways to see a Necker cube, for example (3D cube, 2D image, etc). Of all these active states, no one outcome is yet realized. Each face of the Necker cube is merely potential until one or the other is selected by way of the immediate circumstances. Nevertheless, different weights between nodes make some faces more likely to dominate than others (Maia & Cleeremans, 2005). For instance, the flattened 2D Necker cube is a less likely outcome than 3D illusory cubes.

Constraints determine probabilities within the potential set, so that the potential set realizes a gradient of uncertainty. The gradient distinguishes potential states by their likelihood of being realized. Thus it distinguishes possible from impossible actions and the likelihood of possible acts (see also Fajen, 2005; Warren, 2006). On this basis, we equate the states of gradient potential sets with affordances and effectivities in cycles of action and perception (see also Davia, 2005; Gibson, 1977; Swenson & Turvey, 1991). Affordances are descriptions of the environment directly relevant for action, with reference to an organism and its effectivities; effectivities are descriptions of the organism directly relevant for action, with reference to an environment and what it affords (Turvey & Shaw, 1995). To this we would only add that gradient potential states are also descriptions of the history of an organism, directly relevant for action, with reference to its immediate future.

Gradients of uncertainty are like energy gradients in physical systems in several ways (however see Keijzer, 2003; Turvey & Shaw, 1995). Each gradient is a potential set for action that includes the likelihood of respective actions, without fully prescribing the particular action that will occur. In the Necker cube case, two percepts can even be equally likely (i.e., two orientations of a 3D cube). The behavior that is realized is therefore not ultimately determined by the constraints but by the immediate circumstances that enact behavior (Van Orden et al., in press).

Actions as subtle as eye movements or postural sway change the content of uncertainty gradients. And as the orientation of the perceiver changes, so does the likelihood that any particular act is realized. Facing the doorway will change the likelihood of walking through the doorway, for instance. Actions counter or realize gradients of uncertainty: walking through the door or not. Furthermore, as a cognitive act realizes the instantaneous gradient, it also brings into existence a new or evolved gradient. An action one way or the other changes the uncertainty that pertains to walking through the door, or something else.

Criticality and metastability. Cognition is never fully stable in the traditional sense of the term because gradients of uncertainty are perpetually evolving. Try staring at the Necker cube. Perception of the lines changes spontaneously, even though there is no change in the figure. Change occurs because new potentials for action (and perception) are introduced each time an action is taken, even actions as subtle as eye movements. Actions perpetually update potential sets, which insure a locally unstable system – a system close to a critical state. Critical states are states in which oppositional "forces" – constraints that favor one or another available outcome – are precisely balanced against each other. Critical states are thus a kind of boundary between qualitatively different behaviors.

A system that can stay close to critical states over time is metastable. Metastable states support multiple behavioral options, simultaneously, as the ghosts of attractors – attraction that remains present or potential but that is no longer an attractor (Kelso, 1995). Metastable states are never fully captured by any particular attractor, but instead move among attractors. Just as the connectionist representation of an unresolved Necker cube would be inclusive of different competing percepts, a metastable state is inclusive of different system outcomes. Metastability exists close to critical states, the state in which oppositional constraints are precisely balanced against each other. Near a critical state, metastability can extend sufficiently in time to perpetuate a potential set. This is possible if only relatively small perturbations occur continuously (cf. Jordan, 2003).

Metastability is a central concept for understanding why cognition is so exquisitely context sensitive (Kelso, 1995). Near a critical state, local interactions among component processes are strengthened if they satisfy available contextual constraints. This prunes the available options to those that best suit the context and that best situate the system for future contexts. Criticality is thus a kind of proto-anticipation (Shaw, 2001). In cognition, potential outcomes are a best guess about what will be required in the future (Jordan, 2003). This is particularly clear within language, for example, because context is constitutive of what words mean and how they are pronounced (Elman, 2005; Fowler & Saltzman, 1993; Turvey & Moreno, 2006).

Metastability is a desired property for an organism, a source of flexible and immediate action. Dynamics of the critical state recruit processes to the metastable interaction until local feedback loops extend to the periphery of the system, creating interdependence among all component processes. Interdependence poises processes to act together across the body, as a coherent organism for example. Interdependence also allows that even a small change in constraints, a source of additional minuscule constraint, in favor of one or another behavioral option will enact behavior immediately. In doing so, action both realizes and destroys the situated gradient of uncertainty that anticipated it (cf. Haken, 2000; Schneider & Kay, 1994).

Criticality and metastability are the source of counterintuitive predictions (Kello, Beltz, Holden, & Van Orden, 2007). Extensive feedback allows each component process to affect every other process. The consequent interdependence shows up as power-law behavior, a kind of negotiated compromise between processes' frequency of change and magnitude of change. Frequency and magnitude align themselves in a power-law, in that large amplitude change is less frequent, and low amplitude change is more frequent. Power-laws have a strong association to fractal behavior, the nested self-similarity of structure. Thus fractal and power-law behavior figure prominently in the evidence for self-organized criticality.

Critical states and metastable states also share properties with traveling waves. Traveling waves are themselves self-sustaining potential functions (e.g. Iliev, Khristov, & Kirchev, 1994; Infeld & Rowlands, 1990). Traveling waves are special in this regard because they are configurations of matter, energy, or uncertainty in which no fundamental difference exists between the current state and the potential to act.

Self-organization, by itself, would end in an attractor state, whether the attractor is a fixed point, a limit cycle, a taurus, or a strange attractor. By contrast, self-organized criticality does not end in an attractor. Instead it is attracted to metastable states that dance in the neighborhood of attractors without fully realizing them. Rather than being drawn in to an attractor, self-organized criticality is drawn toward a critical state, among the possible attractors rather than within a single attractor. Living systems appear to be drawn to such metastable states of criticality. Why remains unknown (Kauffman, 2000).

Evidence for Self-Organized Order in Cognition

While it is not entirely clear why living systems are attracted to an inherently unstable state, specific predictions can be made about the consequences of self-organized criticality. We mentioned already the ubiquitous power-law and fractal behavior that results from the general coordination among component processes. Other kinds of strongly nonlinear behavior include, for example, *sudden jumps* – qualitative changes in behavior in response to small changes in the balance of constraints. One can also observe *hysteresis*, wherein the balance of constraints that could yield a sudden jump is biased by recent history. Hysteresis is a form of bi-stable inertia that predisposes cognitive activity to repeat itself until a threshold is crossed, suddenly moving the system to the alternate state of inertia. And if such sudden jumps yield thermodynamically favored states, we would expect to see changes in the disorder or *entropy* of behavior. Entropy should increase prior to the sudden jump, followed by *negentropy* after the jump, a decrease in entropy accompanied by an increase in order. Finally, metastability and criticality imply interdependent components and time scales, yielding scale-free behavior, for example in the timing and the size of sudden jumps.

In the remaining sections of this chapter, we review evidence that is consistent with the predictions of complexity theory in cognition. While our organization follows different cognitive activities, the evidence we review is not specific to a particular cognitive activity. Instead, it corroborates fundamental assumptions about strong nonlinearity, emergence, and criticality in cognition overall. It lays the groundwork for reliable analyses, new intuitions, and different questions about cognitive activities.

Reasoning

We start with reasoning, cognitive activity that solves a problem, makes an inference, or creates knowledge. A typical problem with which to study reasoning is the gear turning task. Participants are presented with a daisy chain of meshed gears and the turning direction of the first gear. They then predict the turning direction of the last gear in the chain. A creative solution to the problem is to simply note whether there is an even or an odd number of gears in the chain. Knowing this, the direction of the last gear follows the simple rule: same direction as the first gear in an odd-numbered chain, and opposite direction as the first gear in an even-numbered chain. Most adults and children discover this more elegant solution, after first using a less efficient strategy, namely tracing the direction of each gear one after the other (Dixon & Bangert, 2002; Dixon & Dohn, 2003; Trudeau & Dixon, in press).

If the creative solution is self-organizing, a clear prediction can be made about the system's internal entropy before and after the novel solution emerges. One expects to see an increase in entropy before the emergence of the novel strategy. And one expects to see negentropy immediately after the emergence of the novel strategy. The predictions come from viewing the novel strategy as a potential state, prior to discovery. In each trial of the gear-turning task the potential strategy becomes more equal in probability to the gear-tracing strategy. This, in turn, increases uncertainty, which we see as an increase in entropy. At some point, gear-counting becomes sufficiently more probable than gear-tracing, and the sudden jump to the new strategy occurs. Past that point, the probability of the old strategy plummets and uncertainty is reduced in the bargain. The dominance of the new strategy is revealed in negentropy..

The prediction has been corroborated in the gear-turning task (Stephen, Dixon, & Isenhower, 2007a; Stephen, Boncoddo, Dixon, & Magnuson, 2007b). Detailed motion data were captured for participants' finger movements as they traced gears. Entropy was calculated consistent with the basic formula of Shannon and Weaver (1948) within a recurrence quantification analysis (Webber & Zbilut, 2005). As predicted, entropy in finger movements increased over trials and peaked right before the creative solution emerged. At that point entropy changed to negentropy – a reduction in entropy with a corresponding increase in information (order). Eye-movements showed the same increase in entropy and change to negentropy, parallel with the finger movements (Stephen et al., 2007b).

Repeatedly measuring finger or eye-movements during problem solving made it possible to reconstruct the intrinsic dynamics of problem solving, the rise and fall of entropy predicted the emergence of the new solution in the gear-turning task. But what is the logic by which an index of finger or eye-movements is transparent to cognitive changes? It originates in the idea of interdependent processes mentioned earlier. A truly interdependent system allows variation in each component to reflect variation in every other component. Massive interdependence ensures that a well-chosen repeated measurement will contain information about the entire system. In the ideal, one may reconstruct the dynamics of a possibly higher-dimensional system from a one-dimensional time-series of repeated measures (Mañé, 1981; Takens, 1981).

No prior studies have successfully captured the coming into existence of a new problem solving solution, or any other "executive" activity (however see Kelso, 1995, for emergence of percepts and motor coordination). Should these studies prove reliable, and should they generalize to comparable phenomena, their authors will have established, for the first time, a basis for origins of order in high-level creative cognition. As entropy accumulates, the gear-tracing problem-solving strategy comes apart, so to speak, and makes way for the emergence of the new strategy of counting the gears.

Listening

The key historical issue in speech perception has been categorical perception: how to parse the stream of sounds into meaningful units, whether on the level of phonemes, the level of words, or the level of sentences. No apparent category boundaries exist in speech (Klatt, 1989), and environmental noises, changing speaking rates, as well as co-articulation make it difficult to find invariant acoustic properties that could be used to define such boundaries. Yet, even infants ably distinguish between phonemes (e.g., Dehaene-Lambertz, Dehaene, & Hertz-Pannier, 2002).

Category boundaries emerge in non-linear interactions. This is why no one-to-one correspondence exists between acoustical signals and percepts; nonlinear interactions produce many-to-one and one-to-many relations in addition to one-to-one. Consequently, for some regions of the acoustic signal, on the one hand, we can expect to see no change in the percept despite relatively large changes in the acoustic signal. For other regions of the acoustic signal, on the other hand, we can expect sudden jumps in percept from one category to another, despite only minimal changes in the acoustic signal. The occurrence of sudden jumps can also be expected to vary with the context, constrained by the immediate history of the participant.

Each of these predictions has been confirmed using artificial speech (see Tuller, 2005, for a review). An artificial acoustic continuum was created for the word "say," for example, by inserting an increasingly longer gap of silence after the phoneme /s/. Participants' task was to report whether the acoustic signal sounded like "say" or "stay." Across participants, a short gap of silence (0-20ms) yielded perception of "say," and a greatly expanded gap of silence (60-80ms) yielded perception of "stay." These stable regions of the acoustic signal yielded the same percept ("say" or "stay," respectively) despite changes in the gap duration of the acoustic signal.

Intermediate regions (30-50 ms) yielded sudden jumps from one percept to the other. Minimal changes in the acoustic signal induced categorically different percepts. For example, as the gap of silence increased in small increments, a participant abruptly switched from perceiving "say" to perceiving "stay." The sudden jump is highly dependent on context however, and it does not correspond to a static threshold. For example, when the acoustic signal changes in direction from shorter to longer gaps of silence, the sudden jump will occur at a longer gap than when the acoustic signal changes in the reverse direction, from longer to shorter gaps of silence. In sum, whether the participant heard “say” or “stay” on the previous trial affects what is heard on the subsequent trial. The immediate history of the perceiver matters.

This pattern of context dependency in the say/stay example demonstrates hysteresis. This dynamic originates in ranges of acoustic parameters or input dimensions that are ambiguous (not unlike the Necker cube), and are therefore sensitive to the embedding context. In other words, hysteresis emerges out of a kind of dynamical instability that amplifies available constraints, manifesting as context sensitivity. Thus, although hysteresis behavior involves threshold behavior, hysteresis thresholds are not fixed, but rather are emergent properties: The locations of thresholds, or more precisely the locations of critical points, are entirely context dependent.

By changing the available constraints, one can even invert the gradient of uncertainty to turn hysteresis into its opposite, a contrastive effect. In this case, the sudden jump from “say” to “stay” (or vice versa) comes extra early, rather than being delayed. The relative likelihood of the contrastive effect (versus hysteresis) is increased, for instance, by over exposure to the particular stimulus continuum. In the say/stay example, incrementally changing stimuli in sequence will now pull the category threshold toward the “say” end of the continuum so that the “say” perception looses stability earlier and transitions earlier to “stay” (or vice versa if the order of presentation is reversed).

Sudden jumps, hysteresis, and contrastive effects are not visible when trials are randomized, or when participants’ performance is averaged over trials to eliminate sequence effects (compare Fechner’s method of limits). Once understood, however, these reliable empirical flags of self-organization serve to advance the research program of speech perception. For example, one may write differential equations for emergent order to characterize and explore the intrinsic self-organization of speech perception. By tracking stability within these differential equations, one can flesh out the detailed interplay of context, history, and agent in perception.

Such interplay can even account for the fact that speakers returning from years abroad will have an accent in their native language. Learning non-native phonemes changes the topology of phoneme perception and production in a native language (Sancier & Fowler, 1997). Clearly, agents are themselves changeable reservoirs of constraint on speech perception, and embodied constraints work in concert with contextual constraints. The critical ratios of constraints that define boundaries in perception "adjust flexibly with factors such as phonetic context, the acoustic information available, speaking rate, speaker, and linguistic experience" (Tuller, 2005, p. 355).

The speech perception examples we have reviewed catalog clear progress toward understanding speech perception. Furthermore, they are not rarified laboratory phenomena. They are laboratory analogs of ordinary, variable, and creative perception. Ordinary perception must recognize the same word produced by males, females, speakers of different ages, and with different dialects and accents, "and by the same speaker in markedly different linguistic and intentional contexts, even when the listener has had no prior experience with the other individual's speech patterns. Thus, perceptual stability coexists with perceptual flexibility." (Tuller, 2005, p. 355).

Speaking

Language is more than just parsing a continuous stream of acoustic signals into meaningful units. It requires able communication: the ability to say words and sentences and sustain a conversation with a partner. In what follows, we review evidence for self-organization and criticality in three domains of speech: metastability of producing words, coordination of articulators to pronounce words, and coordination of partners in a conversation.

Power-law scaling in repeated speech. As discussed above, the coordination dynamics near critical states (i.e., metastability) express a fractal pattern. Such patterns have a redundancy in their composition (or self-similarity) that makes them appear similar, no matter how one decomposes the composite whole. For example, power-law scaling refers to a fractal pattern in which large amplitude, low frequency oscillations embed intermediate amplitude, intermediate frequency oscillations that, in turn, embed low amplitude, high frequency oscillations (across an indeterminate number of timescales). Because fractals express self-similarity as a statistical kind of self-affine structure, we should find the same kind of statistical self-affine structure in any way that measurements divide up behavior.

Take for example a spoken word. If this behavior is fractal, then no matter where or how we take repeated measures of a spoken word, they should exhibit the same kind of fractal pattern. Kello, Anderson, Holden, and Van Orden (in press) tested this hypothesis by asking adult participants to repeat the word "bucket" over 1000 times, tracking a metronome beat for when to speak. Each spoken "bucket" was recorded and divided into its syllables "buck" and "ket." Spectographs of "buck" and "ket" were parsed further into 300 Hz frequency bands evenly spaced from 150 Hz to 13,350 Hz. This parsing resulted in 45 component frequency bands per syllable, and the intensity of the acoustic signal was tracked at each of the 45 bands. In total, the speech signal was parsed to yield 45 measurements per syllable and 90 measurements per participant, which were tracked across 1100 repetitions of "bucket." The point of all this was to create many measurements of a cognitive activity and track them all across very many repetitions.

Each participant's 90 measurement series were examined one by one for fractal structure. Fractal scale-free behavior was estimated using spectral analysis, then using relative dispersion analysis, and finally detrended fluctuation analysis (Holden, 2005). These methods estimate the extent to which data points in a series are correlated over the long-term. Each method can yield a fractal dimension, in this case dimensions close to the ideal value 1.2, the fractal dimension of metastable dynamics. The important finding was that each of the 90 measurement series of every participant yielded estimates close to the ideal scaling value, just as metastability predicts. This finding is unique because it established fractal patterns in so many simultaneous measurements. It provides strong evidence for metastability in speech production.

Coordination of articulators. In proper speech, movements of tongue and lips are coordinated in intricate patterns that take years to perfect. If such coordination results from dynamical self-organization, it should obey the laws of coordination dynamics, laws that were first established for the coordination of limbs (Kelso, 1995; Turvey, 1990). For example, a 1:1 limb coupling (e.g. oscillating one index finger at the same frequency as the other index finger) is more stable than a 2:1 coupling (e.g., oscillating one index finger twice as fast as the other index finger). The more complex 2:1 coordination becomes weaker as frequency of oscillation increases, such that only the simpler mode (1:1) remains stable. If the same principles of coordination hold for speech, then speech errors should also result from failures of coordination.

This prediction was tested in speech production by analyzing the kinematic data of tongue and lip movements during articulation (Goldstein, Pouplier, Chen, Saltzman, & Byrd, 2007). Adults repeated two-word phrases such as "cop top" at varying speeds elicited by a metronome (80 to 120 beats per minutes). Note that the /p/ sound requires lip constriction, while the /k/ and /t/ sounds require tongue constrictions (raised tongue dorsum for /k/, and raised tongue tip for /t/). The lip-tongue coordination in "cop top" therefore follows a 2:1 pattern: Two cycles of lip oscillation (/p/) are completed for every one cycle of tongue dorsum constriction (/k/) or for every one cycle of tongue tip constriction (/t/).

By increasing the frequency of the metronome, Goldstein et al. (2007) induced speech intrusion errors: blends of the intended and intruding articulator movements for the /k/ and /t/ sounds. Tongue dorsum constrictions intruded on tongue tip constrictions to produce a novel combination of articulator movements. Interestingly, the new sound does not occur in English – hence it cannot be explained by some fatigued executive planning faculty. But more importantly, the new sound brought into existence a more stable coordination pattern of lip and tongue movement. Every "word" now had a tongue tip and a tongue dorsum motion at the beginning, and a lip motion at the end. As a result, one cycle of lip oscillation was completed for one cycle of tongue dorsum constriction – an anti-phase 1:1 coordination.

Similar results were reported for induced speech errors that did not require any overt repetition of words but instead relied on the visual presentation of primes (Motley & Baars, 1976; see also Pouplier & Goldstein, 2005). Together, they strongly reject the idea that speech errors are random, for example a result of a simple mix-up of articulator motions in the linear sequencing of speech. Instead they show that speech errors, and therefore the system of articulators, follow laws of dynamical stability. At comfortable speeds, coordination of articulators can be more complex, but with increasing speeds, coordination is captured by the most intrinsically stable 1:1 mode.

Coordination in conversations. More than one person can participate in coordinated cognitive activity. Talking with each other requires substantial coordination and cooperation to sustain the conversation as an ordered structure in time. Indeed speakers in a conversation have a tendency to converge in dialect (Giles, Coupland, & Coupland, 1991), speaking rates (Street, 1984), vocal intensity (Natale, 1975), and pausing frequencies (Cappella, 1981). Thus the conversation, itself, can be characterized as a self-organizing system (Shockley, Santana, & Fowler, 2003).

But self-organized criticality goes beyond the coordination of speech activities. It makes the unintuitive prediction that measures of non-speech activity are coordinated as a conversation takes place. Criticality strictly implies interdependence among component activities across the entire system. As we saw for reasoning with the gear task, a well-chosen repeated measurement should contain information about the entire system. If so, then repeated measures of the non-speech activity of conversing agents may be coordinated in the course of the conversation.

To test this prediction Shockley et al. (2003) measured bodily sway along the anterior-posterior dimension, at the hip, while participants stood and conversed. Participants' task was to figure out, how two subtly different versions of a cartoon differ. Each participant could only see one of the two versions, so a conversation was necessary to cooperatively find differences between the cartoons. Conversing participants either faced each other or they had their backs turned to each other (facing confederates). In a control condition, the two participants faced each other but conversed with a confederate whom they could not see.

Shockley et al. (2003) found more entrainment of hip movements when participants conversed with each other than when they faced each other but conversed with the confederate. Most importantly, participants hip movements were entrained even when the participants could not see each other. The act of conversation between two participants was enough for their respective hip movements to become coordinated. These findings are particularly interesting because hip movements are not directly related to speech, yet hip coordination emerged despite lack of visual feedback. It suggests that a spontaneous coordination between two conversing people can permeate non-language behaviors.

Taken together, findings from speech acts corroborate that interdependent components coordinate cognitive activities. We thus expect unintuitive coordinations of measurements taken across the human body and compared between participants, so long as there is a coordinating event. Information processing and task-specific components do not appear to be the originators of order, whether it is in repeated speech, in speech errors, or in conversations. Rather, the origin of order in speech is coordination, the self-organization of nonspecific component activity in linguistic behavior.

Reading

Reading has long been a kind of poster-child for cognition because it includes aspects of perception, language, memory, and so on, and because it is a culturally derived capacity (Huey, 1908). Thus, what we learn about reading is likely to generalize to cognition at large. Conveniently, there is also a vast literature on this topic, research that has largely come to focus on questions about the ambiguity of written symbols.

Ambiguities in linguistic discourse exist at all scales of description, from visual features, to letters, phonemes, and syllables, and to the meanings and pronunciations of words, sentences, paragraphs, and narratives (Langacker, 1987). Moreover, multiple scales of ambiguity can accrue, as ambiguity exists among multiple relations, including graphemes and phonemes, spelling bodies and pronunciation rimes, whole-word spellings and pronunciations, morphological structures and meaning, prosody and pragmatics, surface forms and deep forms of phrases, sentences, and narratives, and so on (Van Orden et al., 1990). The way in which the system resolves ambiguity is thus an important test of the fundamental workings of the system.

Note that accumulation of ambiguity is a slowly driven process. One requires experience with multiple words and sentences before ambiguity can come into existence. For example, one needs exposure to several kinds of uses of the word “lead” (e.g. lead::metal vs. lead::guide vs. lead::principal vs. lead::ahead, and so on) before “lead” becomes ambiguous in meaning and pronunciation. Thus ambiguity accumulates more slowly than on-line linguistic experience. Once in place, ambiguity mimics metastability in that both make multiple options for action available. To resolve ambiguity, or to resolve metastability, one needs a sufficiently disambiguating context.

Because ambiguity implies more than one potential outcome, the details of ambiguity define the gradient of uncertainty. Ambiguity works like static friction in a physical system in the sense that it must be resolved or overcome before reading comprehension or performance can occur. Ambiguity resolution is thus a reduction in uncertainty, and an increase in information. Comparatively more ambiguity implies greater reductions of uncertainty, an increased likelihood of extended time in the metastable potential set, which is revealed in power law behavior. Thus the degree of ambiguity predicts the extent of power law behavior.

Power-law behavior in reading sentences. Sentence reading can be examined by presenting one word at a time, and requiring a key press after each word before the next word is shown. Participants are instructed to read as naturally a possible, but they must press a key after reading a word in order to see the next word. This method collects self-paced reading times and can be focused on the disambiguating region of an ambiguous sentence – the specific words where the reader confronts the ambiguity.

If reading comprehension is a product of self-organized criticality, we expect to see power-law behavior in the frequency distribution of reading times. A power-law distinguishes itself from other frequency distributions (i.e., log-normal) in the exaggerated or "stretched" slow tails of the distribution. Recall that in a power-law, the frequency of an event's occurrence and its magnitude (or amplitude) are aligned in a linear fashion on a log-log scale. In the case of reading times, the frequency of a response time (how often a particular response-time occurs) aligns with magnitude of the response time (how very slow a response time is). Faster response times are more likely than slower reaction times, with an overall linear relation between likelihood and magnitude on log-log scales. This linear relation is evaluated using the slope of the line that relates likelihood and magnitude. More shallow slopes reflect more extreme values of magnitude – more extremely slow reading times for instance.

The expected relation between ambiguity and power law behavior was corroborated in an experiment that used increasingly ambiguous sentences (Schultz & Tabor, 2008). For example, the sentence: "As the author wrote the story she envisioned *grew rapidly in her* mind," is more ambiguous than the sentence: "As the author wrote the story that she envisioned *grew rapidly in her* mind," which in turn is more ambiguous than the sentence: "As the author wrote the story *grew rapidly in her* mind." The italicized words are the words in the disambiguating region. Participants' reading times for the words in the disambiguating region were measured and plotted as frequency distributions.

As predicted, the distribution tails at slow reading times became shallower, extending to more extreme slow response times as ambiguity increased. In other words, the slope estimate of the linear relation revealed a direct relation between ambiguity and stretched tails. One implication of this outcome is that the same processes are entailed in all instances of sentence reading. Undifferentiated power-law behavior presents no joints at which to carve out isolated workings of particular reading components. Power laws may imply that all sources of constraint on the reading performance are present in all instances of the reading performance; there are no qualitative differences between slow reading responses and fast reading responses in the power law. Different reading tasks may self-organize different dynamics, but tasks and stimulus words do not selectively activate causal properties of reading components.

Power-law behavior in reading words. As discussed earlier, linguistic ambiguity is not limited to sentences, but extends in a nested fashion to graphemes, spelling bodies, and whole words. Whole word ambiguity always nests within it spelling body ambiguity, and spelling body ambiguity always nests grapheme ambiguity, but not vice versa. If a state of ambiguity resembles a critical state, then we expect power law-behavior in reading tasks that require disambiguation. Also, if nested structures of linguistic ambiguity comprise fractals, then power-law behavior should be more present to the extent that particular words exhibit the fractal nesting.

These predictions were tested using lists of words that differed in the scale at which their spelling was ambiguous: whole-word ambiguity (“lead”) versus spelling-body ambiguity (“bead”) versus grapheme ambiguity (“beat”) (Holden, 2002). After a key press, a word appeared on a computer screen, and the participant named it as quickly as possible. As predicted, frequency distributions of naming times differed in their slow tails, in that slow tails were most exaggerated for words with whole-word ambiguity, followed by words limited to spelling body ambiguity, followed by words limited to grapheme ambiguity. Although the modes of the naming time distributions did not change, the slow tails of the distributions were differentially stretched, attendant on differences in word ambiguity.

Scale of ambiguity also exaggerated slow tails held when participants decided whether strings of letters were English words or not (Holden, 2002). The larger the limiting scales of ambiguity, the more stretched were the slow tail of the distribution. Together, these results support the idea that linguistic ambiguity, itself, has fractal structure. Words that represent this structure to a larger degree, such as whole-word ambiguous words, also produce the clearest power law behavior. Interestingly, when response times are combined across the ambiguity conditions, they combine in a common power-law. This is suggestive, though not conclusive, that all reading performances in this case sample a common power law. If so, then the same sources of constraint are present in all reading performances, though they may be differently emphasized in different tasks (see Holden, 2002). Just as in sentence reading discussed above, there is no empirical basis for individuating separate reading processes that act in isolation.

Power-law vs. lognormal behavior. Holden (2002) and Schultz and Tabor (2008) manipulated word and sentence properties, the influence of laboratory context upon the gradient of uncertainty. In contrast, Holden, Van Orden, & Turvey (2008) conducted analyses to emphasize the influence of participant history. They examined differences among participants in response time distributions. The task was again a word-naming task, this time using a list of 1100 words, identical across participants. For a few readers, there was clear evidence of very shallow power-law behavior in the exaggerated slow tails of their distributions. They produced widely dispersed naming times. But for other readers, the distribution of word naming times closely fit a lognormal distribution. Naming times for these readers were more narrowly dispersed. Most readers fell somewhere in between, with naming time distributions that were well fit by a mix of power law and lognormal behavior.

The difference among readers reflects their propensity for reducing the degrees of freedom in word naming. Holden et al. (2008) speculate that the three kinds of distributions – lognormal dominant, mixed, and power law dominant – map onto the complexity of the reading task for particular readers. The basic idea is that reading is more complex (e.g., less constrained) for some readers than others, which means that more uncertainty exists for some readers than others. For instance, as readers gain experience, they accrue constraints word-by-word that "grease the wheels" of cognitive dynamics, so to speak. They reduce or eliminate thresholds of ambiguity as word-specific constraints of more and more words dominate performance (yielding lognormally distributed naming times). Thus for some highly experienced readers, word naming times may simply express a kind of word naming "inertia," in line with the task-induced intention to read words aloud.

Skilled word naming has long been thought of as an automatic behavior. However, no trustworthy context-free criteria exist for automatic behavior (Fearing, 1970; Tzelgov, 1997). Even a canonical automatic effect such as the Stroop effect (Stroop, 1935) can be reduced or eliminated by only small changes in laboratory method (Bauer & Besner, 1997; Besner & Stolz, 1999a, b; Besner, Stolz, & Boutilier, 1997). These concerns are inconsequential if skilled performance is seen as the interplay of sufficient internalized and external constraints. Sufficiently constrained performances sample response times from a lognormal distribution, otherwise they sample response time from a power-law.

In this light it is noteworthy that the equivalent of "stimulus inertia" in a strictly feed-forward multiplicative system will produce finishing times that are lognormally distributed (as the number of multiplicative steps approaches infinity). This is the multiplicative version of the central limit theorem. Also, one can create a strictly feed-forward system from a feedback system by adding sufficient constraints to make feedback redundant. In other words, a sufficiently constrained complex system will produce finishing times dispersed in the lognormal pattern, as a product of serial multiplicative interactions among random variables (Farmer, 1990).

Remembering

Complex systems exhibit scale invariance and abundant data suggest that scale invariance occurs in remembering (Brown, Neath, & Chater, 2007). For instance, if people are asked what they did or will do in a week, month, or year's time, they will perform almost identically across the different scales of time. They do not remember better or in greater quantity the events from a recent week than events from a recent year. In these examples, memory appears the same at all time frames for things to be remembered. Memory appears to have no preferred scale; it works the same at all timescales (Chater & Brown, 1999; Maylor, Chater & Brown, 2001).

Evidence for scale-invariance also exists in error data. Patterns of transposition errors in order reconstruction tasks remain constant across many timescales of delay between events, from milliseconds to weeks (Huttenlocher, Hedges, & Prohaska, 1992; Nairne, 1991, 1992; Neath, 1998). Also, the proportion of errors in serial recall tasks for each serial position is invariant to absolute parameters such as inter-stimulus presentation interval, time between trials, familiarity and meaningfulness of the material, and degree of learning (Braun & Heymann, 1958; McCrary & Hunter, 1953).

Evidence for scale invariance also shows up in the dynamics of serial position curves. A typical finding is that increasing the retention interval on recall tasks abolishes recency effects (Glanzer & Cunitz, 1966; Postman & Phillips, 1965). However, when the spacing between presented items is increased proportionally with the retention interval, the recency effect stays (Bjork & Whitten, 1974). That is, recency effects scale with the ratio of inter-stimulus interval and retention interval and the structure of serial position effects holds at many scales of measurement. Further evidence comes from the finding that when items are grouped within a list, items at the beginning and end of each group show improved recall relative to items in the middle of the list in addition to primacy and recency effects present for the entire list (Frankish, 1985, 1989; Ryan, 1969a; b; Hitch, Burgess, Towse, & Culpin, 1996). This sort of nested self-similarity is a hallmark of scale-invariant systems.

Finally, power-law forgetting curves may be indicative of scale-free memory. It is well established that the magnitude of the rate of change for forgetting diminishes with time – the type of decay expected from power-law forgetting. Power-law behavior is characteristic of self-organizing systems, which demonstrate scale invariance in their behavior. Complex systems orient themselves towards states of self-organized criticality, where the potential for action of the system is evenly distributed over all its scales of measurement (Turvey & Moreno, 2006). The result of this organization is power-law behavior when the system is perturbed. Note however that some argue that forgetting curves are not actually best described as power-laws (Rubin, Hinton, & Wenzel, 1999; Wickens, 1999).

A scale-invariant model of memory would naturally suit the application of complex systems theory to cognition. Brown et al. (2007) describe such a model (SIMPLE for scale invariant memory, perception, and learning) that accounts for many esoteric and fine-grained details of memory performance. While SIMPLE is not a complex system per se, it is a useful model for building intuitions about how complex systems theory would begin to reframe memory phenomena. For instance, it demonstrates how similar principles can account for various memory phenomena over many different time scales and tasks. It emphasizes interference (interactions) among temporal neighbors to account for forgetting. And its memory traces are organized at logarithmic distances from a point of reference.

Recent work by Rhodes and Turvey (2007) was inspired by the dynamical implications of the SIMPLE model. They assume that the minimum units of analysis are organism-environment systems. In that regard, they applied insights from animal foraging behavior to cognition and constructed a dynamical analogy to repeated memory-recall. The basic idea was to predict the pattern of times between successive instances of recall. For instance, repeated category recall occurs when a participant reports all the animals they can remember or all the world's capital cities. This kind of category recall typically includes briefly interspaced reports of semantically related clusters (e.g., eagle, hawk, pigeon), separated by lengthy time intervals before another report (Bousfield & Sedgewick, 1944).

Animal foraging behavior is also characterized by short path clusters of search activity separated by lengthy search paths, when food is sparsely dispersed at locations that are unknown prior to foraging. In these conditions an animal travels an overall search path that strongly resembles a Lévy flight. This is a random walk or diffusion process in which the step sizes are distributed as a power law. Short steps on the search path are very common and very long steps are rare. Consequently, the frequency of step sizes is inversely proportional to their length. This search pattern is found in foraging for a variety of species, including humans, and may be optimal for an environment with fractal properties (Viswanathan, et al., 2001).

Like foraging behavior, the overall scale-free relation in the recall data is well modeled by Lévy distributions, shallow sloped inverse power-laws with infinite variance. These results allow that memory resembles a sparsely populated landscape with fractal structure. Perhaps memory is an adaptation that is homologous to foraging, for instance, an adaptation to a search domain with fractal properties (Rhodes & Turvey, 2007).

Acting with Purpose

The previous sections all discuss data collected in laboratories, data that concern laboratory performances done at the behest of a scientist. But most cognitive activity is self-directed, reflecting the intentions of the actor. What then does laboratory performance say about the origin of order in intentional activity? Arguably this is the first question of cognition, the question that qualifies how other details of cognitive performance are interpreted.

A scientist must instruct participants to behave in line with the purpose of the experiment. In this way, participants' intentions become tied to task instructions and the purpose of the laboratory preparation (Vollmer, 2001). In a word naming task, for example, participants must insure that they respond with articulated speech; they must be vigilant and attentive to read aloud each word as soon as it appears; and they must keep in mind inherent limits on their behavior – to say aloud only the words presented and only at the time they are presented. In other words, in taking on the purpose of the experiment, a participant situates, or scales, variation in mind and body to stay within constraints made explicit in task instructions.

If intentionality and purposeful behavior are products of self-organized criticality, as Juarrero (1999) has argued, task instructions equal sources of constraint that self-organize a gradient of uncertainty for performance. Purposeful behavior is anticipated in the gradient potential as a metastable state. In turn, we expect the nested fractal pattern of 1/f scaling exhibited by metastable states (e.g., a particular power-law relationship). In the nested fractal pattern, low-amplitude high-frequency variation is nested within intermediate-amplitude intermediate-frequency variation that, in turn, is nested within large-amplitude low-frequency variation, and so on to an indefinite number of larger amplitudes and lower frequencies. The resulting fractal pattern is scale-free. Amplitude (magnitude) and frequency of variations, across an indefinite number of timescales, form a line with slope close to -1 on log/log scales, the slope of the scale-free 1/f scaling relation.

But what laboratory method justifies equating variation in behavioral measurements with variation in the intention to act? First, variation in intentions fluctuates on slower time scales than the pace of measurement trials (as does 1/f scaling). Although a participant bears in mind the intention to perform on each and every trial, variation in intentions is seen across trials, on slower timescales. Second, examining variation in the intention to act requires that each measurement trial repeat identical conditions for the intention to act – task demands, stimulus, response and all else that makes up a laboratory trial. An ideal task is therefore a production task in which participants repeatedly produce lines of one inch in length, for example, or say the same word, or repeatedly estimate the passing of a second. These production tasks entail the same purpose and action on every trial so changes in performance across trials better emphasize fluctuations in the participant's intentions.

In the time-estimation production task, participants press a key to signal that a second has passed, for a total of 1100 estimates. The relevant data form a pattern of variation in each participant's performance across all of the trials. Spectral analyses of this pattern (and other complementary analyses) reveal the predicted 1/f scaling – the fractal pattern of nested variation. That is to say, within and across participants, the spectral analysis yields slopes close to -1, the slope of 1/f scaling. As we have explained, the fractal scaling cannot be attributed to changes in experimental task, so it instead reveals the intrinsic fluctuations of acting with purpose. In that regard, production tasks produce the clearest examples of fractal 1/f scaling in cognitive performance (Gilden, 2001; Kello et al., in press; Thornton & Gilden, 2005).

The finding of 1/f scaling in repeated measures – fractal variation across an indefinite number of timescales – indicates that acting with purpose comprises an indefinite number of psychological dimensions. A reasonable claim is to equate these dimensions with sources of constraint that oscillate across an indefinite range of frequencies (Van Orden et al., 2003). Constraints recur across trials and change at their respective natural frequencies to support acting with purpose. Because they combine nonlinearly, resulting in scale-free behavior, they therefore yield the emergent property, 1/f scaling. Also, as we have noted for other power law behavior, there are no empirical joints along which to divide intentional behavior. No intelligible means exist to isolate one source of constraint from another (Thorton & Gilden, 2005). The same sources of constraint are present in each instance of intentional behavior.

Variation in intentional behavior exhibits the emergent property of 1/f scaling, and so intentional behavior itself is emergent. Acting with purpose originates in metastable cognition, an anticipatory "posture of readiness, like that of a runner poised for a quick start" (Woodward & Schlosberg, 1965; p. 830), or muscles that "become tense in preparation for a task" (Bills, 1943; p. 11). Sources of constraint for purposeful action exist throughout the body, including anticipatory movements and excitations, as well as dynamical processes of the body such as head position, posture, respiration, and digestion. At the extreme, sources of constraints even include the details of capillary red blood flow, and local oxidation in muscle tissue. All these constraints combine to reduce the degrees of freedom for cognitive behavior, and thus contribute to the fractal variation in cognitive activity.

Summary and New Questions

What are the origins of order in cognitive activity? Or as Neisser (1968) asked, "[1] How is the raw material [of constructions] organized? [2] What determines what is constructed? And [3] what purpose does it serve?" (p. 280). In this chapter, we have addressed these questions, applying concepts of complex systems to cognitive activities. Accordingly, we argued (1) that raw material of cognition self-organizes to stay near critical points, yielding metastable states that anticipate contextually appropriate actions. Furthermore, we agued (2) that potential sets, the system's best guesses about action in the future, are determined by situational and historical constraints, updated each time an action takes place. Finally, we argued (3) that organization serves the purpose of consuming of gradients of uncertainty (creation of information).

We then reviewed evidence for the fundamental assumptions and laws of complex systems within various cognitive activities. We chose the studies for this review to illustrate the breadth of corroboration for fundamental assumptions. This evidence includes, for example, nonlinear disproportionate relations between changes in the environment and changes in behavior, sudden jumps, hysteresis and contrastive effects, dynamics that obey laws of coordination, and ubiquitous power law behavior and 1/f scaling, the signature behaviors of complex systems. Our sample of this evidence comes from reasoning, speaking, reading, and remembering, and, most generally, acting with purpose. As with any basic assumptions and their confirmation, the results surveyed here should not be seen as ends in themselves. Valid assumptions are like a license to drive, existence proofs for the legitimacy of the complex systems perspective.

Having corroborated basic assumptions, we now face new questions about cognitive activity. These questions pertain to the control of cognitive activities and their dynamics. Answers to control questions will make explicit otherwise implicit conflicting constraints, and discover the tradeoffs that change the qualitative outcomes for perception and action (Kelso, 2003). Control, from this perspective, is not the traditional notion of control, such as control by stimuli or homunculi, rather control emerges in the interplay of constraints combining agent and environment. This idea of control pertains to human error, for example (Vincente, 1999), and to human creativity.

The question of dynamics concerns changes in the interplay among constraints, over and above simply demonstrating which constraints are in play. It addresses more subtle indications of change, quantitative changes in fractal dimension, for example, or the slopes of power laws. To understand details of dynamics will require fleshing out the types and relations among forces at play in cognitive activity, the dynamics at the heart of task difficulty and individual differences, for example, and the conflicts among constraints that bring new capacities for behavior on line.

References

Aks, D. J., & Sprott, J. C. (2003). The role of depth and 1/f dynamics in perceiving reversible figures. *Nonlinear Dynamics, Psychology, and Life Sciences, 7,* 161-180.

Anderson, J. R., Lebiere, C., Lovett, M., & Reder, L. (1998). ACT-R: A higher-level account of processing capacity. *Behavioral and Brain Sciences, 21,* 831-832.

Allport, A. (1989). Visual attention. In M. I. Posner (Ed.). *Foundations of cognitive science* (pp. 631-682). Cambridge, MA: MIT Press.

Aslanidi, O. V., & Mornev, O. A. (1996). Soliton-like regime in the Hodgkin-Huxley equations. *International Conference on Mathematical Models of Nonlinear Excitation Processes*, Tver, Russia, July 2-5.

Azuma, T., Williams, E. J., & Davie, J. E. (2004). Paws + cause = pause? Memory load and memory blends in homophone recognition. *Psychonomic Bulletin & Review, 11*, 723-728.

Bartlett, F. C. (1932). *Remembering*. New York: Cambridge University Press.

Bauer, B., & Besner, D. (1997). Mental set as a determinant of processing in the Stroop task. *Canadian Journal of Experimental Psychology, 51*, 61-68.

Beaumont, J., Davidenko, N., Davidenko, J. M., & Jalife, J. (1998). Spiral waves in two-dimensional models of ventricular muscle: Formation of a stationary core. *Biophysical Journal, 75*, 1-14.

Berkinblit, M. B., Feldman, A. G., & Fukson, O. I. (1986). Adaptability of innate motor patterns and motor control mechanisms. *Behavioral and Brain Sciences. 9*, 585-638.

Bernstein, N. (1967). *The coordination and regulation of movements*. London, England: Pergamon Press.

Besner, D., & Stolz, J. A. (1999a). Unconsciously controlled processing: The Stroop effect reconsidered. *Psychonomic Bulletin & Review, 6,* 449-455.

Besner, D., & Stolz, J. A. (1999b). What kind of attention modulates the Stroop effect? *Psychonomic Bulletin & Review, 6*, 99-104.

Besner, D., Stolz, J. A., & Boutilier, C. (1997). The Stroop effect and the myth of automaticity. *Psychonomic Bulletin & Review, 4*, 221-225.

Bills, A. G. (1943). *The psychology of efficiency*. New York: Harper & Brothers.

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology, 6,* 173-189.

Bousfield, W. A., & Sedgewick, C. H. W. (1944). An analysis of restricted associative responses. *Journal of General Psychology, 30*, 149-165.

Braitenberg, V. (1984). *Vehicles*. Cambridge, MA: MIT Press.

Braun, H. W., & Heymann, S.P. (1958). Meaningfulness of material, distribution of practice, and serial-position curves. *Journal of Experimental Psychology, 56,* 146-150.

Brown, G. D. A., Neath, I., & Chater, N. (2007). *A ratio model of scale-invariant memory and identification*. Submitted manuscript.

Cappella, J. N. (1981). Mutual influence in expressive behavior: Adult-adult and infant-adult dyadic interaction. *Psychological Bulletin, 89,* 101-132.

Chater, N., & Brown, G. D. A. (1999). Scale-invariance as a unifying psychological priniciple. *Cognition, 69*, 17-24.

Clark, A. (1997). *Being there: putting brain, body, and world together again*. Cambridge, MA: MIT Press.

Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science, 298,* 2013-2015.

Davia, C. J. (2005). Life, catalysis and excitable media: A dynamic systems approach to metabolism and cognition. In J. Tuszynski (Ed.), *The Physical basis for consciousness*, 229-260. Heidelberg, Germany: Springer-Verlag.

Davydov, A. S. (1979). Solitons, bio-energetica, and the mechanism of muscle contraction. *International Journal of Quantum Chemistry, 16*, 5-17.

Depew, D. J., & Weber, B. H. (1995). *Darwinism evolving: systems dynamics and the genealogy of natural selection*. Cambridge, MA: MIT Press.

Dixon, J. A., & Bangert, A. S. (2002). The prehistory of discovery: Precursors of representational change in solving gear-systems problems. *Developmental Psychology, 38,* 918-933.

Dixon, J. A., & Dohn, M. C. (2003). Redescription disembeds relations: Evidence from relational transfer and use in problem solving. *Memory & Cognition, 31,* 1082-1093.

Duke, T., & Julicher, F. (2003). Active traveling wave in the cochlea. *Physics Review Letters, 90,* 158101.

Elman, J. L. (2005). Connectionist models of cognitive development: Where next? *Trends in Cognitive Sciences, 9*, 112-117.

Englander, S. W., Kallenbach, N. R., Heeger, A. J., Krumhansl, J. A., & Litwin, A. (1980). Nature of the open state in long polynucleotide double helices: Possibility of soliton excitations. *Proceedings of the National Academy of Science, 77,* 7200-7226.

Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. (2001). A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience, 13,* 1-10.

Fajen, B. R. (2005). The scaling of information to action in visually guided braking. The *Journal of Experimental Psychology: Human Perception and Performance, 31*, 1107-1123.

Farmer, J. D. (1990). A rosetta stone for connectionism. *Physica D, 42*, 153-187.*f*

Fearing, F. (1970). *Reflex action: A study in the history of physiological psychology*. Cambridge, MA: MIT Press. (Original work published 1930)

Fodor, J. A., & Pylyshyn, Z. W. (1981). How direct is visual perception? Some reflections on Gibson's "Ecological Approach". *Cognition, 9,* 139-196.

Fowler, C., & Saltzman, E. (1993). Coordination and coarticulation in speech production. *Language and Speech, 36*, 171-195.

Frankish, C. (1985). Modality-specific grouping effects in short-term memory. *Journal of Memory and Language, 24,* 200-209.

Frankish, C. (1989). Perceptual organization and precategorical acoustic storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 469-479.

Freeman, W. J. (2000). *How brains make up their minds*. New York: Columbia University Press.

Gibbs, R. (1999). *Intentions and the experience of meaning*. New York: Cambridge University Press.

Gibbs, R. (2006). *Embodiment and cognitive science*. Cambridge, MA: MIT Press.

Gibson, J. J. (1977). *The theory of affordances*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gilden, D. H. (2001). Cognitive emissions of 1/f noise. *Psychological Review, 108,* 33-56.

Giles, H., Coupland, J., & Coupland, N. (1991). *Contexts of accomodation*. New York: Cambridge University Press.

Glanzer, M., & Cunitz, A.R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5,* 351-360.

Godfrey, K. B., & Swindale, N. V. (2007). Retinal wave behavior through activity-dependent refractory periods. *PLoS* ***??*** *Computational Biology, 3,* 2408-2420.

Goldinger, S. D., & Azuma, T. (2003). Puzzle solving science: The quixotic quest for units in speech perception. *Journal of Phonetics, 31,* 305-320.

Goldstein, L., Pouplier, M., Chen, L., Saltzman, E., & Byrd, D. (2007). Dynamic action units slip in speech production errors. *Cognition, 103,* 386-412.

Grill-Spector, K., & Kanwisher, N. (2006). Visual recognition: As soon as you know it is there, you know what it is. *Psychological Science, 16,* 152-160.

Grossberg, S. (1980). How does the brain build a cognitive code? *Psychological Review, 81(1)*, 1-51.

Haken, H. (2000). *Information and self-organization: A macroscopic approach to complex systems*. Berlin: Springer.

Haselager, W. F. G., Bongers, R. M, & van Rooij, I. (2003). Cognitive science, representations and dynamical systems theory. In W. Tschacher, & J-P. Dauwalder (Eds.), *The dynamical systems approach to cognition*, 229-241. River Edge, NJ: World Scientific.

Hausdorff, J. M., Zemany, L., Peng, C.-K., & Goldberger, A. L. (1999). Maturation of gait dynamics: Stride-to-stride variability and its temporal organization in children. *Journal of Applied Physiology, 86*, 1040-1047.

Hitch, G. J., Burgess, N., Towse, J. N., & Culpin, V. (1996). Temportal grouping effects in immediate recall: A working memory analysis. *Quarterly Journal of Experimental Psychology*, *49A,* 116-139.

Hofstede, B. T. M., & Kolk, H. J. (1994). The effects of task variation on the production of grammatical morphology in Broca's aphasia: A multiple case study. *Brain and Language, 46*, 278-328.

Holden, J. C. (2002). Fractal characteristics of response time variability. *Ecological Psychology, 14,* 53-86.

Holden, J. C. (2005). Gauging the fractal dimension of response times from cognitive tasks. In M. A. Riley & G. C. Van Orden (Eds.), *Contemporary nonlinear methods for behavioral scientists: A tutorial,* 267-318. Retrieved March 1, 2005, from http://www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jsp

Holden, J. C., Van Orden, G. C., & Turvey, M. T. (2008). *How component processes interact in reading aloud.* (submitted for publication).

Huey, E. B. (1908). *The psychology and pedagogy of reading*. Cambridge, MA: MIT Press.

Humphreys, M. S., Burt, J. S., & Lawrence, S. (2001). Expecting *dirt* but saying *dart*: The creation of a blend memory. *Psychonomic Bulletin & Review, 8*, 820-826.

Huttenlocher, J., Hedges, L.V., & Prohaska, V. (1992). Memory for day of the week: A 5 + 2 day cycle. *Journal of Experimental Psychology: General, 121,* 313-326.

Iliev, I. D., Khristov, E. K., & Kirchev, K. P. (1994). *Spectral methods in soliton equations*. New York: Longman Scientific & Technical.

Infeld, E., & Rowlands, G. (1990). *Nonlinear waves, solitons, and chaos*. Cambridge MA: Cambridge University Press.

James, W. (1890). *The principles of psychology*. New York: Henry Holt and Co, Inc.

Jordan, J. S. (2003). The embodiment of intentionality. In W. Tschacher, & J-P. Dauwalder (Eds.), *The dynamical systems approach to cognition*, 201-227. River Edge, NJ: World Scientific.

Juarrero, A. (1999). *Dynamics in action.* Cambridge, MA: MIT Press.

Kauffman, S. A. (1993). *The origins of order*. New York: Oxford University Press.

Kauffman, S. A. (2000). *Investigations*. New York: Oxford University Press.

Keijzer, F. (2003). Self-steered self-organization. In W. Tschacher, & J.-P. Dauwalder (Eds.), *The dynamical systems approach to cognition*, 243-259. River Edge, NJ: World Scientific.

Kello, C. T., Anderson, G. G., Holden, J. G., & Van Orden, G. C. (in press). The pervasiveness of 1/f scaling in speech reflects the metastable basis of cognition. *Cognitive Science*..

Kello, C. T., Beltz, B. C., Holden, J. G., & Van Orden, G. C. (2007). Tne emergent coordination of cognitive function. *Journal of Experimental Psychology: General, 136,* 551-568.

Kelso, J. A. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT Press.

Kelso, J. A. (2003). Cognitive coordination dynamics. In W. Tschacher, & J-P. Dauwalder (Eds.), *The dynamical systems approach to cognition*, 45-67. River Edge, NJ: World Scientific.

Klatt, D. H. (1989). *Lexical representation and process*. Cambridge, MA: MIT Press.

Kloos, H., Silverman, A., & Fenstermaker, B. (2007). *Preschoolers' understanding of density: Type of knowledge or effect of task context?* Poster presented at the Cognitive Development Society, Santa Fe, NM.

Kolk, H. H. J., & Heeschen, C. (1992). Agrammatism, paragrammatism and the management of language. *Language & Cognitive Processes, 7,* 89-129.

Koroleva, V. I., & Bures, J. (1979). Circulation of cortical spreading depression around electrically stimulated areas and epileptic foci on the neocortex of rats. *Brain Research, 173,* 209-215.

Kugler, P. N., & Turvey, M. T. (1987). *Information, natural law, and the self-assembly of rhythmic movement*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Landy, D., & Goldstone, R. L. (2007) How abstract is symbolic thought? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 720-733.

Langacker, R. W. (1987). *Foundations of cognitive grammar*. Stanford, CA: Stanford University Press.

Maia, T. V., & Cleeremans, A. (2005). Consciousness: Converging insights from connectionist modeling and neuroscience. *Trends in Cognitive Sciences, 9*, 397-404.

Mandler, G. (1984). *Mind and body: Psychology of emotion and stress*. New York: Norton.

Mandler, G. (1997). *Human nature explored*. New York: Oxford University Press.

Mañé, R. (1981). On the dimension of the compact invariant set of certain nonlinear maps. In D.A. Rand & L.S. Young (Eds.), *Dynamical systems and turbulence, Warwick, 1980*. Lecture notes in mathematics 898, 230-242. New York: Springer-Verlag.

Marr, D. (1982). *Vision*. New York: Henry Holt and Co., Inc.

Maylor, E. A., Chater, N., & Brown, G. D. A. (2001). Scale invariance in the retrieval of retrospective and prospective memories. *Psychonomic Bulletin & Review, 8*, 162-167.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375-407.

McCrary, J. W., & Hunter, W. S. (1953). Serial position curves in verbal learning. *Science, 117,* 131-134.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63,* 81-97.

Motley, M. T., & Baars, B. J. (1976). Laboratory induction of verbal slips: a new method for psycholinguistic research. *Communication Quarterly, 24*, 28–34.

Necker, L. A. (1932). Observations on some remarkable phenomena seen in Switzerland and an optical phenomenon which occurs on viewing of a crystal or geometrical solid. *Philosophical Magazine, 3,* 329-337.

Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition, 18*, 251-269.

Nairne, J. S. (1991). Positional uncertainty in long-term memory. *Memory & Cognition, 19*, 332-340.

Nairne, J. S. (1992). The loss of positional certainty in long-term memory. *Psychological Science, 3,* 199-202.

Natale, M. (1975). Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology, 32,* 790-804.

Neath, I. (1998). *Human memory: An introduction to research, data, and theory*. Pacific Grove, CA: Brooks/Cole.

Neisser, U. (1968). *Cognitive psychology*. New York: Appleton-Century-Crofts.

Odell, G.M. (1980). Biological Waves. In L.A. Segal (Ed.) *Mathematical models in molecular and cellular biology*, 523-567. New York: Cambridge University Press.

Oyama, S. (2000). *The ontogeny of information: Development systems and evolution*. Durham, NC: Duke University Press.

Petterson, M. A. (1994). Object recognition processes can and do operate before figure-ground organization. *Current Directions in Psychological Science, 3*, 105-111.

Postman, L., & Phillips, L. W. (1965). Short-term temportal changes in free recall. *Quarterly Journal of Experimental Psychology, 17,* 132-138.

Pouplier, M., & Goldstein, L. (2005). Asymmetries in the perception of speech production errors. *Journal of Phonetics, 33,* 47–75.

Rauschenberger, R., Peterson, M. A., Mosca, F., & Bruno, N. (2004). Amodal completion in visual search: Preemption or context effects? *Psychological Science, 15,* 351-355.

Rhodes, T., & Turvey, M. T. (2007). Human memory retrieval as Lévy foraging. *Physica A, 385*, 255-260.

Rock, I. (1957). The role of repetition in associative learning. *The American Journal of Psychology, 70*, 186-193.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 803-814.

Rubin, D. C., Hinton, S., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review*, *103,* 734-760.

Ryan, J. (1969a). Grouping and short-term memory: Different means and patterns of grouping. *Quarterly Journal of Experimental Psychology, 21,* 137-147.

Ryan, J. (1969b). Temportal grouping, rehearsal, and short-term memory. *Quarterly Journal of Experimental Psychology, 21,* 148-155.

Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics, 25*, 421-436.

Sataric, M. V., Zakula, R. B., Ivic, Z., & Tuszynski, J. A. (1991). Influence of a solitonic mechanism on the process of chemical catalysis. *Journal of Molecular Electronics, 7,* 39-46.

Schneider, E. D., & Kay, J. J. (1994). Life as a manifestation of the second law of thermodynamics. *Mathematical and Computer Modeling, 19*, 25-48.

Schultz, A., & Tabor, W. (2008). *Reaction time distributions in sentence processing*. Unpublished manuscript.

Shallice, T. (1988). *From neuropsychology to mental structure*. New York: Cambridge University Press.

Shannon, C. E., & Weaver, W. (1948). *The mathematical theory of communication*. Chicago: University of Illinois Press

Shaw, R. E. (2001). Processes, acts, and experiences: Three stances on the problem of intentionality. *Ecological Psychology, 13*, 275-314.

Shockley, K., Santana, M. V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 326-332.

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

Stephen, D. G., Dixon J. A., & Isenhower, R. W. (2007a). *Dynamics of representational change: Entropy, action, and cognition*. (Submitted for publication).

Stephen, D. G., Boncoddo, R. A., Dixon, J. A., & Magnuson, J. S. (2007b). *The dynamics of insight: Phase transitions and entropy in mathematical discovery*. (Submitted for publication).

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276-315.

Street, B. V. (1984). *Literacy in theory and practice*. New York: Cambridge University Press.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-661.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Swenson, R. & Turvey, M. T. (1991). Thermodynamic reasons for perception–action cycles. *Ecological Psychology, 3,* 317–48.

Takens, F. (1981). Detecting strange attractors in turbulence. In D.A. Rand & L.S. Young (Eds.), *Dynamical Systems and Turbulence, Warwick, 1980*. Lecture notes in mathematics 898, 366-381. New York: Springer-Verlag.

Thornton, T. L., & Gilden, D. L. (2005). Provenance of correlations in psychological data. *Psychnomic Bulletin & Review, 12*, 409-441.

Trudeau, J., & Dixon, J. A. (in press). Embodiment and abstraction: Actions create relational representations. *Psychonomic Bulletin and Review*.

Tuller, B. (2005). Categorization and Learning in Speech Perception as Dynamical Processes. In M. Riley & G. C. Van Orden (Eds.), *Tutorials in contemporary nonlinear methods for the* *behavioral sciences,* 353-400. Retrieved March 1, 2005, from http://www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jsp

Turvey, M. T. (1990). Coordination. *American Psychologist, 45*, 938-953.

Turvey, M. T. (1992). Affordances and prospective control: An outline of the ontology. *Ecological Psychology, 4*, 173-187.

Turvey, M. T. (2004). Impredicativity, dynamics, and the perception-action divide. In V. K. Jirsa & J. A. S. Kelso (Eds.), *Coordination dynamics: Issues and trends: Vol. 1. Applied complex systems*, 1-20. New York: Springer-Verlag.

Turvey, M. T., & Moreno, M. A. (2006). Physical metaphors for the mental lexicon. *The Mental Lexicon*, *1,* 7-33.

Turvey, M. T., & Shaw, R. E. (1995). Toward an ecological physics and a physical psychology. In R. Solso, & D. Massaro (Eds.), *The science of mind: 2001 and beyond,* 144-172. New York: Oxford University Press.

Tzelgov, J. (1997). Specifying the relations between automaticity and consciousness: A theoretical note. *Consciousness and Cognition, 6*, 441-451.

Van Gelder, T., & Port, R. F. (1995). It's about time: An overview of the dynamical approach to cognition. In R. F. Port & T. van Gelder (Eds.) *Mind as motion*, 1-43, Cambridge, MA: MIT Press.

Van Orden, G. C., Kello, C. T., & Holden, J. G. (in press). Situated behavior and the place of measurement in psychological theory. *Ecological Psychology*.

Van Orden, G. C, & Kloos, H. (2005). The question of phonology and reading. In M. S. Snowling, & C. Hulme, (Eds.). *The science of reading: A handbook.* Oxford: Blackwell Publishing, 61-78.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (2001). What do double dissociations prove? *Cognitive Science, 25,* 111–172.

Vincente, K. J. (1999). *Cognitive work analysis: Toward safe, productive, and healthy computer-based work*. Mahwah, NJ: Erlbaum.

Viswanathan, G. M., Buldryev, S. V., Afanasyev, V., Havlin, S., da Luz, M. G. E., Raposo, E.P., et al. (2001). Levy flights search patterns of biological organisms. *Physica A*, 295, 85-88.

Vollmer, F. (2001). The control of everyday behaviour. *Theory and Psychology, 11*, 637-654.

Walter, G. (1953). *The living brain*. London, England: Penguin Books.

Warren, W. H. (2006). The dynamics of perception and action. *Psychological Review, 113,* 358-389.

Watson, J. M., Balota, D. A., & Roediger, H. L. (2003). Creating false memories with hybrid lists of semantic and phonological associates: Over-additive false memories produced by converging associative networks. *Journal of Memory & Language, 49*, 95-118.

Webber, C. L., Jr., & Zbilut, J. P. (2005). Recurrence quantification analysis of nonlinear dynamical systems. In M. A. Riley & G. C. Van Orden (Eds.), *Tutorials in contemporary nonlinear methods for the behavioral sciences*, 26-94. Retrieved March 1, 2005, from http://www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jsp

Wickens, T. D. (1999). Measuring the time course of retention. In C. Izawa (Ed.) *On human memory: Evolution, progress, and reflections on the 30th anniversary of the Atkinson-Shiffrin model,* 245-266. Mahwah, NJ: LEA.

Wijnants, M. L., Bosman, A. M. T., Hasselman, F., Cox, R., & Van Orden, G. C. (2008). *Massed practice of precision aiming and 1/f scaling*. (submitted for publication).

Woodward, R. S., & Schlosberg, H. (1965). *Experimental psychology: Revised Edition.* New York: Holt, Rinehart & Winston.

Yakushevich, L. V. (2001). *Nonlinear physics of DNA*. Chichester, England: John Wiley and Sons.

Figure Caption: Necker Cube

Fix your eyes on the green dot and watch the cube fluctuate between different orientations.