FORMALIZING LEARNING AS A COMPLEX SYSTEM:

SCALE INVARIANT POWER LAW DISTRIBUTIONS IN GROUP AND INDIVIDUAL DECISION MAKING

Thomas Hills, The University of Texas at Austin*
Andrew C. Hurford, The University of Texas at Austin
Walter M. Stroup, The University of Texas at Austin
Richard Lesh, Indiana University

* corresponding author - The University of Texas at Austin, Department of Curriculum and Instruction, 1 University Station D5705, Austin, TX 78712-0382; phone: (512) 232-9683; email: wstroup@mail.utexas.edu

For every complex problem there is an
answer that is clear, simple, and wrong.

– H. L. Mencken

Introduction

There is a tendency in education to answer the complex problem of learning, including learning in group contexts like classrooms, with “clear” and “simple” pedagogical methods like direct instruction (Silbert et al., 1997). Rather than fall into the trap outlined by Mencken, we look to move in a quite different direction. We start by acknowledging and respecting that learning is complex. As a result, we specifically look to investigate the use of complexity theory in education. An important distinction needs to be made, as there are two broad ways in which complexity theory is engaged in the education literature. One way complexity theory is discussed relates to teaching and learning *about* complex systems (e.g., Resnick & Wilensky, 1998, Jacobsen, 1998, Hmelo et al., 2000; Wilenksy & Stroup, 1999;). The other relates to learning *as* a complex system (e.g., Ennis, 1992; Hurford, 1998; Thelen & Smith, 1996). Although “learning about” and “learning as” are apt to be mutually informative and learning about complexity and systems analyses is arguably an increasingly important candidate for school curricula, the primary focus in this discussion is on learning *as* a complex dynamical system.

Consistent with the tenor of some of the prior analyses, we begin with a top-level discussion of why complexity-based approaches might be seen to better fit the phenomenology of learning. What is distinctive is that we then go on to take up one of the more challenging aspects of making the case for the applicability of complexity theory to learning: identifying and formally investigating instances of scale invariance (or self-similarity) in classroom-situated decision making. Using some of the formalisms associated with complexity theory, we identify and investigate learning behaviors that are self-similar across scale from individuals, to groups, and to groups of groups. We believe one of the accomplishments of this line of work is a shift from saying learning and cognition are, in certain ways, *like* complex systems to showing that they actually satisfy some of the more stringent formal conditions for *being* complex systems.

In a sense then, we are responding to the challenge of formalization made by one of the theorists of complexity theory:

> …common usage of the term *complex* is informal. The word is typically employed as a name for something that seems counterintuitive, unpredictable, or just plain hard to pin down. So if it is a genuine *science* of complex systems we are after and not just anecdotal accounts based on vague personal opinions, we're going to have to translate some of these informal notions about the complex and the commonplace into a more formal, stylized language, one in which intuition and meaning can be more or less faithfully captured in symbols and syntax. (Casti, 1994, p. 270)

Can we move beyond "anecdotal accounts based on vague personal opinions" and translate some of our informal notions about learning as a complex system into "formal, stylized language, … more or less faithfully captured in symbols and syntax"? In order for the efforts to engage learning as a complex system to gain traction, we believe we must respond to this challenge. Accordingly, the data sets and analyses we present here are, so far as we know, the first time human behaviors have been shown, in formalized language, to be scale-invariant both at the level of the group and at the level of the individual and in ways that are specifically related to underlying cognitive processes predictive of success in a problem-solving environment.

To the extent that the results presented here generalize to learning in other environments, we are inclined to ask in what ways is learning complex? We attempt to provide an answer for this based on existing theoretical work in other complex sciences. This has implications for understanding the combination of internal and external constraints that are most likely to engage cognition as a complex adaptive system, and what implications for activity design might follow from designing with and/or for complexity. This allows to further engage the question of when learning is not complex. By engaging the formal features of the learning system that would result in the emergence of the observed power law relations we present here, we are further able to address both the developmental features of learning as a complex adaptive system and the kinds of environments that support that system.

We start with a discussion of why we think approaching learning as a complex system rather than as linear and exclusively individualistic process is appropriate. We then present our empirical results and conclude with a discussion of these results.

Plausibility Arguments for Treating Learning as a Complex System

The great majority of theory building in constructivist and cognitive learning theories has been focused on the learning of an individual. Behaviorist perspectives (Skinner, 1954; Stein, Silbert, & Carnine, 1997), information processing perspectives (Anderson, 1983; Anderson, Reder, & Simon, 2000; Mayer, 1996), novice-expert perspectives (Chi, Feltovich, & Glaser, 1981; NRC, 1999, chap. 2; Reiner, Slotta, Chi, &

Resnick, 2000), schema-theoretic perspectives (Derry, 1996; diSessa, 1993), and constructivist perspectives (Cobb, 1994; Ernest, 1996; Piaget[i], 1923/1959, 1924/1969,1929/1951; Vygotsky, 1987, chap. 6) are all predominantly individualistic views of learning. These efforts have provided many insights and have been very successful in helping researchers to build useful models of learning. Although individualistic approaches to learning have been quite productive they also have several limitations.

Limitations of Individualistic and Linear Theories of Learning

The first limitation of individualistic theories of learning is that they do not "scale" well–that is, the learning of a classroom of students is not very profitably described as the linear combination of a number of individual learners. This type of scaling to whole classrooms of learners does not and cannot take into account the complex interactions and synergetic effects derived from the properties of groups. Very little of what goes on in classrooms can be understood in terms of straightforward cause and effect relationships among simple aggregations of individual learners.

Individualized and linear models of learning tend to be more static than dynamic. Behaviorist models (e.g., Stein, Silbert, & Carnine, 1997, pp. 3-29) assume that learning is the simple accumulation of fixed and appropriately sized knowledge bits that are taken in as given, without much, if any, active adaptation or interpretation on the part of the learner. Another line of (individualized) learning research posits that learners possess relatively static conceptual structures and that teaching and learning constructs knowledge structures and then repairs or replaces "misconceptions" (Reiner, Slotta, Chi, & Resnick, 2000, p. 7) with increasingly "expert" structures. Though, not enough is said in this literature about how these transformations actually take place. In each of these lines of research, knowledge is envisioned as bits of information stored in and accessed from static conceptual structures internal to individual "knowers".

Individualized approaches also tend to focus on a learner at the expense of the learner's context and her or his membership in a learning community. There are many aspects of the surrounding contextual situation that influence how learning takes place and what gets learned (Lave, 1988; Lave &Wenger, 1991; Wertsch, del Rio, & Alvarez, 1995). Students and teachers are embedded in a wide variety of social, historical, and cultural systems (cf., Bowers, Cobb, & McClain, 1999; Hiebert, et al., 1996, p. 19; Lave & Wenger, 1991, pp. 67-69) that profoundly affect learning (Cobb, Perlwitz, & Underwood, 1996) and individualized approaches to learning generally overlook these important and complex influences. The sociocultural historical milieu of the classroom can be seen as part of the environment relative to which adaptation occurs. At any given moment, classrooms and learners are immersed in a wide variety of interconnected and often competing activities and goal structures. Theories of learning that focus on individuals generally do not take these kinds of complex and ubiquitous learning conditions into account.

Finally, individualized accounts of learning do not offer very much to teachers in the way of helping them to make sense of or design for whole-classroom activities. Although teachers may develop individualized educational plans, they almost never design classroom activities with a single individual in mind. Classroom activity is inherently a *group activity*, and there is very little in the language and ideas of individualized cognitively-based or even standard constructivism-based learning theory that enables teachers to make sense of the activities of groups of learners.

Affordances of Systems-theoretic Approaches to Learning

In contrast with individualized approaches, dynamical systems-theoretical perspectives have much to offer in terms of helping teachers, researchers, and others to focus on and make sense of learning at the level of the group. First, a systems perspective enables thinking about classroom learning in terms of a dynamic, continuously changing "dance" between the group, its members, and the contextual situation. Second, as discussed above, classrooms are much more than a linear sum of individual learners, and a systems perspective enables thinking about the synergetic affordances and "lever points" (Holland, 1995, p. 39) inherent in classrooms. Third, it may well be that the most important affordance of systems-theoretical approaches to learning is in the *language* of complexity itself.

Complex systems terminology and the ideas that terminology represents are increasingly finding their way into the discussions and literature of cognitivist, constructivist, and sociocultural learning theory camps. It seems that the use of the language of complexity is preceding more rigorous and careful application of systems-theoretical tenets–Casti's (1994) "formalization" (pp. 274-276)–to perspectives on learning. The hope is that use of the language will serve as precursor and enabler of more systematic complexity-based modeling in future education research. In any event, the combination of the fundamental ability to address dynamic and complex interrelations, the ability to provide a powerfully descriptive language for talking about what happens cognitively and in classrooms, and the potential for formalization of these things into useful models all serve to demonstrate the potential of systems-theoretical approaches for research into learning.

Systems perspectives *do* scale well in terms of considerations of group activity. In one way or another, every dynamical systems viewpoint addresses both the individual and the aggregate. For example, from the structuralist perspective of Jean Piaget[ii] (1968/1970, chap. 2), the group and the elements of the group are mutually constitutive. That is, the dynamic creation of a group, the activity of individual members of the group, and the group's context each influence the other, forming a complex adaptive system[iii]. An example of this in a classroom is when students are aware (or quickly become aware) of their status within the larger group, and those status considerations have powerful effects on the students' and the group's subsequent activities (cf. Empson, 2003). Complex systems analyses (Casti, 1994; Camazine, et al., 2001; Clark, 1997; Holland, 1995, 1998; Stroup & Wilensky, 2000; Prigogine, 1984, 1997) focus on higher-level patterns (e.g., aggregation, flows) that are generated by activity and adaptation at the level of individuals whose behaviors are based solely on the local environment and the individual's own internal models. In contrast to individualized theories of learning, systems-theoretical points of view are fundamentally concerned with viewing learners and groups as mutually constitutive agents whose behavior is to be understood in the context of their larger patterns of activity.

Systems-theoretical points of view tend to be very dynamic–characterizing activity in terms of evolving *patterns*. Rather than focus on static "snapshots" of individual student's or the group's learning, such as quiz grades or end-of-year tests, we focus our investigations of learning in the context of on-going problem-solving behavior. To this end, we recorded the decisions made by all students in a multi-player network simulation as they engaged an open-ended problem concerning the nature of unseen resource distributions. Our results show individuals and groups both exhibited power law distributions for inter-decision intervals of longer than three seconds, and the length of these inter-decision intervals was correlated with overall success in the game. Individuals and classrooms also showed variation in their fit to a power law (Zipf) distribution and our analysis suggests that this variation is informative with respect to learning and individual strategies.

Introduction to Power Laws

Power law distributions are characterized by self-organized arrangements with many small events and few large events. The distributions are often called scale-free because there is no limit to the largest size event and no natural scale of measurement. The relationship between neighboring elements or events in a distribution is preserved regardless of the scale of observation. Scale-free phenomenon that are ordered in this way spatially are called fractal. Temporal distributions of similar arrangement exhibit a behavior known as one-over-f (1/f) noise.

Power law distributions are found in diverse self-organizing systems ranging from forest fires to web links (e.g., forest-fires, Malamud et al., 1998; corporate firm sizes, Axtell, 2001; avalanches, Bak et al., 1988; earthquakes, Johnston & Nava, 1985; biological extinctions, Raup, 1986; stock-market fluctuations, Mandelbrot, 1982; volcanic activity, Diodati et al., 1991; the organization of the world-wide-web, Huberman & Adamic, 1999; city sizes, Ioannides & Overman, 2000; turning-times in fruit flies, Cole, 1995; also see Bak, 1994). Power law distributions have also been observed in human reaction times when learning procedural tasks, for example, how to roll cigars ("the power law of learning", Neves & Anderson, 1981; Anderson, 1982; Logan, 1988), and in retention ("the power law of forgetting", Anderson & Schooler, 1991; Rubin & Wenzel, 1996; Wixted & Ebbesen, 1991).

While power law distributions have not yet been described for higher level cognitive processes (i.e., above reaction times) or for social problem solving, power law relationships have been observed for distributions of words in human language. This is designated as Zipf's law (Zipf, 1949) and is mathematically representative of a pareto power law distribution (Adamic & Huberman, 2002). When words of a given language are ordered by decreasing frequency, the frequency of the *i*th word, P(*i*), is proportional to i^{-A}, where A is approximately equal to one; this is mathematically identical to 1/f noise described above. This is suggestive of cognitive constraints associated with human symbolic manipulations, where the use of a new symbol is based on the frequency-dependent relationship it shares with the symbolic population in which it originates. Mechanisms for the evolution of power laws in human communication have been described by Cancho & Sole (2003) and these authors suggest that power law distributions may be "required by [all] symbolic systems."

Methods

To investigate patterns in human 'thinking' about complex problems, we collected data from five separate classrooms participating in a biology related foraging simulation (Hills & Stroup, 2004) built in the Hubnet NetLogo environment (Wilensky, 1999; Wilensky & Stroup, 1999; Wilensky & Stroup, 2000) in which individuals controlled the search patterns of individual foragers (avatars) in a 2-D playing field , hunting for invisible pixels that were grouped in small clumps (Figure 1). Students sat at individual laptops and could move their avatars up, down, left, or right for each individual move. The group space was projected at the head of the classroom and the simulation was fast enough that avatars responded in less than 500ms. Each run of the simulation lasted at least three minutes. During this participatory simulation, the students were unaware of the position of the pixel 'resources' but did know the location of their individual avatar as well as the position of other players (Figure 1, shows orange food 'pixels'(individual boxes), but students could not see them).

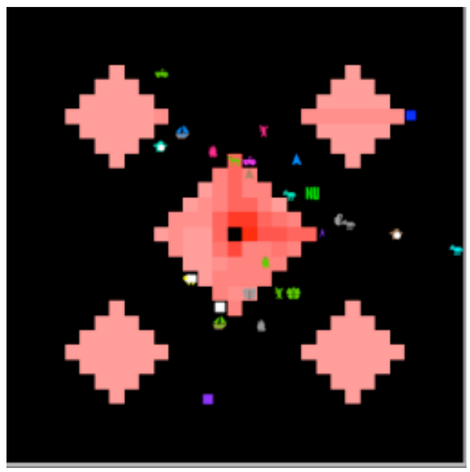

Figure 1. The resource distribution in which students foraged.

As a consequence of moving their avatars around, their computers reported to them when they encountered food (but still did not make the food visible), thus presenting them with the challenge of cognitively reconstructing the size and arrangement of resource clumps. Using this information, they could then adapt their foraging behavior to maximize resource collection.

Students could also listen to the verbal reports and actions of other students in the classrooms. All students shared the same resource patterns, so information from other students could be informative about distributions. Separate resources were allocated for each individual, so students were not competing for resources. However, students were asked at 30 second intervals to vocally report how many pixels of food they had acquired (reported on their individual monitor). With this information students could alter their strategies in accordance with their relative success or failure during reporting intervals. During the simulation we recorded all information about avatars, including position, cumulative resources found, and wait times between individual actions.

Results

Figure 2 shows the typical relationship observed between average wait times for individual decisions and the cumulative resources found over the entire simulation. We use the term “wait times” here because “reaction time” often implies a minimum time to arrive at a correct answer (“How fast can one make a decision?”). Figure 2 shows that students who took longer to make decisions did better in the simulation.

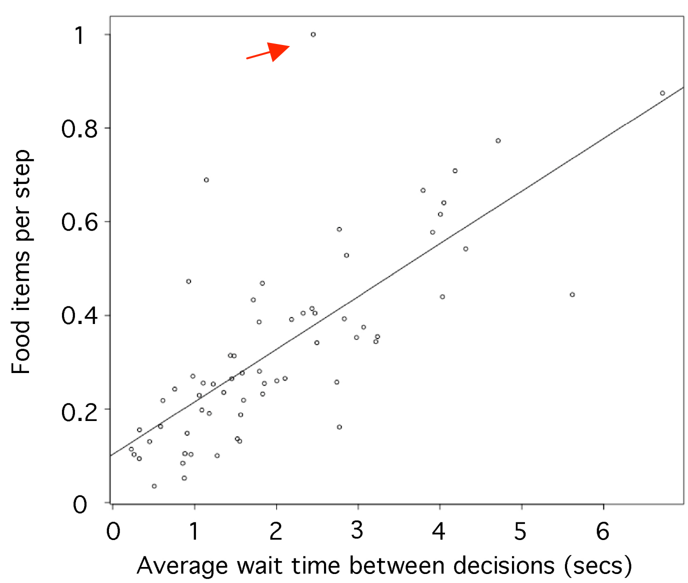

Figure 2: Students who took longer to make decisions, found more resources per step in the game. Only one student took fewer than 20 steps (marked by red arrow). All other students took more than twenty steps and were observed to actively search for the duration of the simulation.

Students who were interacting with the simulation with intervals under one second were observed to be primarily interested in moving as quickly as possible through the resource space with little attention as to direction or distribution of the underlying resources. After the simulation, students self-reported that longer wait-times were used to “think about where the resources might be.” Using the student’s terminology, longer periods of thinking correlated with higher success in the problem environment.

To determine if students’ behaviors were associated according to a power law distribution, we took the log of individual decision times and plotted them against the log of the rank of individual decision times. The rank is simply a rank ordering of decision times by their duration. Figure 3 shows the Zipf plot for all decisions made by students in Figure 2.

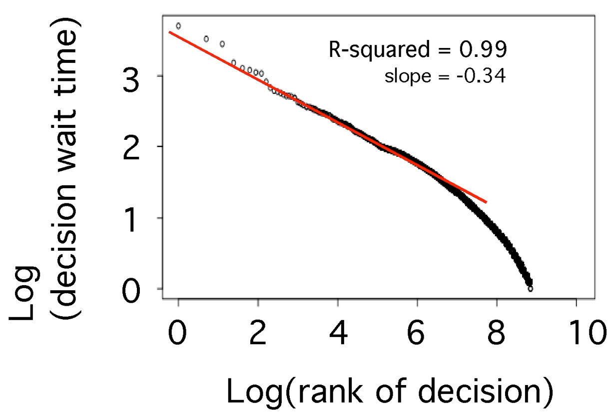

Figure 3: The Zipf distribution for all individual decisions (n = 23,637). Circles represent wait times between individual decisions. A linear regression fit to decisions longer than 3.3 seconds fits with an R-squared of 0.99.

This wait-time graph treats all individual behaviors as members of the group behavior and is highly consistent with a Zipf distribution, characteristic of a scale-free power law distributions.

To determine the robustness of the power law distribution, we also averaged decision times for individuals and plotted these in a similar fashion in figure 4. This plot is less suggestive of a power law distribution, but does reveal a similar relationship above 3.3 seconds. It is also possible that these distributions fit a log-normal distribution better than they fit a log-log distribution (the distribution for a power law). We tested this possibility by calculating the R-squared for a log-normal distribution. For individual averages, the distribution fits a log normal with a R-squared of 0.97 over the entire distribution, but the log-log distribution fits with a R-squared of 0.90. This is not the case for the other log-log distributions shown in this paper, where log-normal distributions fit more poorly. We discuss further the interpretations of the shapes of the distributions below.

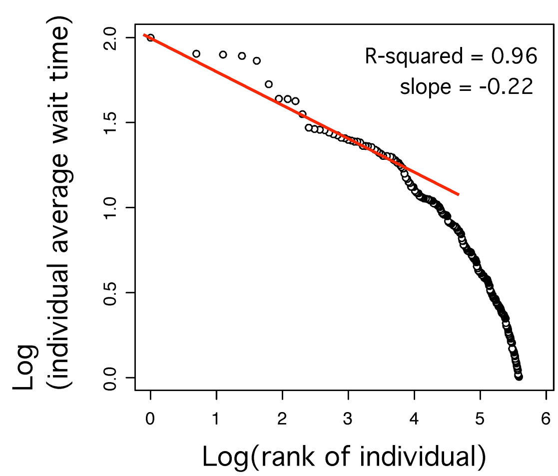

Figure 4: The Zipf distribution for average wait times per student in each simulation (n = 297). Circles represent the average wait times for individual students for a given resource distribution. A linear regression for wait times longer than 3.3 seconds fits with an R-squared of 0.96.

If individual decisions over multiple simulations fit a power law distribution with such accuracy as in Figure 1, we were also interested in understanding how individuals distributed their decisions within the class, and what an individual class looked like during a simulation. Figure 5 shows the distribution for an individual class and reveals that a student's decisions (marked with same colored circles) may show considerable variance within the class.

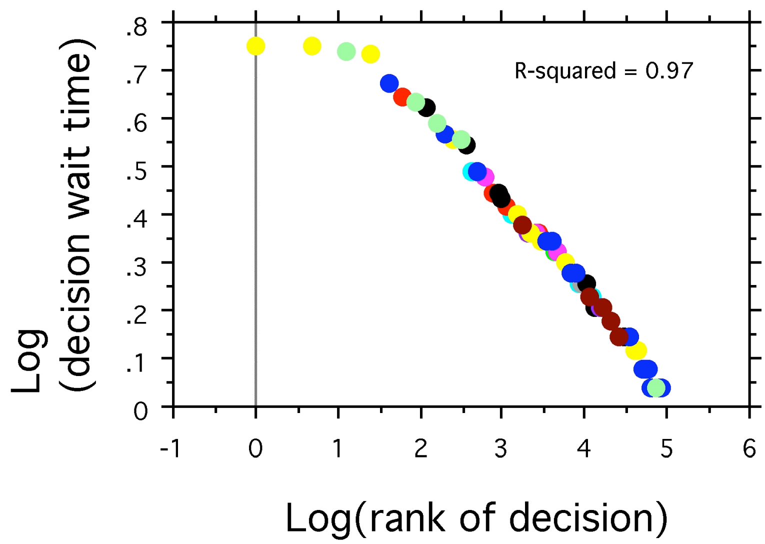

Figure 5: The Zipf plot for a single class during a single run (resource distribution). Individual decisions are represented by circles (n=139). A regression for all decision times longer than 1 second fit with an R-squared or 0.97. A single shading represents all decisions made by one individual.

This suggests that most of the variance seen in the simulation is within individuals, not between them. In other words, individuals sample from a similar subset of decision times, and are more alike than individual averages would indicate. This is shown most dramatically in Figure 5 by observing the overlap of dark and light circles. There are many students who show considerable overlap in decision timing, despite the fact that their averages are likely to be different.

To further assess the scale invariance in decisions, we plotted the log-log distributions of individual students. Figure 6 depicts the Zipf distribution for a student from Figure 5.

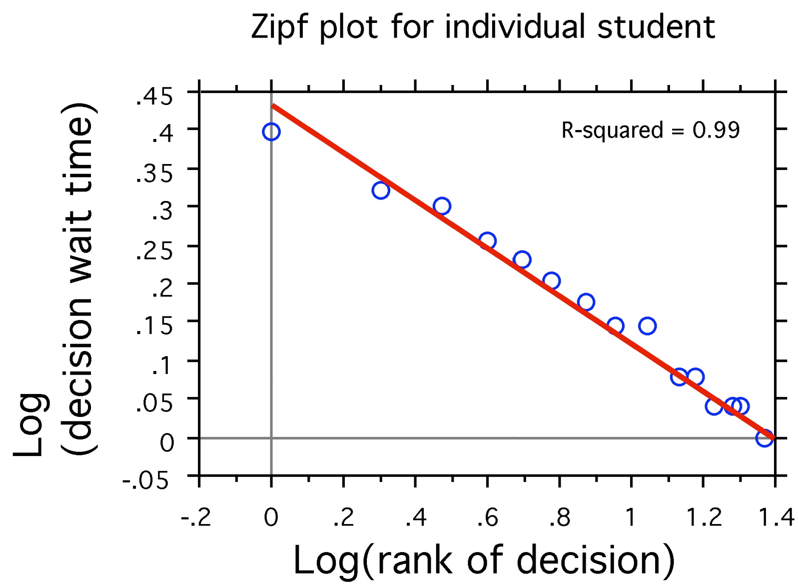

Figure 6: A Zipf plot for a single student during one simulation run. Circles represent decision times for individual decisions. A regression for all decisions longer than 1 second fit with an R-squared of 0.99.

Like many students, this student fit a power law distribution with an R-squared of 0.99. This is one of the better fits. On average students fit a log-log regression with an R-squared of 0.95. In some cases log-normal fit better, in other cases log-log distributions made the better fit.

Overall, distributions for all classes, individual classes, and individuals reveal a degree of similarity highly suggestive of scale invariance. As stated above this is one of the hallmarks of formal complex systems. Yet, log-log distributions also reveal other characteristics about the nature of dynamic problem-solving. Figure 7 shows a typical log-log plot for an individual student. For this student, the figure reveals a distinct elbow in the distribution at approximately 2 seconds. Numerous elbows at different times were observed in many of the data sets. Figures 4 and 5 both show evidence of similar elbows and it is possible that Figure 3 contains two elbows. In many cases the visual confirmation of the elbow is dramatic (as in Figure 7 and Figure 5) and it is not isolated to students or classrooms.

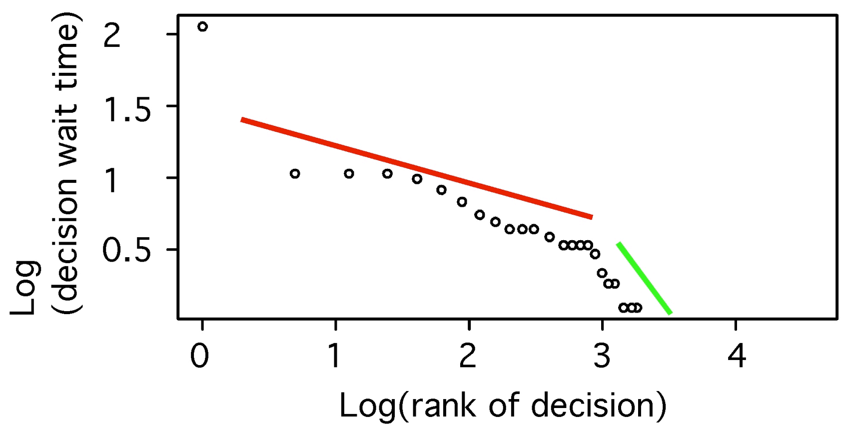

Figure 7: Zipf elbow for a single student.

The elbows appear to be associated with different kinds of behavior in the simulation. For example, in Figure 7, the steeper slope on the right is associated with shorter wait times and may therefore be associated with straighter runs across the landscape, as opposed to high angled turns, which would be more characteristic of longer wait times on the shallower sloping distribution, to the left. As a preliminary test of this hypothesis, we compared the average turning angles after decisions lasting longer than 3 seconds and after decisions which took less than 3 seconds to make. Decisions taking less than 3 seconds led to significantly straighter movement than decisions lasting longer than 3 seconds. Decisions lasting longer than 3 seconds ended on average in a right angled turn.

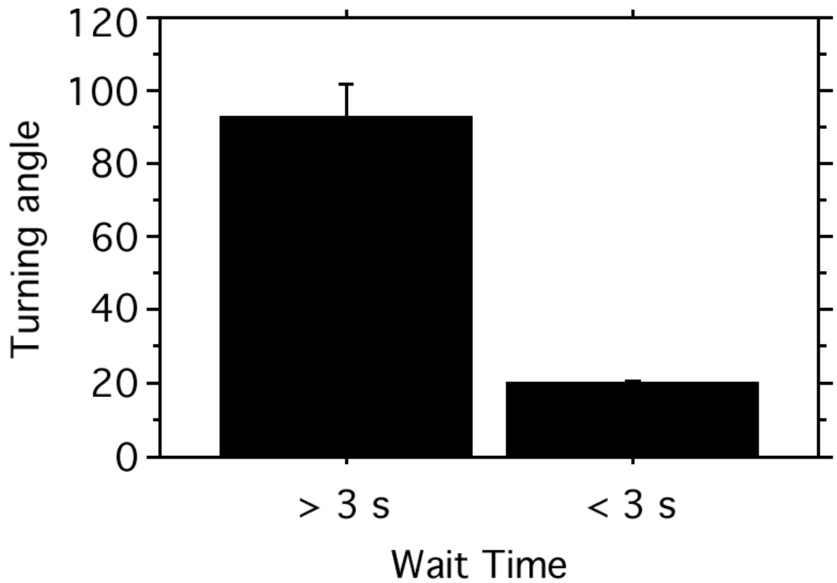

Figure 8. Comparison of turning angles taken after wait times of varying length. Decisions that took longer than 3 seconds to make are represented by '> 3 s', whereas those taking shorter than 3 seconds are represented by '< 3 s'.

It is also useful to note that individuals who performed most poorly in the simulations overall (see Figure 2), sampled from faster temporal distributions on average, which would therefore lie more on the vertical line to the right of Figure 3. Individuals who performed well in the simulations, sampled primarily form the slower temporal distributions, to the left of Figure 3.

The presence of a power law distribution was ubiquitous regardless of individual students' performances. Only five students failed to show a power law distribution for decision times lasting longer than 3 seconds. The presence or absence of a power law does not therefore appear to reveal much about the success of students in the simulation, however, the shape of the overall distribution is suggestive of other ways to characterize learning. For example, Figure 9 shows the log-log distribution for an individual student during a single run of the simulation. The presence of the elbow during the first minute of the simulation (represented by crosses) is suggestive of a binary decision making process, in which the student is choosing between two strategies, one representing local foraging with high-angled turns and the other representing straight runs with no or very low turning angles (as shown in Figure 8). However, the second minute shows a much less distinctive elbow (represted by circles), and as a whole fits a power law distribution better than the first minute, when all decision times are taken into consideration.

One possible explanation for this change in the shape of the distribution is that the student is adapting to the depletion of resources in the second minute by choosing a mixed search strategy that combines temporal features of both strategies used during the first minute. As resources are depleted, the patchy quality of the original distribution (Figure 1) begins to decay, such that the distribution of food is less spatially auto-correlated—finding one pixel is less informative about nearby pixels in the second minute than it is in the first minute. Therefore, the appropriate cognitive and behavioral response is to forage less locally around found resources, but also to rely less on running into large clumps of resources during straight runs. This is a hypothesis that would explain the difference in the early and late log-log distributions shown in Figure 9. Still, this hypothesis is based on a very early approach to cognition as a complex system, and much work remains to be done in formalizing that approach.

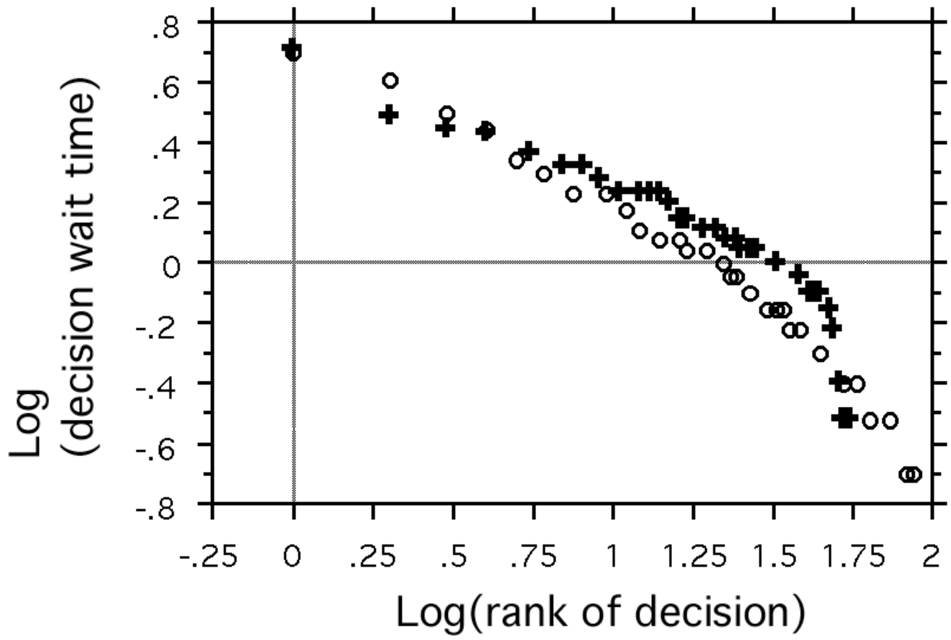

Figure 9: Data for an individual student during a single run of the simulation. Crosses represent decisions made during the first minute of the simulation, circles represent decisions made during the second minute of the simulation.

Discussion

The evidence presented here is to our knowledge the first time human behaviors have been shown to be scale-invariant both at the level of the group and of the individual and, moreover, related to underlying cognitive processes that are predictive of success in a problem solving environment. That these behaviors follow a power law distribution implies that these distributions are self-similar and cannot be accurately described by a Gaussian distribution. Gaussian distributions have exponentially decaying tails and make behaviors much larger than the mean very unlikely. However, the power law distribution gives finite probability to individuals with wait times much larger than the mean. This means that the relationship between the timing of decisions is preserved regardless of the time scale one chooses to observe the phenomenon. Thus the self-similarity.

This work also suggests that behaviors are less alike than individuals, meaning that behaviors are constructed from a more strongly multiplicative organizational scheme than are the individuals themselves, who, though different, are more alike than any two randomly chosen behaviors might indicate. This supports individualistic approaches to learning, in the sense that studying single individuals can be very informative with respect to the general strategies used by all individuals. However, it also points out the limitations of treating individuals as averages. The data we present show that the cognitive processes involved in decision making reveal higher-order patterns that have not yet been investigated. These traits are inaccessible to static or individualistic approaches that fail to treat real-time thinking processes as dynamic representations of different 'kinds' of thinking. Averaging over thinking processes eliminates the ability to differentiate between different ways of thinking, and leads to the false implication that different ways of thinking are different primarily in quantity, not quality. The data we present here hardly support this claim. Figure 9 suggests this is not true even for one individual during a two-minute time span.

The Scope of the Complexity Claim

Not all dynamic spatiotemporal processes meet the requirements of formal complex systems (Table 1). Many processes, such as traffic at airports or bank account growth, are constructed of variables represented by gaussian or other less common distributions. Future work investigating cognition as a complex system must naturally grapple with the divide between power law thinking and other kinds of thinking that are not formally complex.

In our analysis of the above data, we observed that for most individuals performance in the foraging simulation was not correlated with the quality of the log-log fit—but appeared to be correlated with other characteristics of the distribution (see the discussion of elbows above). We checked this in a given classroom where individuals repeatedly foraged in the same resource distribution. While food per step increased two-fold on average, R-squared fits to a log-log distribution were unchanged (data not shown). The only evidence that R-squared fits are related to performance is seen in individuals who show very poor fits to the log-log distributions (R-squared < 0.85). These individuals performed far more poorly than their peers in food-per-step. Though they represent but a small fraction of the individuals in our data set (n=5), we are inclined to believe that non-power law thinking holds ample information about the nature of cognition as a complex adaptive system. We are also inclined to believe that, to the extent that cognition is allowed to engage itself in tasks that nurture its formally complex features, these tasks will promote problem solving abilities in general.

Designing for complex thinking may not be trivial, but our investigations do point to specific characteristics of design that may inhibit complex thinking. Foremost, problems with limited scope, that contain but one correct answer, are likely to prevent the kinds of distributions we witness here. This is likely to be especially true of kinds of strategy switching associated with elbows. An aspect of problem solving that is seldom discussed is the ability of individuals to switch cognitive strategies on-demand. This is different from giving up on failed attempts. Maintaining and sampling from several cognitive strategies may be more closely associated with attending to multiple aspects or features of a problem simultaneously. In the foraging simulation, the resource space consisted of patches of resources and spaces in between. Being able construct and represent this space cognitively without having seen it, and to simultaneously switch between appropriate strategies while navigating that space, is clearly a feat of cognitive skill distinctly different from remembering the capital of Indiana.

Power-laws are common, but not everywhere

WHAT IS…	WHAT ISN'T…
• Corporate Firm Sizes • City Sizes • Species Extinction Rates • Avalanches • Movement times in insects • Distribution of tree sizes in forests • Links in world-wide web • Coauthorship networks	• Traffic at airports • power grid for Southern California • Standardized test scores • Social networks in small communities (high school or church) • Connections between movie actors • Growth in your bank account

Table 1. Systems that do and do not exhibit power law relationships

Structural Analysis of the Emergence of Power Laws

A number of theoretical explanations for power law distributions have been proposed. Two that may be relevant with respect to cognitive processes are the growth with preferential attachment theory proposed by Barabasi & Albert (1999) and the random stopping of exponential growth processes proposed by Reed & Hughes (2002). The Barabasi & Albert model (1999) suggests that scale-free networks emerge when networks grow with preferential attachment to vertices that are already well attached. As new vertices arrive, they enrich previously well-connected centers, creating clusters with many nodes with few attachments, and few with many attachments. The Reed & Hughes model (2002) suggests that if processes grow with exponential expectations, but are randomly stopped, then scale-free distributions will be the result. The two models are probably not mutually exclusive, but may depend on whether relationships are spatial or temporal, respectively.

It is possible that cognitive processes behave similarly to many other phenomena described by power laws because they share similar underlying growth and organizational structures. For example, at the level of neural assemblies, if there are temporally weighted distributions associated with behavior, then these distributions, when active, may fatigue at exponential rates (temporal mechanisms of this type are reviewed in Hills, 2003). If other distributions are competing for activation, some of which may lead to terminal decisions and others of which may lead through distributed networks characterized by clustering, these processes taken together may generate power law distributed decision times, as in Figure 3. If time spent 'searching' the network is correlated with greater appropriateness of the final decision, this may in turn generate the relationship observed in Figure 2.

Growth with preferential attachment also has the characteristic ring of constructivist theories of learning, which suggest that the development of cognitive schema is an organic growth process that depends on schema that are already in the system, so to speak. Learning does not come from nowhere, but must marry itself to cognitive representations that are already available to the learner. For example, the ability to conceive of different kinds of horses is facilitated by the ability to recognize horses as a species. However, upon grasping the differences between an Irish Draught horse and a Welsh Mountain pony, one gains considerable cognitive leverage with respect identifying horses in general and associating functions with their identity, which in turn opens up frames for characterizing other things and relationships. If horses are an important part of one's cognitive landscape, then new learning, and especially that kind of learning that happens outside of schools, is most likely to find itself in a cognitive relationship with horses.

The observation that individuals and classrooms share similar organizational properties with the forest fires and city sizes is promising, in multiple respects. For one, scale invariant phenomena allow one to make predictions about global behavior based on local observations. One can make predictions about the relationships between larger city sizes in the future based on observations of smaller city sizes now (Gabaix & Ioannides, 2004). As well, as we begin to formalize our understanding of cognition as a complex system, the growing literature with respect to small world networks, power laws in nature, and self-organized criticality become a valuable resource in the understanding of cognition.

References

Adamic, L. A., & Huberman, B. A. (2002). Zipf's law and the Internet. *Glotoometrics 3,* 143-150.

Anderson, J. R. (1982) Acquisition of cognitive skill. *Psychological Review 89*, 369-406.

Anderson, J.R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J, R., Reder, L. M., and Simon, H. A. (2000). Applications and misapplications of cognitive psychology to mathematics instruction. *Texas Education Review 1*, 29-49.

Anderson, J. R., & Schooler, L. J. (1991) Reflections of the environment in memory. *Psychological Science 2*, 396-408.

Axtell, R. L. (2001) Zipf distribution of U.S. firm sizes. *Science 293*, 1818-1820.

Bak, P., Tang, C., & Wiesenfield, K. (1988) Self-organized criticality. *Physica Review A 38,* 364-374.

Bak, P. (1994) Self-organized criticality: a holistic view of nature. In G. Cowan, D. Pines, and D. Meltzer, (Eds) *Complexity: Metaphors, Models, and Reality*, *SFI Studies in the Sciences of Complexity, Proceedings Vol. XIX.* Massachusets: Addison-Wesley.

Barabasi, A., & Albert, R. (1999) Emergence of scaling in random networks. *Science 286*, 509-512.

Bowers, J., Cobb, P., & McClain, K. (1999). The evolution of mathematical practices: A case study. *Cognition and instruction 17*, 25-64.

Camazine, S., Denoubourg, J., Franks, N., Sneyd, J., Theraulaz, G., & Bonabeau, E. (2001). *Self-Organization in Biological Systems*. Princeton, NJ: Princeton University Press.

Cancho, R. F., & Sole, R. V. (2003) Least effort and the origins of scaling in human language. *Proceedings of the National Academy of Sciences 100*, 788-791.

Casti, J. L. (1994). *Complexification: Explaining a paradoxical world through the science of surprise*. New York: HarperCollins Publishers, Inc.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science 5*, 121-151.

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

Cobb, P. (1994). Constructivism in mathematics and science education. *Educational Researcher 23*, 13-20.

Cobb, P., Perlwitz, M., & Underwood, D. (1996). Constructivism and activity theory: A consideration of their similarities and differences as they relate to mathematics education. In H. Mansfield, N. A. Pateman & N. Bednarz (Eds.), *Mathematics for tomorrow's young children: International perspectives on curriculum*. Boston: Kluwer Academic Publishing.

Cole, B. J. (1995) Fractal time in animal behavior: the movement activity of Drosophila. *Animal Behavior 50,* 1317-1324.

Derry, S.J. (1996). Cognitive schema theory in the constructivist debate. *Educational Psychologist 3*, 163-174.

Diodati, P. F., Marchesoni, F., & Piazza, S. (1991) Acoustic emission from volcanic rocks: an example of self-organized criticality. *Physical Review Letters 67,* 2239-2242.

diSessa, A. (1993). Toward an epistemology of physics. *Cognition and Instruction 10*, 105-225.

Empson, S., (2003). Low-performing students and teaching fractions for understanding: An interactional analysis. *Journal for Research in Mathematics Education 34*, 305-343.

Ennis, C. D. (1992). Reconceptualizing learning as a dynamical system. *Journal of Curriculum and Supervision 7*, 115-130.

Ernest, P. (1996). Varieties of constructivism: A framework for comparison. In L.P. Steffe & P. Nesher (Eds.), *Theories of mathematical learning*, 335-350. Mahwah, NJ: Erlbaum.

Gabaix, X., & Ioannides, Y. (2004) Evolution of city size distribution. In J.V. Henderson and J.-F Thisse, (Eds.) *Handbook of Regional and Urban Economics,* vol IV (North Holland, Amsterdam).

Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K., Human, P., Murray, H., Olivier, A., & Wearne, D. (1996). Problem solving as a basis for reform in curriculum and instruction: The case of mathematics. *Educational Researcher 25*, 12-21.

Hills, T. (2003) Towards a unified theory of animal event timing. In W. Meck (ed.) *Functional and Neural Mechanisms of Interval Timing*. New York:CRC Press.

Hills, T., & Stroup, W. (2004) *Cognitive exploration and search behavior in the development of endogenous representations.* Paper presented at the Annual Meeting for the American Educational Research Association, San Diego, CA.

Hmelo, C.E., Holton, D.L. & Kolodner, J.L. (2000). Designing to Learn about Complex Systems. *Journal of the Learning Sciences 9,* 247-298.

Holland, J. H. (1995). *Hidden order: How adaptation builds complexity*. New York: Addison-Wesley Publishing Company.

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

Huberman, B. A., & Adamic, L. A. (1999), Growth dynamics of the world wide web. *Nature 401,* 131.

Hurford, A. (1998). *A dynamical systems based model of conceptual change.* Paper presented at the Annual Meeting of the International Association for the Education of Teachers of Science, Minneapolis, MN.

Ioannides, Y. M. & Overman, H. G. (2003). Zipf's law for cities: an empirical examination. *Regional Science and Urban Economics 33*, 127-137.

Jacobsen, D. M. (1998). *Adoption Patterns and Characteristics of Faculty Who Integrate Computer Technology for Teaching and Learning in Higher Education.* Doctoral Dissertation, Educational Psychology, University of Calgary. [On-line]. Available: http://www.acs.ucalgary.ca/~dmjacobs/phd/diss/

Johnston, A. C., & Nava, S. J. (1985) Recurrence rates and probability distribution estimates for the New Madrid seismic zone. *Journal of Geophysics Research B 90,* 6737-6751.

Lave, J. (1988). *Cognition in practice: Mind, mathematics, and culture in everyday practice*. New York: Cambridge University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.

Logan, G. D. (1988) Toward an instance theory of automization. *Psychological Review 95,* 492-527.

Malamud, B. D., G. Morein, & D. L. Turcotte, Forest fires: An example of self-organized critical behavior. *Science 281,* 1840–1842, 1998.

Mandelbrot, B. (1982) *The Fractal Geometry of Nature.* San Francisco: W. H. Freeman.

Mayer R. E, (1996). Learners as information processors: Legacies and limitations of educational psychology's second metaphor. *Educational Psychologist 31*, 153-55.

National Research Council, (1999). *How people learn: Brain, mind, experience, and school*. Committee on Developments in the Science of Learning. J. D. Bransford, A.L. Brown, and R.R. Cocking, (Eds.). Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Neves, D. M., & Anderson, J. R. (1981) Knowledge compilation: mechanisms for the automization of cognitive skills. In J. R. Anderson (Ed.), *Cognitive Skills and Their Acquisition* (pp. 57-84). Hillsdale, NJ: Erlbaum.

Piaget, J. (1923/1959). *The language and thought of the child*. London: Routledge and Kegan Paul.

Piaget, J. (1924/1969). *Judgment and reasoning in the child*. Totowa, NJ: Littlefield Adams.

Piaget, J. (1929/1951). *The child's conception of the world*. London: Routledge and Kegan Paul

Piaget, J. (1968/1970). *Structuralism*. New York: Basic Books, Inc.

Prigogine, I., & Stengers, I. (1984). *Order out of chaos: Man's new dialogue with Nature*. New York: Bantam Books.

Prigogine, I., & Stengers, I. (1997). *The end of certainty: Time, chaos, and the new laws of nature*. New York: The Free Press.

Raup, M. D. (1986) Biological extinction in earth history. *Science 251*, 1530-1532.

Reed, W. J. & Hughes, B. D. (2002). From gene families and genera to incomes and internet file sizes: why power laws are so common in nature. *Physical Review E 66*, 067103.

Reiner, M., Slotta, J. D., Chi, M. T. H., & Resnick, L. B. (2000). Naive physics reasoning: A commitment to substance-based conceptions. *Cognition & Instruction 18*, 1-34.

Resnick, M., & Wilensky, U. (1998). Diving into complexity: Developing probabilistic decentralized thinking through role-playing activities. *Journal of the Learning Sciences 7*, 153-171.

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

Silbert, J., Stein, M., & Carnine, D. (1997). *Designing effective mathematics instruction: A direct instruction approach (3rd ed.)*. Upper Saddle River, NJ: Prentice-Hall Inc.

Skinner, B.F. (1954). The science of learning and the art of teaching. *Harvard Educational Review 24*, 86-97.

Stein, M., Silbert, J., & Carnine, D. (1997). *Designing effective mathematics instruction: A direct instruction approach, 3rd Ed.* Upper Saddle River, NJ: Prentice-Hall, Inc.

Stroup, W. M., & Wilensky, U. (2000). Assessing learning as emergent phenomena: Moving constructivist statistics beyond the bell curve. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research in mathematics and science education*. Mahwah, NJ: Lawrence Erlbaum Associates.

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

Wertsch, J. V., del Rio, P., & Alvarez, A. (Eds.), (1995). *Sociocultural studies of mind*. New York: Cambridge University Press.

Wilensky, U. (1999). NetLogo. *http://ccl.northwestern.edu/netlogo.* Center for Connected Learning and Computer-Based Modeling. Northwestern University, Evanston, IL.

Wilensky, U. & Stroup, W. (1999). Learning through Participatory Simulations: Network-based Design for Systems Learning in Classrooms. *Proceedings of the Computer Supported Collaborative Learning Conference, Stanford University,* *December.*

Wilensky, U., & Stroup, W. (2000) Networked Gridlock: Students Enacting Complex Dynamic Phenomena with the HubNet Architecture. *Proceedings of the Fourth Annual International Conference of the Learning Sciences, Ann Arbor, MI, June*.

Wixted, J. T., & Ebberson, E. B. (1991) On the form of forgetting. *Psychological Science 2,* 409-415.

Vygotsky, L. S., (1987). *The collected works of L. S. Vygotsky: vol. 1, Problems of general psychology. Including the volume thinking and speech.* (N. Minick, Ed. & Trans.) New York: Plenum.

Zipf, G., (1949) Human Behavior and the Principle of Least Effort. Cambridge, MA: Addison-Wesley.

---

[i] Although there is controversy over whether or not Piaget was actually an individual-constructivist most of the work done in the Piagetian tradition is decidedly focused on individuals.

[ii] Here, we are considering the systems-theoretical, non-individualistic aspect of Piagetian constructivism.

[iii] We would also agree with Chapman's observation that "The kinship between Piaget's theory of equilibration and contemporary theories of self-organizing systems is particularly promising …" (1988, p. 340).