Change, Self-organization and the Search for Causality in Educational Research and Practice

MATTHIJS KOOPMANS
Mercy College (USA)

Causality is an inextricable part of the educational process, as our understanding of what works in education depends on our ability to make causal attributions. Yet, the research and policy literature in education tends to define causality narrowly as the attribution of educational outcomes to intervention effects in a randomized control trial context. This reduction of the educational process to simple input – output relationships leaves important questions unattended about how change is produced in educational systems and how observed results can be predicted based on the propensity toward change in the baseline settings of those systems. This paper considers these questions from a Complex Dynamical Systems perspective, and concludes that answers to them can qualify the findings from at least some experiments.

Our understanding of what works and what does not work in education relies on our ability to make causal attributions. Parents, teachers, school administrators and policy makers use such attributions all the time to determine the impact of their efforts. The dominant methodological paradigm in educational research to establish cause-effect links empirically is the randomized control trial design (RCT), including related designs such as quasi-experimental ones. Results from RCT studies generally carry greater weight in the educational policy arena than causal explanations emanating from other research designs. In a highly influential report, the National Research Council (NRC, 2002) recommends an expansion of RCT designs in education to remove some of the ambiguity of quasi-experimental results, the design that predominated at the time, and research funding priorities have shifted accordingly.

The NRC report acknowledges that RCT studies are particularly well suited when the causal hypotheses are relatively simple, and when group means can be used as an

Complicity: An International Journal of Complexity and Education
Volume 11 (2014), Number 1 • pp. 20-39

expression of the causal processes of interest. A randomized assignment of teachers to incentive pay vs. control conditions to assess whether incentive pay results in higher teacher motivation (Yuan *et al*., 2013) is a good example of such a study. The clearly defined difference between experimental and control conditions makes RCT designs optimally suitable to investigate the causal relationship between payment and motivation.

The NRC report makes a case for methodological diversity in education, however, and notes the healthy tension it creates between explanatory paradigms. Raudenbush (2005) acknowledges the need for this diversity in educational policy research, but makes the following qualification:

> Questions drive methodological choices, and randomized experiments provide the clearest answers to causal questions arising in social science. (Raudenbush, 2005, p. 25)

There is by no means consensus in the literature that a rigorous empirical confirmation of causal and effect relationships necessarily requires RCT, or related designs (Maxwell, 2004; 2012), and different theoretical orientations may inspire different kinds of questions about causality in education. Maxwell (2004) particularly takes issue with the fact that the NRC report favors RCT and its focus on the measurement of the relationships between variables describing input and output conditions, rather than the processes underlying those conditions and relationships. RCT enables the investigator to infer causality from observed regularities in the relationship between variables, but it does not require a description of the underlying causal processes, nor are those processes necessarily specified in the research questions RCT seeks to answer. While qualitative research is equipped to examine those processes, it is not explicitly informed by dynamical theories about the processes that constitute cause and effect. Complex Dynamical Systems (CDS), also referred to in the literature as Nonlinear Dynamical Systems or Complex Adaptive Systems paradigms, offers an alternative conceptualization of causality, which focuses on how the interaction between elements of a system generates cause and effect relationships in the behavior of that system (Doll Jr., 2012; Sulis, 2012). With respect to causality in education, the status of CDS vis-à-vis RCT has not been discussed in great detail in the research literature. The purpose of this paper is to initiate a remedy to this situation by comparing the linear model of causality that underlies RCTs to the causality model that typically informs work taking place from a CDS perspective. The paper first reviews some of the original foundations of dynamical thinking in education (Piaget and Vygotsky), and then outlines four principles of contemporary dynamical thinking, followed by a discussion of three broad questions whose answers will help to explain how dynamical processes contribute to the relationship between educational interventions and outcomes.

Linear and Recursive Approaches to Causality

RCT and CDS approaches represent differing perspectives on what constitutes causality. RCTs are designed to confirm a *linear* or directional causal relationship. This perspective organizes causes and effects in a sequential manner, with causes preceding effects (Pearl, 2009; Smith, 2011). An example of linear causality is the relationship between smoking and lung cancer. CDS approaches tend to focus on the description of *recursive* causal relationships. This perspective organizes causes and effects as a feedback loop, where cause and effect constantly feed into each other in an ongoing interrelationship (Pearl, 2009; Strotz & Wold, 1960). An example of recursive causality is the relationship between the size of predator and their prey populations in a given habitat (Zeeman, 1976). The two perspectives are not mutually exclusive, of course, but they articulate a different angle of incidence to causality. The section that follows discusses in greater detail the design implications from these divergent perspectives for educational research.

Linear Causality in Experimental Designs

Smith (2011) identifies two requirements for the establishment of linear causality: a stable association between variables, and the successful elimination of selection factors. A stable association can be said to exist between two variables X and Y if an ordered relationship between the two is robust, and not negated by the full range of intervening variables Z. Determining whether a relationship meets this criterion typically involves establishing that the relationship is sufficiently strong, and can be reproduced in independently conducted studies. In addition, outcomes should show a predictable relationship with dosage. Also, to establish a causal connection between interventions and outcomes, the study should provoke changes in a natural setting, with outcomes that are behaving appropriately if potential cause is applied, removed and reinstated. Furthermore, the findings should be consistent with subject matter knowledge and be predicted by a theory (Smith, 2011).

To establish causality according to this model, selection factors need to be ruled out in the comparison of the outcomes of different treatment regimes. Strictly, causality requires a comparison between the outcomes for a subject participating in an intervention to those of that same subject had he or she not participated in the intervention (i.e., the counterfactual). Since a person cannot be in two treatment conditions at the same time, a comparison is made to a similar person in a different treatment condition (Dunn, 2002). To rule out differences between the subjects according to their pre-existing characteristics, or an interaction of those pre-existing characteristics with treatment, a random assignment of a sufficiently large number of subjects to treatment conditions is required to ensure a random distribution of these pre-existing characteristics across the treatment groups, which will allow us to attribute differences in outcomes to exposure to different treatment conditions.

In experimental research in education, typically, clusters rather than individuals are randomly assigned to treatment or control conditions to avoid contamination within schools, classrooms, or communities. For instance, the implementation of new initiatives in some classrooms in a given school may affect how teachers in that same school who are not participating in the initiative deliver their instruction, due to their exchanges with participating colleagues. A cluster-based approach would minimize this effect by assigning schools rather than classrooms to treatment conditions, allowing one to statistically model the commonalities in school and classroom experiences, thereby enhancing the reliability of the estimation of treatment effects (Bryk & Raudenbush, 1992; Murray, 1998). A causal relationship between treatment and outcomes can be established if no selection factors at the student or school level compromise the comparison between treatment groups.

However, to effectively describe the effect of changes provoked by RCT designs on measured outcomes, we need to know more about how the changes manifested themselves in the course of the implementation of the intervention, and understand their impact on the systemic behavior at the district, school, or classroom level. By design, RCT limits itself to the input and the output parts of this process without requiring a detailed analysis of the processes that mediate the transitions that ultimately result in certain treatment outcomes. Those processes may or may not respond to the external adaptive demands imposed by the RCT setting due to factors unrelated to their treatment status in the research design, so knowledge of the mediating processes at the systemic level is an essential part of understanding causal relationships between interventions and outcomes. Addressing this problem, for which the term 'black box problem' has been coined in the cybernetic literature, requires a reconstruction of the dynamical system based on observed input output relationships. This can be done if a fully determinate relationship is observed between input and output (Ashby, 1957; Wiener, 1961). However, determinacy is typically absent in the world of education, a state of affairs that creates the need for three additional types of information to establish cause and effect: a characterization of the distribution of outcomes in terms of central tendency and variability to test the validity of the expectations regarding the output, random assignment including a sufficient number of analytical units (students, classrooms, schools) to permit exclusion of selection factors in the analysis, as well as an understanding of the self-regulating behavior in the system producing these outcomes to help us understand the interdependency of input and output.

Recursive Causality in CDS

CDS argues that systems are shaped by the behavior of their constituent components (students, teachers, administrators, other stakeholders), while the behavior of those components is, in turn, delineated by the behavior at the higher systemic level (Koopmans, 1998; 2009). This feedback relationship between higher-level systems and their component parts constitutes *recursive causality*, as the behavior of the system is causally attributed to that of its constituents, while the behavior of those constituents can be attributed to the systemic affiliation that is expressed through their interactive behavior. Thus, recursive causality describes how the systems perpetuate themselves through social interaction, i.e., the relationships that enable communicative exchanges to take place between members are the expression of the mechanisms through which systems maintain and modify themselves (Kampis, 1991). One consequence of this approach is that CDS researchers include different levels of description of the system's unit of analysis and discuss them interchangeably.

While traditional logical-positivistic designs focus on the establishment of causal relationships at the research design level, CDS takes a primary interest in the question of how changes occur at the systemic level, and the dynamical underpinnings of this change process, as stated by the theory (e.g., positive feedback, CUSP catastrophes, transitions from chaos to order, self-organized criticality, etc.). The CDS approach acknowledges the hierarchical nature of the context in which educational outcomes are produced, but its primary interest is in the analysis of the process through which hierarchical structures (classrooms, schools districts) are established and perpetuated, rather than the static classification of hierarchical components.

The measurement of systemic changes provoked by the researcher requires a definition of the system, including its agents, the implementation conditions, and the outcome variables that have been identified for this endeavor (e.g., learning outcomes as a function of the instructional process in a specific classroom). Questions about the propensity of change within those systems at the time that the RCT is undertaken, as well as questions about system stability are consequential to the outcome of the RCT because they help answer the question of why interventions produce changes in given treatment conditions, or why they do not. The absence of a treatment effect may in fact be attributable to intransigence in the system rather than the intrinsic merits of the intervention. In typical RCT reports, the absence of information about the degree of stability in the system complicates the interpretation of the findings.

The Time Factor

There is an implicit time dimension to cause and effect relationships inferred from traditional RCT findings: interventions need to occur before their effects can be observed. This orders implementation and outcome measurement events in a sequential manner, but, without further specification, it places predictor – outcome relationships out over a time-scale in an arbitrary manner, rather than explicitly modeling the effect of time on those relationships (van Geert, 2009). While this arbitrariness with respect to the time variable does not preclude causal attributions about input – output relationships, it obfuscates the feedback loops through which intervention variables and outcome variables mutually and continuously affect each other. Such feedback loops play out in the classroom, for example, where teachers implement instructional strategies to which students respond on an ongoing basis while the teachers adjust and fine-tune their strategy based on those responses.

Modeling the time factor is also important to account for the possibility that outcomes that may be due to interventions do not necessarily occur within the time frame of the RCT study. If the interplay between equilibrating and dis-equilibrating forces in the system result in a slower pace of change, or a later onset, the observed result may fall outside of this time window, leading to a confirmation of the null hypothesis of no effect, while in fact the impact of the intervention may occur later. While RCT do not preclude use of wider time windows to examine treatment effects, the longer term is typically not considered unless the research questions that guide the experiment are explicitly concerned with long term effects (see e.g., Hoxby & Turner, 2013; Olds, Henderson, & Kitzman, 1994; Olds, *et al.*, 1997). On the other hand, a CDS perspective would take an interest in the description of delayed impact scenarios to find out whether treatment effects are sustained over the long run in the interplay between forces of stability and change, and to determine whether or not changes produced by interventions are reversible or not.

The analysis of equilibrium and disequilibrium in systems requires a higher level of resolution than is offered in traditional longitudinal designs (e.g., Singer & Willett, 2003), because there is typically (but not necessarily) an insufficient number of measurement occasions in these designs, as well as insufficient focus on the cyclical patterns over time that would describe the dynamics of the system, and on the increased within-subject variability that CDS predicts would generate systemic transformation. [1] An example of a detailed assessment of the dynamical processes in a system is the Bassano and van Geert (2007) study, which reports on detailed recordings of sentence production in two French-speaking children during their second and third years. The study shows how increased variability in the frequency that the sentences of a given complexity are observed marks critical developmental transitions toward the use of more complex sentences. This finding illustrates an important dynamical principle, namely that increased variability often accompanies qualitative transformation (see below for further discussion). This study also illustrates that observing changes over a long and detailed time scale yields understanding of the feedback relationship between contingencies and behavioral responses, as well as confirming the mechanism that CDS posits regarding the nature of the transformations in those relationships. However, change processes of this kind are taken for granted when between-group and rudimentary within-group comparisons are used to merely infer them, a neglect of what Dewey (1929) calls "the factor most important in education, namely the longitudinal, the temporal span of growth and change" (pp. 67-68).

Archetypical Dynamical Structures in Education

While the CDS approach is underutilized as a research paradigm in education, there is a long tradition of dynamical thinking at the theoretical level. Two important theoretical foundations to the dynamical description of educational processes are Vygotsky's Zone of Proximal Development, and Piaget's concept of discontinuous learning. Vygotsky defines the Zone of Proximal Development as "…the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (Vygotsky, 1978, p. 86). Resolving the discrepancy between what the learner is capable of doing on his or her own and under

[1] Rigorous qualitative research can also achieve the level of resolution that would be required to study dynamical processes. A well-known example is Brown's (1973) detailed study of the language acquisition process in three pre-school children over multi-year periods. See also the Laidlaw et al, (2013) study discussed in further detail below, which is explicitly concerned with dynamical principles.

guidance of a more capable peer or adult becomes a central feature of pedagogical interactions as is the effective use of these discrepancies in learning situations to make such interactions productive.

Piaget's (1967) acknowledges a discontinuity of the learning process, with or without stage transitions, and describes the mechanisms underlying this discontinuity (growth in brain size, disequilibrium in the adaptive process followed by accommodation or reflective abstraction). Piaget's work is a prime example of how dynamical principles can be used to construct a theory. While the empirical confirmation of this theory, for instance of the timing of the stage transitions, does not necessarily require nonlinear dynamical systems methodology, the theory does explicitly articulate time as a contributing factor to the change process, as well as the process of self-organization that underlies cognitive development. Piaget's genetic epistemological position on causality (Piaget, 1971) is essentially an extension to his developmental theory in the sense that the abstract representation of causal relations not only drives the intellectual development and adaptive behavior of children, but also our scientific understanding of our surroundings as adults (see also Doll Jr., 2008).

A CDS perspective describes behavioral and educational outcomes as part of a process of systemic self-organization of individuals adapting to their environments, as well as the self-organization in the instructional system of which both students and teachers are part. The CDS literature is rich with examples that empirically address these archetypical dynamical processes as fundamental building blocks of the teaching and learning process. Van der Maas and Molenaar (1992), use catastrophe theory to present a measurement model for Piagetian stage transitions; Steenbeek and van Geert (2013) operationalize the Zone of Proximal Development in contemporary nonlinear dynamical terms and describe observed learning trajectories as they materialize in that interactive space. In the context of outcomes measurement, this research is productive because it describes the underlying learning processes and the conduciveness of the interactive context to these processes.

When evaluating the impact of interventions on teaching and learning behavior, discontinuity takes students' achievement or comprehension beyond its initial settings, producing a qualitative transformation that is not reversible. Piaget's stage-wise change in the development of children is but one example of such irreversible change. At the district and school building level, institutional internalization of intervention characteristics in a sustainable manner constitutes a transformation in educational systems. According to CDS, the irreversibility of observed changes needs to be determined to convincingly demonstrate program impact.

When making causal attributions about systemic change, CDS makes a distinction between the endogenous and exogenous variables characterizing the system. Endogenous variables are those that are internal to the system, exogenous variables refer to the influences that are external to the system, i.e., the impact of the environment on the system as defined. In a study of the effect of a given intervention on educational outcomes such as achievement and persistence, the time of onset of the intervention, as well as the implementation characteristics (e.g., professional development, use of incentive pay for teachers) are examples of exogenous variables, while the ongoing adaptive behavior of the system, and the parameters we use to describe, it are endogenous. Chaos theory (Prigogine & Stengers, 1984; Sprott, 2003), for example, takes an interest in the analysis of endogenous variability in the system to detect processes of spontaneous transformation, such as, for instance, sensitive dependence on initial conditions, i.e., the amplification of very small changes through which turbulence is produced in the system later on.

CDS challenges us to effectively describe the endogenous process and understand the system's predisposition toward change. The traditional understanding of self-organization in systems is the notion that systems are in principle in a state of equilibrium with occasional periods of turbulence that accompany qualitative change in the system (see e.g., Lewin, 1947). In this context, Watzlawick, Weakland and Fish (1974) distinguish first and second order change. First order change refers to gradual transformations that do not challenge the systemic configurations, whereas second order change refers to qualitative transformations that occur when a system reorganizes itself. This traditional understanding implicitly assumes that systems are, in principle, in a state of equilibrium, unless they are in a period of transformation from one state to another, in which case their behavior will display turbulence. Jörg (2011) effectively summarizes this understanding of how systems work by defining the three basic requirements of any analysis of self-organizing behavior in human social systems as 1) the study the social interaction between individuals, 2) re-thinking causality as being part of a recursive self-organizing process, and 3) the unit of study as being defined in terms of systemic constellations rather than individuals in isolation. The interrelationship between individuals as constituents of those systems is important because the dynamics of self-organization and the dynamics of the behavior of the individuals within those systems are inextricably connected (Koopmans, 1998; 2009). A focus in research on classroom interaction to understand learning behavior is an example of this approach, which describes learning in terms of recurring teacher-student and student-student interactions, rather than as outcomes attributed to the presence or absence of given characteristics of teacher-student relationships.

The shift from focusing on individual learning outcomes to a contextual view of the learning process has also been articulated in the qualitative research literature (see e.g., Candela, Rockwell & Coll, 2004; Fasse & Kolodner, 2000; Maxwell, 2004; 2012), which appreciates the need to understand behavior in context, as well as the hierarchical nature of that context (e.g., classrooms embedded in schools embedded in districts embedded in state and federal policy). Qualitative research also shares with CDS a focus on the *when* and *how* rather than *whether* and *to what extent*, when it comes to the assessment of learning and instruction and its outcomes (Maxwell, 2012). Although there is a historical affinity between qualitative research and the analysis of systemic change (e.g., Bateson, 1935/1972), the qualitative research literature in education is not typically informed by the CDS terminological framework, nor has it focused on the discovery of theoretically formulated dynamical principles in the educational process emanating from that framework. However, recent work also shows that qualitative research can be productively used to explore the contextual mechanisms through which novelty occurs in the education setting. Based on a two-year study, Laidlaw, Makovichuck & Wong (2013) show the transformative capability of digital technology (iPads, You-tube) in the preschool years at home and at school due to the unpredictable nature of the interactions of young children with these virtual information sources.

The description of the time dimension over which observed intervention – outcome relationships play out has been identified as a primary concern to dynamical approaches to behavior in the early cybernetic literature (Ashby, 1957; Wiener, 1961). Fasse and Kolodner (2000) usefully distinguish discrete time and continuous time in their discussion of the importance of the time dimension to qualitative research. In the educational literature, the quantitative measurement of outcomes over a continuous time scale has received very little attention, in spite of the fact that the potential of this approach has been appreciated some time ago (Glass, 1972) and some empirical work (Koopmans, 2011; 2013) notwithstanding. To elucidate the dynamical principles underlying the educational process, modeling continuous time is an important aspect of the discovery process because it allows for the distinction between the gradual and qualitative aspects of transformation, a central concern to CDS scholarship, which cannot be effectively addressed when systemic behavior is analyzed based on a comparison of processes or outcomes taking place within crudely sampled chunks of time. The time aspect is largely ignored in the study of causality in education, except for the behavior modification literature (see e.g., Hall *et al.*, 1971, and many others), where the impact of an intervention is measured on a time scale, after a long string of baseline measurements is conducted. The group comparisons on which causal attribution in RCT studies is based does not concern itself at the design level with the dynamical processes outlined above, leaving those open to investigation by others. The section below describes some of the processes that would be of primary interest in such research.

Principles of Self-organization and Change

CDS has a rich vocabulary to describe processes of the systemic self-preservation and change (Koopmans, 2009), but several basic principles characterize most of the scenarios described in the literature. The principles described below are the ones that potentially have the greatest relevance to education. These scenarios have in common that they all that take the hierarchical character of educational systems into account, i.e., the interaction between higher-level systems such as districts, schools and classrooms, and their lower-level components (teachers, administrators and students).

Principle 1: Systems acquire and maintain their integrity through self-organization.

Self-organization can be defined as a process through which systems create and maintain order through interaction between constituent components. This general formulation of the self-organization process distinguishes "living" systems from closed mechanical ones (Varela, Maturana & Uribe, 1974), and it presupposes an adaptive process through which systems are capable of maintaining their integrity in the face of exogenous processes impacting the system, and modify themselves in response to contingencies (Kampis, 1991). As noted above, CDS postulates an ongoing hierarchical interrelationship in the communication between actors and the systems that they are part of through which systems regulate themselves internally as they negotiate external demands.

The distinction between systems coming into being and systems perpetuating themselves is immaterial in this context, as self-perpetuation through interaction is an ongoing formative process (Varela, Maturana & Uribe, 1974). Thus, systems perpetuate their integrity through the interactive process (Solé & Bascompte, 2011), and through this process, they maintain their distinctiveness. Furthermore, systems are capable of remaining distinct as a coherent organization in the face of turbulence. (*Cf* autopoiesis; Mingers, 1995; Varela, 1979; Varela, Maturana & Uribe, 1974). Given this broad definition of self-organization, schools and classrooms are examples of such systems.

The dynamical understanding of systems as outlined above contrasts the CDS from a more static approach that depicts systems in terms of a hierarchical classification of their components, as is done in cluster-randomized experiments. The implication of this dynamical point of view to the concept to causality in education is the fact that transformations in the system are understood in terms of changes in this self-organizing process. Thus, changes in measured educational outcomes are related back to this process. For example, if implementation of new programs leads to improved instruction, the interaction between teachers and students mediate the relationship between program implementation and the desired outcomes. Therefore, these interaction processes are a critical aspect of the causal attribution of outcomes to implementation in educational setting. As an example of this approach, Pennings, *et al.* (in press) empirically investigated how the analysis of real-time teacher interaction with their individual students helps us understand how students and teachers co-construct the relational patterns that uniquely define their classroom systems.

Principle 2: Change is accompanied by perturbation in the system.

Catastrophe theory and of chaos theory are two of the most highly influential conceptual frameworks informing current thought in CDS. Catastrophe theory (Zeeman, 1976) posits that systemic change is accompanied by turbulence as fluctuations in the behavioral expressions of the system approach a mathematically unacceptable region, resulting either in a retreat of the system to its original state, or a reorganization into a novel state (second order change). Second order change is accompanied by turbulent patterns in the outcome variable due to the proximity of the predictors to this unacceptable region. Hysteresis refers to the delay effect that precedes second order change, reflecting an initial tendency in the system to retain its original state when the determinants of its behavior are shifting. When considering whether interventions transform the educational system, this delay effect may account for the possibility that hypothesized changes fall outside of the time window of an RCT study, because even if change is imminent, the behavioral determinants may still be close to the unacceptable region. A valid interpretation of the observed findings would require knowledge of the behavior of these parameters. Conversely, a gradual change in the outcomes may lead one to conclude an effect, whereas in fact there is none, reflecting a lack of transformation at the systemic level, as indicated by the absence of change in the underlying parameters, which is to say there is change without reform (i.e., first order change).

In their study of developmental stage transitions, van der Maas and Molenaar (1992) found that one empirical indicator of hysteresis in a system is a bimodal distribution of learning outcomes, which may suggest, for instance, that some learners are still in the original state of the system while others have already moved to the new state of the system. A bimodal distribution has also been detected in the achievement of high school students on a science test suggesting such a transformation in the learning processes measured through these outcomes (Stamovlasis, 2006; Stamovlasis & Tsaparlis, 2012).

The notion of perturbation as an accompaniment of transformation is also central to chaos theory. Prigogine & Stengers (1984) describe how perturbation may retreat or spread in a system depending on whether the size of the region affected by the initial perturbation has reached a critical value (nucleation threshold). Past that critical value, second order change is predicted to occur. For example, it could be that instruction in the classroom has more of an impact on the level of discourse in a classroom as a whole once a certain number of students in that class have internalized a higher level of understanding of the material, as a result of which the level of discourse gets lifted to bring the other students along. The identification of those threshold values and the question whether such specific threshold values generalize across classrooms, are empirical issues to be addressed in future research. However, a retroactive determination that change has taken place does not by itself reveal anything with regards to the process of self-organization that is responsible for the changes. At most, the process that produced the outcomes can be hypothesized.

Principle 3: Novelty is a product of the interaction between higher level systems and their constituent members.

Ever since the introduction of the feedback loop into our understanding of how social interaction and systemic self-regulation are connected (e.g., Bateson, 1967/1972), systemic change is seen as requiring an interaction between the higher level systems and their individual component parts. Families and organizations have been described as such systems (Bütz, Chamberlain & McCown, 1997; Goldstein, 1994; McKelvey, 2004; Koopmans, 1998). Classrooms and school districts can be described as such systems as well, which is to say that the structure of these systems is maintained through the social interaction of constituent members, amongst themselves as well as with players that are external to the system. To find the origins of novelty in instruction and learning, both bottom-up and top-down causality in classroom behavior need to be acknowledged. Research on emergence processes tends to focus on the bottom-up part, i.e., how shifts in the behavior of individual members of the system produce change in the higher level system of which they are part (e.g., McKelvey, 2004). The emergence of higher levels of understanding of lesson material as a result of teacher – student and student – student interactions is an example of such novelty. We need to know how the system (classroom) self-organizes prior to the emergence of novel behavior patterns to understand the system's predisposition toward transformation (Goldstein, 2002; Holland, 1998). Knowing this propensity facilitates interpretation of RCT results, as the absence of change may reflect a low propensity toward transformation and a high propensity may result in changes that are not sustained over the long term because of subsequent transformations.

Principle 4: The propensity toward change is a crucial process to consider in the determination of cause and effect relationships in education.

Many systems scholars have dropped the assumption that systems are stable in principle, unless they are in transformation from one state to another. Instead, it is argued that many systems are probably in a state that can be better described as a far-from-equilibrium (Goldstein, 1988), where ongoing turbulence creates continuous opportunities for transformation. This shift in our thinking raises the question how instabilities in the baseline condition of the system predispose it toward transformation.

In his theory of self-organized criticality, Bak (1996) postulates that a cumulative effect of the input conditions in a system can cause a system to evolve into a critical state, where minor disturbances result in higher-level transformations. The textbook example of this process is the sand pile model. When pouring sand on a flat surface, an accumulation of sand will occur with occasional avalanches at critical moments. As the continued supply of sand to the pile produces friction between grains, the avalanche reorganizes the pile reducing that tension. The state of the sand pile at the point where the avalanche occurs is defined as self-organized criticality, the moment when reorganization is imminent (Bak, 1996).[2]

The concept of self-organized criticality is important for education because it describes a particular type of transformative process where continued input to the system drives a transformative process. Vygotsky's zone of proximal development is an example of self-organized criticality in the learning process, where the 'critical state' of the system is the potential of learners to surpass their inherent capabilities as they benefit from the prolonged guidance of others.

Another aspect in our understanding of baseline conditions in the system is the determination of whether the system displays sensitivity to initial conditions (Sprott, 2003), an indicator that the system is chaotic. Sensitivity to initial conditions is said to occur if tiny shifting patterns at a given point in time may result in major turbulence and

[2] In spite of its continued usefulness as a heuristic metaphor, it should be noted that replications of the sand pile experiment failed to confirm self-organized criticality as Bak defined it, because the texture of sand is too dense relative to the friction between the grains that would produce the criticality. It turns out that experiments with rice piles are more successful in this regard, provided that elongated grains are used (Jensen, 1998).

unpredictability later on. The importance of this notion to the description of any system is the fact that it is part of the endogenous process, and in that sense can be seen as a 'spontaneous transformation' that does not require an impact from forces that are external to that system. The Brazilian butterfly that flaps its wings to produce a hurricane later on is an archetypical (albeit empirically unconfirmed) example of this process.

Establishing sensitivity to initial conditions in a system is important because it describes its capability for drastic transformation. Koopmans (2013) challenged the assumption of high school attendance as a stable pattern by showing sensitivity to initial conditions in the daily attendance rates in one urban high school over a seven year period. The importance of examining sensitivity to initial conditions is that it establishes the propensity toward change in the baseline conditions. Variability in this propensity from one system to the next may complicate the interpretation of experimental findings, because the RCT design does not require an estimation of its contribution to those results. However, the generalization of RCT findings requires constancy among systems in this regard.

The scenarios discussed above have several important features in common. They describe change as a qualitative rather than a gradual process. Change is defined as a process through which a system transforms from one state to a qualitatively distinct other state. Such transformation is hard to predict, and when it occurs, it is typically, but not necessarily accompanied by turbulence and high levels of variability in the data. Of more immediate relevance to the question of establishing causality in education is the distinction between endogenous or exogenous processes: change can occur as part of the normal behavior of systems that are not stable, or it can occur as an adaptive response to influences that are external to the system, such as interventions into educational systems. The causal attribution of systemic change as part of the endogenous process requires a description of the self-organizing behavior of the system because the causal agent lies in the internal dynamics of the system rather than in an external provocation. Similarly, in the case of transformations as an adaptive response to exogenous influences, the parameters that describe the self-organization in the system are instrumental in determining whether or not the outside impact causes transformation in the system. Therefore, the relationship between interventions and outcomes is not straightforward and requires a description of the self-organizing processes through which second order change is produced. The attribution of change to interventions and the study of change as a self-organizing process are different but complimentary endeavors in the search for causality in education.

Three Questions for the Field

Research based on a comparison of group outcomes, such as RCT, offers an incomplete understanding of causality because the investigation of the process of change is not incorporated into the design. An effective triangulation of findings from experimental or quasi-experimental designs with those from qualitative research on the implementation process go a long way toward addressing this shortcoming, and RCT studies are in fact often supplemented with a qualitative research component to describe the implementation story, although the fact that this information is collected does not necessarily mean that it is also utilized to inform the findings (Newman *et al.*, 2012). CDS takes a specific interest in the applicability of a set of mathematically well-defined scenarios of stability and change to the education and schooling process, which neither RCT nor qualitative research are well-equipped to address. We should take an interest in those scenarios because our causal attributions of program effects are strengthened if processes of transformation and self-preservation in the system are actually observed in detail rather than retrospectively inferred from observed differences in outcomes between groups exposed to different treatment conditions. In keeping with one of the most basic insights derived from CDS, the principles outlined above all focus on the processes of stability and change. Among the questions that address the implications of CDS for the study of how instructional and policy innovations cause change in educational outcomes are those discussed below.

Is the system in a highly stable or in a far from equilibrium state when the intervention is introduced?

The absence of a treatment effect in an RCT study begs the question whether this is due to the fact that the theory of change behind the treatment is false, or whether there is a lack of propensity in the system to innovation in response to external influences, regardless of the merits of the intervention. If an intervention consistently does not produce effects, we would conclude that the theory of change is false regardless of the propensity for change in the settings to which the RCT results are generalized. Variability in RCT outcomes due to variability in the propensity toward change in the system would qualify the theory of change from setting to setting. It is important to distinguish these two situations.

When addressing cause and effect questions in education, the insights from CDS work can fruitfully infuse quantitative RCT designs, and well as qualitative research efforts to provide context to those findings. The use of the tools and expertise of qualitative research can enable researchers to make detailed observations of the interaction between individual players in the system to describe the process of self-organization within the system, and the propensities toward stability and transformation that can be detected in those exchanges. While qualitative research provides the tools to study these processes in great detail, it does not necessarily impose a conceptual framework on the interpretation of the observed findings, while CDS is particular about how interactive processes in the system generate cause and effect relationships.

CDS can be used to analyze baseline conditions in the systems of interest ('business-as-usual'), as well as the propensity toward transformation that is produced by and reflected in the interactions between the participants in the system. The same methodological principles can be used to study the transformation in response to exogenous processes invoked by RCT participation. Such transformation can include changes in classroom interaction patterns, as well as self-organizing behaviors at higher levels of the educational system. Koopmans (2001) proposes a two-tiered approach to the analysis of interacting social systems, which includes detailed observations of the interactions between members, as well as discussions with participants about how they see their systems as being organized, and about their perception of their influence on that organization through their interaction. While this analysis focused on family systems, it will be productive to explore the applicability of this methodology in educational settings other than the family, and to use it to study responses to interventions.

Are the changes that are inferred from RCT comparisons of an irreversible nature?

CDS makes a distinction between reversible and irreversible change. Concluding a treatment effect from RCT findings does not resolve the question whether the observed changes are of an irreversible nature or not, raising the possibility of a return to the baseline after the study is completed. Answering the reversibility question is important because it helps to decide whether ongoing treatment produces lasting transformation. Determining the sustainability of treatment effects requires a prolonged measurement process following the introduction of treatment to the system. To conclude the irreversibility of the observed changes (Prigogine & Stengers, 1984), it is necessary to distinguish first and second order change processes in the system in response to interventions (Watzlawick, Weakland & Fish, 1974). It could be that a classroom moves from a traditional floor plan with students in rows facing the teacher, to roundtables to facilitate collaborative learning processes as part of a new initiative. This qualitative transformation may support a sustaining effect of the condition, reducing the likelihood of a return to a baseline where collaborative learning activities are gradually replaced by traditional pedagogy.

How is novelty introduced in the interaction between stakeholders and the larger systemic constellations of which they are part?

Both linear and non-linear approaches to causality acknowledge the hierarchical structure of educational systems. In the experimental context, linear designs have been expanded to include cluster-based rather than individual-based random assignments, such as those by school and by classroom. In dynamical models, the interrelationship between individuals and higher systemic levels involves an assessment of how the social interaction between individual agents is constrained by the systemic context, and how that same interactive process is capable of producing higher-level systemic transformation. The interaction between individuals and the higher level systems that they are part of is of particular relevance in the emergence literature, which seeks to understand novelty in systemic processes in terms of bottom-up processes in which innovative features in the interaction process between individual components and higher level systemic components produce novelty in the self-organization of the higher level system. It is in keeping with this bottom up approach to seek for the origins of transformation in the interaction between particular individuals.

Emergence as a bottom-up process is complemented by a top-down process in which individuals are reacting to systemic constraints in their behavior, a process called supervenience (Sawyer, 2005). While these constraints may delineate the behavior of individuals within their systems, the agents are always individuals. The interplay between the behavior of individuals and the educational systems they maintain deserves empirical consideration. The school reform literature is not short of examples of how educational systems can be caught between these two forces (e.g., Cuban, 2010; Elmore, 2004; Hess, 2010), but there is a shortage of empirical work that specifically looks at how the dynamic exchanges between individuals within educational systems produces or fails to produce transformation. Holland (1998) posits that systemic redundancies inhibit innovation, and that therefore, propensity toward transformation would be heightened if those redundancies were reduced or eliminated. For example, if curriculum innovations are introduced to supplement existing curricular features, both old and new patterns operate in the system, which, according to this point of view, would in effect reduce the system's propensity toward change rather than enhancing it because of the lack of room within the system to accommodate the new elements (Goldstein, 2002). The utility of this model to help address the widely perceived difficulty to introduce novelty into educational systems (Hess, 2010) would be a productive endeavor. Given the paucity of empirical research on the organizational features of education systems, it would be instructive as well in this context to consider the impact of the dynamical literature concerned with self-organizing processes in for-profit environments (e.g., Allen, Maguire, & McKelvey, 2011), on educational environments.

Conclusion

The CDS approach shifts our focus from questions about what works to questions about the transformative process underlying changes for the better. As a meta-theoretical perspective, it inspires formulation of a new set of priorities by which to assess the validity of findings from traditional RCT studies. These questions include 1) a determination of the stability of the system in which change is attempted to be made through treatment, 2) specification of the systemic effects of the measured RCT outcomes as reversible or irreversible, 3) addressing the possibility that changes in outcomes fall outside of the time window of the RCT study, 4) identifying the point of onset from which systemic change emerges. A review of the growing corpus of RCT research in education in light of considerations such as these would delineate more clearly the role that complex dynamical systems approaches can play to enhance our understanding of causality in education, and further qualify that role in light of different types of experimental results.

There also is a need for further theoretical work within CDS on the question of whether educational systems are necessarily self-organizing. As is the case with larger societal systems (Mingers, 1995), the applicability of constructs such as self-organization and autopoiesis to educational systems is not without detractors. Further discussion of this controversy is worthwhile, but beyond the scope of this paper.

References

Allen, P., Maguire, S. & McKelvey, B. (2011). *The SAGE handbook of complexity and management.* Newbury Park, CA: Sage.

Ashby, W. R. (1957). *An introduction to cybernetics*. London: Chapman & Hall.

Bak, P. (1996). *How nature works: The science of self-organized criticality*. New York: Springer.

Bassano, D. & van Geert, P. (2007). Modeling continuity and discontinuity in utterance length: A quantitative approach to changes, transitions and intra-individual variability in early grammatical development. *Developmental Science*, *10*, 588-612.

Bateson, G. (1935/1972). Culture contact and schismogenesis. In G. Bateson, *Steps toward an ecology of mind: A revolutionary approach to man's understanding of himself*. (pp. 61-71). [Reprinted from *Man*, *XXXV*, Article 199]

Bateson, G. (1967/1972) Cybernetic explanation. In Bateson, *Steps to an ecology of mind,* (pp.399-410). New York: Ballantine. [Reprinted from *American Behavioral Scientist*, *8*, 29-32]

Brown, R. W. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Bryk, A. S. & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.

Bütz, M. R., Chamberlain, L. L., & McCown, W. G. (1997). *Strange attractors: Chaos, complexity and the art of family therapy.* London: Wiley & Sons.

Candela, A. Rockwell, E. & Coll, C. (2004). What in the world happens in classrooms: Qualitative classroom research? *European Educational Research Journal*, *3*, 692-711.

Cuban, L. (2010). *As good as it gets: What school reform brought to Austin*. Cambridge, MA: Harvard University Press.

Dewey, J. (1929). *The Sources of a science in education*. New York: Liveright.

Doll Jr., W. E. (2008). Response to Proulx: Maturana is not Constructivist…..Nor is Piaget. *Complicity: An International Journal of Complexity and Education, 5,* 27-31.

Doll Jr., W. E. (2012). Complexity and the culture of curriculum. *Complicity: An International Journal of Complexity and Education, 8*, 10-29.

Dunn, G. (2002). The challenge of patient choice and non-adherence to treatment in randomized controlled trials of counseling or psychotherapy. *Understanding Statistics*, *1*, 19-30.

Elmore, R. F. (2004). School reform from the inside out: Policy, practice and performance. Cambridge, MA: Harvard University Press.

Fasse, B. & Kolodner, J. L. (2000). Evaluating classroom practices using qualitative research methods: Defining and refining the process. In B. Fishman & S. O'Connor-Divelbiss (Eds.) Fourth International Conference of the Learning Sciences. Mahwah: NJ: Erlbaum. (pp. 193-198)

Glass, G. V. (1972). Estimating the effects of intervention into a non-stationary time-series. American Educational Research Journal, 9, 463-477.

Goldstein, J. (1988). A far-from-equilibrium systems approach to resistance to change. *Organizational Dynamics*, *17*, 16-26.

Goldstein, J. (1994). *The unshackled organization: Facing the challenge of unpredictability through spontaneous reorganization*. Portland, OR: Productivity Press.

Goldstein, J. (2002). The singular nature of emergent levels: Suggestions for a theory of emergence. *Nonlinear Dynamics, Psychology and Life Sciences, 6*, 293-309.

Hall, R. V., Fox, R., Willard, D., Goldsmith, L, Emerson, M., Owen, M., Davis, F., & Porcia, E. (1971). The teacher as observer and experimenter in the modification of disputing and talking-out behaviors. *Journal of Applied Behavior Analysis, 4,* 141-149.

Hess, F. M. (2010). *The same thing over and over: How school reformers get stuck in yesterday's ideas*. Cambridge, MA: Harvard University Press.

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

Hoxby, C., & Turner, S. (2013). *Expanding college opportunities for high-achieving, low income students*. Stanford, CA: Stanford Institute for Economic Policy Research. Retrieved from http://siepr.stanford.edu/publicationsprofile/2555 .

Jensen, H. J. (1998). *Self-organized criticality: Emergent complex behavior in physical and biological systems*. Cambridge: Cambridge University Press.

Jörg, T. (2011). *New thinking in complexity for the social sciences and humanities: A generative, transdisciplinary approach*. New York: Springer.

Kampis, G. (1991). *Self-modifying systems in biology and cognitive science: A new framework for dynamics, information and complexity*. New York: Pergamon Press.

Koopmans, M. (1998). Chaos theory and the problem of change in family systems. *Nonlinear Dynamics in Psychology and Life Sciences, 2,* 133-148.

Koopmans, M. (2001). From double bind to n-bind: Toward a new theory of schizophrenia and family interaction. *Nonlinear Dynamics, Psychology & Life Sciences*, *5*, 289-323.

Koopmans, M. (2009). Epilogue: Psychology at the edge of chaos. In S. J. Guastello, M. Koopmans, & D. Pincus (Eds.) *Chaos and complexity in psychology: The theory of nonlinear dynamical systems.* New York: NY: Cambridge University Press. Pp. 506-526.

Koopmans, M. (2011, April). *Time series in education: The analysis of daily attendance in two high schools*. Presented at the annual meeting of the American Educational Research Association. New Orleans, LA.

Koopmans, M. (2013, March). *Daily attendance rates in one urban high school: Three analytical viewpoints.* Presented at the 21st Annual Winter Chaos/Snowflake Conference, Springfield, MA. http://impleximundi.com/tiki-read_article.php?articleId=149.

Laidlaw, L., Makovichuk, L, & Wong, S. (2013, April) *Complexity, pedagogy, play: On using technology within emergent learning structures with young learners*. Presented at the annual meeting of the American Educational Research Association. San Francisco, CA.

Lewin, K. (1947). Frontiers in group dynamics. *Human Relations*, *1*, 5-41.

Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher*, *33*, 3-11.

Maxwell, J. A. (2012). The importance of qualitative research for causal explanation in education. *Qualitative Inquiry*, *18*, 655-661.

McKelvey, B. (2004). Complexity science as an order-creation science: New theory, new method. *Emergence, Complexity and Organization*, *6*, 2-27.

Mingers, J. (1995). *Self-producing systems: Implications and applications of autopoiesis*. New York: Plenum.

Murray, D. M. (1998). *Design and analysis of group-randomized trials*. New York: Oxford University Press.

National Research Council (2002). *Scientific research in education*. Committee on Scientific Principles for Education Research. Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Newman, D., Finney, P.B., Bell, S., Turner, H., Jaciw, A.P., Zacamy, J.L., & Feagans Gould, L. (2012). *Evaluation of the effectiveness of the Alabama Math, Science, and Technology Initiative (AMSTI)*. (NCEE 2012–4008). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Olds, D. L. Eckenrode, J., Henderson, Jr. C. R., Kitzman, H., Powers, J., Cole, R., Sidora, K., Morris, P. Pettitt, L. M., & Luckey, D. (1997). Long-term effects of home visitation on maternal life course and child abuse and neglect fifteen-year follow-up of a randomized control trial. *JAMA*, *278*, 637-643.

Olds, D. L., Henderson Jr., C. R., & Kitzman, H. (1994). Does prenatal and infancy nurse home visitation have enduring effects on qualities of parental caregiving and child health at 25 and 50 months. *Pediatrics*, *93*, 89-98.

Pearl, J. (2009). *Causality: Models, reasoning, and inference*. (2nd Edition) New York: Cambridge University Press.

Pennings, H. J. M., Brekelmans, M., Wubbels, T., van der Want, A. C., Claessens, L. C. A., & van Tartwijk, J, (in press). A nonlinear dynamic systems approach to real-time teacher behavior: Differences between teachers. To appear in *Nonlinear Dynamics, Psychology, and Life Sciences*.

Piaget, J. (1967). *Six psychological studies*. New York: Vintage.

Piaget, J. (1971). *Genetic epistemology*. New York: Norton.

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

Raudenbush, S. W. (2005). Learning from attempts to improve schooling: The contribution of methodological diversity. *Educational Researcher*, *34*, 25-31.

Sawyer, R. K. (2005). *Social emergence: Societies as complex systems*. New York: Cambridge University Press.

Singer, J. D. & Willett, J. B. (2003). *Applied longitudinal analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.

Smith, R. B. (2011). *Multilevel modeling of social problems: A causal perspective*. New York: Springer.

Solé, R. V. & Bascompte, J. (2011). *Self-organization in complex ecosystems*. Princeton, NJ: Princeton University Press.

Sprott, J. C. (2003). *Chaos and time series analysis*. New York: Oxford University Press.

Stamovlasis, D. (2006). The nonlinear dynamic hypothesis in science education problem solving: A catastrophe theory approach. *Nonlinear Dynamics, Psychology and Life Sciences, 10,* 37-70.

Stamovlasis, D. & Tsaparlis, G. (2012). Applying catastrophe theory to an information processing model of problem solving in science education. *Science Education*, *96*, 392-410.

Steenbeek, H. & van Geert, P. (2013). The emergence of learning-teaching trajectories in education: A complex dynamic systems approach. *Nonlinear Dynamics, Psychology and Life Sciences*, *17*, 233-267.

Strotz, R. H. & Wold, H. O. A. (1960). Recursive vs. non-recursive systems: An attempt at synthesis. *Econometrica*, *28*, 417-427.

Sulis, W. H. (2012). Causal tapestries for psychology and physics. *Nonlinear Dynamics, Psychology and Life Sciences, 16,* 113-136.

van der Maas, H. L. J. & Molenaar, P. C. (1992). Stagewise cognitive development: An application of catastrophe theory. *Psychological Review*, *99*, 395-417.

van Geert, P. (2009). Nonlinear complex systems in developmental psychology. In S. J. Guastello, M. Koopmans, & D. Pincus (Eds.) *Chaos and complexity in psychology: The theory of nonlinear dynamical systems*. New York: Cambridge University Press. Pp. 242-281.

Varela, F. J. (1979). *Principles of biological autonomy*. Dordrecht: North Holland.

Varela, F. J., Maturana, H. R., & Uribe, R. (1974). Autopoiesis: The organization of living systems, its characterization and a model. *Biosystems*, *5*, 187-196.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Watzlawick, P., Weakland, J., & Fish, R. (1974). *Change: Principles of problem formation and problem resolution*. New York: Norton.

Wiener, N. (1961). *Cybernetics, or control and communication in the animal and the machine* (2nd Edition). Cambridge, MA: MIT Press.

Yuan, K., Vi-Nhuan, L., Marsh, J. A., Hamilton, L. S., Stecher, B. M., & Springer, M. G. (2013). Incentive pay programs do not affect teacher motivation or reported practices: Results from three randomized studies. *Educational Evaluation and Policy Analysis*, *35*, 3-22.

Zeeman, E. C. (1976). Catastrophe theory, *Scientific American*, *234*, 65-83.

Acknowledgments

This paper is based on a roundtable presentation at the annual meeting of the American Educational Research Association, San Francisco, CA; April 27, 2013.

About the Author

Matthijs Koopmans is an Associate Professor at Mercy College. His current research interests include the application of complex dynamical systems approaches to education, cause and effect relationships and nonlinear time series. Correspondence: mkoopmans@mercy.edu.