

Chapter 8

Ergodicity and the Merits of the Single Case

Matthijs Koopmans

Introduction

While dynamical scholarship has played a significant role in the twentieth century in the development of theories describing the underpinnings of development and instruction (Piaget, 1967; Vygotsky, 1978), the impact of recent understandings in dynamical theory, such as fractals, chaos, catastrophe, and complexity, has been relatively modest up to this point. There have been some major attempts to accommodate the new insights from those dynamical models to our existing knowledge. For example, van der Maas and Molenaar (1992) used catastrophe theory to describe the dynamical underpinnings of Piagetian stage transitions, and Stamovlasis and Tsaparris (2012) similarly applied catastrophe models to problem solving in science education. Steenbeek, Jansen and van Geert (2012) used a complexity perspective to describe the real-time scaffolding dynamics in teacher interactions with students with emotional behavioral disorders, and children's play has been conceptualized and described as an emergent developmental process by Fromberg (2010) and Laidlaw, Makovichuk, Wong and O'Mara (2013).

There have also been a number of more broad based discussions about the need to consult models of complexity when studying educational change (Jörg, Davis, & Nickmans, 2007; Koopmans, 2014b; Lemke & Sabelli, 2008). While significant, these developments have not yet clearly positioned educational research from a dynamical systems or complexity perspective as an alternative paradigm to negotiate the relationship between theory and practice in education. In fact, in the policy arena, educational research has evolved more or less in the opposite direction. There has been an increasing reliance on quantitative information in the service of school and teacher accountability models, and on large scale randomized control trial (RCT)

M. Koopmans (✉)

Mercy College, School of Education, Mercy Hall, Rm. SE 31, 555 Broadway,
Dobbs Ferry, NY 10522, USA

e-mail: mkoopmans@mercy.edu

studies to examine the effectiveness of educational interventions and school reform efforts. Especially the use of RCT designs has been put forward as the preferred method for establishing causal links between educational interventions and their outcomes (Raudenbush, 2005; Slavin, 2002), because it removes the ambiguity that affects the inference of such causality in quasi-experimental and traditional correlation and regression designs (Murray, 1998; National Research Council, 2002).

Several investigators have taken issue with this methodological disposition, most notably Maxwell (2004; 2012), who argues from a qualitative research standpoint that we should scrutinize the processes that generate causality, because they help understand *how* certain outcomes were obtained, rather than merely establishing *that* they were obtained, thereby leaving it up to theory to provide an answer to the *how* question, as RCT often does. From the vantage point of the complex dynamical systems (CDS) paradigm, one would concur with Maxwell's critique because the investigation of the dynamical processes underlying cause and effect potentially offers important qualifications to the knowledge obtained through RCT and related designs (Koopmans, 2014a). Because of their strictly inductive nature (Bogdan & Biklen, 1982), qualitative research methodologies are well suited to uncover such dynamical processes, and they have been successfully used to address a wide range of complex processes in education, such as school reform (White & Levin, 2013), literacy (Laidlaw et al., 2013) and leadership development (Combes & Patterson, 2013). Studies such as these illustrate the usefulness of the qualitative perspective addressing complexity questions. The richness of the data makes it possible in these instances to observe and describe the dynamical processes in detail.

Establishing a unidirectional causal relationship between interventions and outcomes addresses only a narrow part of the causal process. It relies on aggregated information (comparison of group averages) without typically providing a detailed description of the evolution of the behavioral changes that constitute an effect or the self-organizing process through which change manifests itself systemically. CDS argues that the behavior of agents in the educational system (teachers, students, administrators, policy makers) needs to be understood in relation to that of other agents in the same systemic context. The isolation of the behavior of individual units that is needed for aggregation of findings discourages consideration of the systemic interactions that define those larger units. Examples of such interactions are teacher–student interactions, student–student interactions, and the interactions between teachers and the principal's office. Information is also needed about recursive processes that tie the behavior of individuals to that of the systemic constellations in which they interact. Causal attribution requires an understanding of the mutual contingency of the behaviors of individuals in relation to those of larger systemic constellations of which those individuals are part (Koopmans, 1998; Sawyer, 2005). Examples are the relatedness of students and their classroom systems, and school administrators within the systemic confinements of their districts and school buildings.

It is important to note that to meet the external validity needs that come with policy research, the large scale data collection and the generation of replicable findings is required to the same extent in complexity research as it is elsewhere, because a high degree of resolution is needed in the data to detect the dynamical

processes of interest. This situation has created a need for rigorous statistical methods that are specific to the discovery and description of complex dynamical processes such as self-organized criticality (Bak, 1996; Jensen, 1998), qualitative transformations (Watzlawick, Weakland, & Fish, 1974); sensitive dependence on initial conditions (Sprott, 2003) and emergence (Goldstein, 1988, Chap. 4). In many academic disciplines, such designs have been fully incorporated by now into the research practice (e.g., Guastello & Gregson, 2011). However, they have been underutilized in education until recently, when a cascade of highly promising work appeared utilizing a wide variety of empirical approaches to uncover the dynamical underpinnings of teacher–learner interaction (Steenbeek et al., 2012), the interaction between the learner and the task (Garner & Russell, 2014), the impact of economic conditions on schooling outcomes (Guevara, López, Posh, & Zúñiga, 2014), the dynamics of collaboration among school administrators (Marion et al., 2012), interpersonal teacher–student dynamics (Pennings et al., 2014), the impact of arousal and motivation on academic achievement (Stamovlasis & Sideridis, 2014), self-similarity in high school attendance (Koopmans, 2015), and the adaptive process of motor learning (Tani et al., 2014). This research illustrates a high level of methodological sophistication in the empirical work done in this area that was virtually absent as recently as a decade ago.

Merits of the Single Case

In a *New York Times* article entitled *Why Doctors Need Stories*, Peter Kramer, the author of *Listening to Prozac* expresses his concern about the disregard of the individual case study as a legitimate source of evidence in medicine (Kramer, 2014). He argues for a revitalization of the case study as a necessary complement to aggregated data that are used in randomized control trial (RCT) studies. In a climate that was still quite hostile to case-based field research (Campbell & Stanley, 1963; Scriven, 1967), Smith and Geoffrey (1968) likewise advocated for the use of such designs in the social sciences to help generate hypotheses to be verified through experimental or correlational studies, and it has often been argued since that using ethnographic research in conjunction with comparative designs can help rule out alternative explanations to RCT-based findings (National Research Council, 2002; Yin, 2000), and strengthen explanations of observed effects that fall outside of the purview of the confirmatory study (National Research Council, 2002).

In education, the study of the particularities of individual cases has traditionally been relegated to qualitative researchers using ethnographic designs to immerse themselves into the system to observe the processes of interest unfold in real time and provide detailed description of the interaction between the various components of the system. The use of ethnography provides “thick descriptions” that may uncover the processes through which transformation takes place in classrooms and school buildings (e.g., Bogdan & Biklen, 1982; Lincoln & Guba, 1985). Ethnographic designs also allow for a triangulation of findings from experimental

designs and other designs with thick descriptions of the implementation story, allowing for a deeper understanding of what makes given interventions effective, and what processes ultimately explain the outcomes of an experiment. Ethnographic designs can also be utilized by themselves to investigate causal processes (Maxwell, 2004, 2012), and examine the way antecedent and consequent events play out over time (Miles & Huberman, 1994). There is a long-standing affinity between ethnography and the dynamical literature going back to Gregory Bateson's (1935/1972) anthropological work, and Kurt Lewin's (1947) description group dynamics in terms of stability and change in the systems, a development that has formed the basis for the infusion of qualitative research with more updated dynamical systems concepts (Bloom, Chap. 3; Bloom & Volk, 2007; Hetherington, 2013; Laidlaw et al., 2013).

Nomothetic Versus Idiographic Perspectives

The literature sometimes describes research in the behavioral sciences as representing either one of two distinct epistemological perspectives, often referred to as *nomothetic* and *idiographic* (Burns, 2000; Burrell, 1979). The nomothetic perspective represents the search for generally applicable principles and regularities in the relationship between variables that can be used to make inferences about the population based on what is observed in a sample. It is therefore often equated with quantitative research. In education, conventional RCT designs and quasi-experimental designs would fall under the nomothetic approach. This method derives its utility from the representativeness of the sample to the population. However, information about the particularities of the individual cases gets lost in the aggregation that is required to estimate population characteristics. Hence, there is a trade-off between the rigor derived from conducting observations over a large number of individuals, allowing for a generalization from sampled groups to the populations they represent, and the rigor derived from the accumulation of a large number of sequentially ordered observations of an individual case permitting the detailed estimation of the dynamics of stability and transformation of behavior. This latter approach represents the idiographic approach (Allport, 1960; 1961), which capitalizes on the richness of detail that can be accessed through detailed study of the individual case, and is therefore often equated with qualitative research. Allport presented the idiographic approach in the context of personality psychology, a field of inquiry that takes great interest in the search for stability in the personality traits of individuals over time in the face of shifting environmental conditions, or *stimulus fields*.

Stimulus fields can vary a great deal as the number of variables is large enough to yield an ecologically valid description of the types of influences affecting behavioral outcomes. Stouffer (1941) invites us to an instructive "thought experiment" involving a contingency table with multiple categorical dimensions and notes that as the number of variables increases, the number of cells in the contingency table

increases very rapidly as well. Imagine, for instance, how students from lower SES or non-lower SES backgrounds (2 categories) who are male or female (4 categories), take instruction under treatment or control conditions (8 categories), with effective, somewhat effective, or non-effective teachers serving both conditions (24 categories), who all have been classified as belonging in one of the following reading proficiency categories: below basic reading proficiency, approaching proficiency, proficient, advanced (96 categories). If one distinguishes the home environments for each student as supportive or non-supportive, the number of possible ratings an individual can obtain on the variable constellation doubles to 192. In that light, it is not hard to appreciate how little individuals might have in common when it comes to the environmental and other baseline conditions under which learning and instruction takes place. This makes the individual case an understandable and intuitively appealing methodological choice, when the stimulus field harboring these environmental influences comes under closer scrutiny.

The Temporal Dimension

The idiographic perspective carries two major implications. One is that it identifies the need to focus on the situational specificity of behavior of in relation to stimulus fields, which dynamical scholars would refer to *exogenous processes*. The second implication is that “one must account for the *recurrences* and *stabilities* in personal behavior” (Allport, 1961, p. 312) over time. This temporal aspect of behavior, i.e., behavior relative to itself at a previous point time, is referred to in the dynamical literature as *endogenous* processes, and studying them facilitates understanding of its connection to a constantly evolving antecedent stimulus field.

In a linear causal model, causes are assumed to precede effects (Pearl, 2009). In a recursive causal model, causes and effects are assumed to precede each other in an ongoing interrelationship (Sawyer, 2005). In both instances, time is a critical aspect of our understanding of behavior. The argument also been made in the psychological literature that behavioral measures that ignore the temporal dimension are potentially misleading as they disregard the uniqueness of responses to contingencies that are time dependent, as well as the periodicity in behavioral variability both at the individual and the collective level of description. In psychology, periodicity has been measured in such variables as muscle activity cycles (electrocardiograms) and electrical activity recorded from the scalp (electroencephalograms) and eye movements (Barrett, Johnston, & Pennypacker, 1986). In education, studying the temporal dimension has productively informed our understanding of the effectiveness of behavior modification processes in the classroom in terms of the impact of teacher actions on student behavior, linking behavioral outcomes to their antecedent conditions (Hall et al., 1971; Neef, Shade, & Miller, 1994; O’Leary, Becker, Evans, & Saundargas, 1969).

Dewey (1929) describes the temporal span as the most important aspect of the educational process, yet the contributions of time to education have not often been

systematically investigated. Glass (1972) laid out time series analysis as a methodological framework for doing so, to deal specifically with the correlations between in time-ordered within-subject observations, the estimation of the constancy of statistical properties of a time trajectory over a longer time period, as well as the perturbation of interventions on a trajectory of outcome measurements. There has been little follow-up in educational research to meet the challenges he put forward (Koopmans, 2014a).

Ergodicity

In 2004, the Dutch psychologist Peter Molenaar published a paper entitled *A Manifesto on Psychology as Idiographic Science: Bringing the Person Back into Scientific Psychology, This Time Forever*. It appeared in the journal *Measurement* (Molenaar, 2004) and it was accompanied by seven peer responses. The article takes note of the fact that an orientation toward the individual case ($N = 1$) is almost completely absent from psychology, where conventional research typically observes and analyzes the behavior of large numbers of individuals, and computes central tendency and variability measures to characterize the group (interindividual variation). If a sample is randomly drawn from a population of interest, the sample results permit generalization to that population. There is a shortage of work, however, that takes it as its mission to observe and analyze the behavior of a particular individual over a single time series (intraindividual variation) in order to learn about the particularities of its behavior and how it evolves across the time spectrum of interest. The description of the individual case permits a level of detail in the description of processes and behaviors that would not be possible if those processes and behaviors are aggregated across groups.

As in psychology, applied researchers in education tend to investigate phenomena cross-sectionally, compute measures of central tendency across individuals to characterize group, and estimate variability in terms of the degree to which individual observations deviate from the means in their group. Thus, the effectiveness of educational interventions is estimated in terms of their impact on mean group outcomes. It is implicitly assumed in such designs that the measures used to estimate outcomes and predictors duly characterize the full time spectrum to which the measurements purport to apply. For example, in the course of a school day, or a school year, it is expected that *if* a large number of observations had been taken across the time spectrum, the statistical properties of the trajectory of ordered observations across time would be constant or at least predictable. Moreover, it is typically assumed that individual variation (error) is randomly distributed around the mean of the entire span of observations, as it would be in the measurements across individuals after a successful statistical modeling effort.

Both in the cross-sectional and in the longitudinal case, variances are traditionally defined in terms of the sum of squared differences of individual observations from their sample means, adjusted for sample size. In cross-sectional research, a

population mean is then estimated from the observed sample mean computed across individuals, and the reliability of this estimation is said to increase as the number of individuals who are sampled gets larger. Variability, in this context, is a quantification of the extent to which between subjects outcomes vary from the means of their groups (measurement errors). In the longitudinal case, variability quantifies the extent to which within subjects individual observations vary from the mean of the trajectory of measurements for that particular individual. In both cases, it is typically assumed that these errors are randomly distributed and independent of one another. In either case, it is an empirical question whether this assumption is justified.

In its more general form, the implication of the putative equivalence between the distribution of measurements across cases and within cases across the time spectrum is known as the *ergodic assumption*, which states that the latent variance structure across individuals corresponds to the latent variance structure within individual across the time spectrum. Another way of stating this idea would be to say that the data behavior we can assume in the individual case is somehow captured in the description of the group. The problem with this assumption is that the conclusions that are drawn based on group means do not necessarily apply to all individual cases in that group, and studying the individual cases may reveal idiosyncratic patterns over time that may or may not conform to the causal structure inferred from the cross-sectional group descriptions (Gu, Preacher, & Ferrer, 2014). While it may seem obvious that the ergodic assumption would therefore require empirical confirmation (Molenaar, 2004), the assumption is rarely discussed explicitly in our discipline (but see Gu et al., 2014 and Molenaar, Sinclair, Rovine, Ram, & Corneal, 2009).

The discussion of ergodicity speaks to the important methodological concern of how to account for sources of variability in observed data and the assumptions we are making about relevant data that we typically fail to collect. In part, however, the ergodic question also reflects a substantive concern. Given that traditional fixed effects models may not generalize to anyone in particular (Molenaar, 2004), perhaps the question “Is the program effective?” is the wrong one to ask, and should be rephrased as “For what type of student is what type of program effective?” The latter emphasis allows for different causal structures underlying the behavior of different individuals (Curran & Wirth, 2004).

In its purest form, idiographic science implies that no knowledge can be generalized beyond the single individual, and that the analysis of individuals on a case-by-case basis is the only legitimate basis for the acquisition of knowledge. In their response to Molenaar’s manifesto, Curran and Wirth (2004) rightly note that this position undermines one of the key goals of empirical science, namely external validity, i.e., the applicability of research findings to other people and/or settings than those that were studied. Both Curran and Wirth (2004) and Rogosa (2004) argue in their responses for a conditional between-subjects modeling approach “moving us into the interior of the space demarcated on one side by the strict study of the individual, and on the other side by the sole focus on the characteristics of the overall group (Curran & Wirth, 2004, p. 221).” However, if you define external validity in terms of places and people, as well as time, the challenge of attaining external validity extends to the adequate sampling in all three areas.

We may therefore just not be able to resolve all aspects of external validity within the confines of a single research design.

The use of within-subjects findings to fortify generalizations from between-subjects studies addresses an important external validity issue in conventional RCT designs. A fully reliable randomization of units to various treatment conditions may ensure that causal inferences about the effects of those conditions may be drawn (internal validity, Murnane & Willett, 2011; Murray, 1998). However, since the behavior of individuals or clusters within the sample for those studies may not necessarily be governed by a uniform causal structure, information coming from individual cases may provide important qualifiers to the group results, either as counterexamples, or salient illustrations of the group level processes (Cook & Campbell, 1979).

The possibility of using more refined cross-sectional models to study time dependencies deserves careful consideration. Latent growth curve modeling (e.g., Bryk & Raudenbush, 1992; Willett, 1988) allows for the estimation of nonlinear within-subjects processes (e.g., quadratic, cubic models), and it uses individual growth trajectories as input, and therefore, it can be argued that in many instances, growth modeling allows for valid inferences about both within and between subjects processes. The method also permits an analysis of individual growth trajectories before the cross-sectional components are added to the model. However, the growth modeling currently typically found in the educational research literature does not allow for the more refined estimation of periodicity, a-periodicity and non-periodic error dependencies, which would require the incorporation into the hierarchical regression framework of autoregressive, frequency domain or state space modeling features to estimate error dependencies over the short term and the long term of the trajectory (Koopmans, 2015; Shumway & Stoffer, 2011). In such cases, we would need a sufficient number of input data points per trajectory to ensure representativeness of findings across the time spectrum. The aggregation of findings across cases also results in the loss of information that could potentially be useful for the provision of counter examples to the interpretation of results from cross-sectional comparisons. This individual-level information cannot be recovered from the aggregated information (Rogosa, 2004). This situation calls for an analytical approach that incorporates a within subjects component to between subjects designs to solidify inferences to the population based on samples (Nesselroade, 2004; Rogosa, 2004).

One could argue that studying education and human development almost by definition calls for a within-subjects approach, notwithstanding our methodological predisposition toward group level analyses. Questions of causality in education almost automatically invoke an RCT framework (Raudenbush, 2005) and thereby also the way in which we tend to phrase our research questions. Complex dynamical systems perspectives are also concerned with the causality question, but they are not similarly inclined to articulate their research questions in terms of differences between group averages. Instead, the complexity angle is more likely to focus on how a behavioral trajectory operationalizes self-organizing processes within the system, and how those processes are affected by events that are external to this process. This interest expresses the causal question in a different way.

An Example of Data Suggesting a Non-Ergodic Structure

Can we assume, based on a single snapshot taken at an arbitrary point on the time spectrum, that all microstates within a given time range are equiprobable and that there would be no dependence between observations, had they actually been measured? Upon further empirical scrutiny, what kind of violations can we expect to encounter to this assumption? Let us take a cross-sectional result with a random distribution of individual observations around the group mean, and then play out two scenarios on the time spectrum, one of which would show such randomness, as the ergodic assumption would lead you to expect, and the other which illustrates non-randomness. Figure 8.1 shows these two scenarios in a simulation of $N = 1000$ successive observations, here called Y_t . The top panel on the left (Fig. 8.1a) shows a time series with a mean of zero and a random distribution of measurement error (white noise). The corresponding histogram (Fig. 8.1b) shows that the distribution

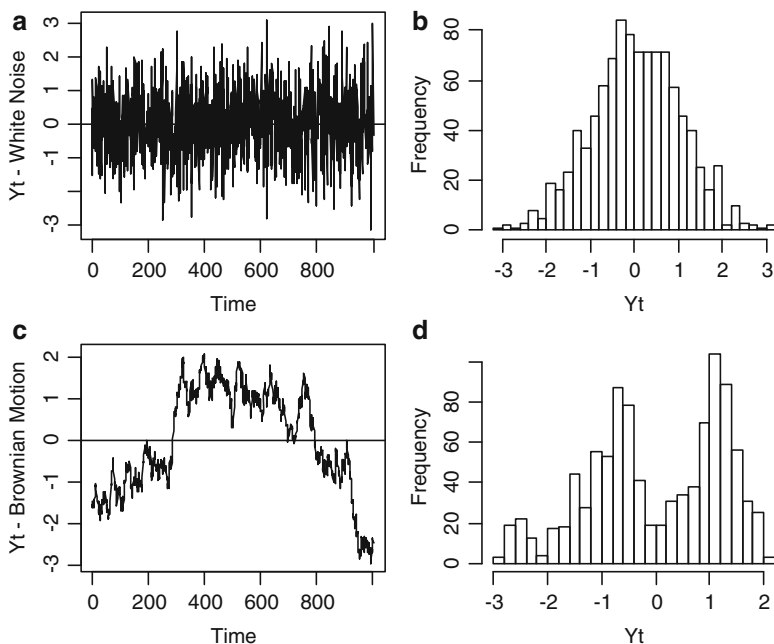


Fig. 8.1 (a) Simulated time series with a zero mean, a standard deviation of 1 and randomly distributed errors (white noise), (b) corresponding frequency distribution, (c) simulated time series with a zero mean, a standard deviation of 1.23 and non-randomly distributed errors (Brownian motion), (d) corresponding frequency distribution. $N = 1,000$ for both simulations

of errors approximates normality quite well. It can also be shown that the measurement errors are uncorrelated in such cases, and that the pattern revealed by this trajectory carries little information about what the trajectory might look like in the future. If we assume cross-sectional white noise, then, these data would be ergodic in relation to it.

Now consider the simulation result shown in the bottom two panels of the figure. While the mean in these distributions equals zero as well, the time series plot in Fig. 8.1c indicates that the errors are not random, but clearly show a pattern in the way individual observations deviate from the mean of the series. The tight clustering of observations to their immediate neighboring ones contrasts with what is shown in Fig. 8.1a, and it indicates a strong correlation between measurement errors. The histogram (Fig. 8.1d) shows that in this simulation, the mean does not characterize the distribution very well due to the bimodality that is also on display in the time series plot. Therefore, if we assume white noise in a collection of cross-sectional data on Y_t , these data are not ergodic in relation to it, as they display a non-corresponding variance structure, and a cross-sectional mean would actually misrepresents the mean of the observations that distributed this way across the time spectrum.

The simulated error scenario shown in the time in Fig. 8.1c is actually known in the dynamical literature as *Brownian motion*, or the *random walk*, and it characterizes unstable systems in which a high level of dependency between individual observations in close proximity is coupled with a high degree of volatility in the trajectory overall. As a result, there is no constancy in the statistical properties (mean, variance) characterizing the series in its entirety, as those properties heavily depend on the location of the observations on the trajectory. Therefore, conducting valid outcome measurements would require in this situation that the status of the ergodic assumption be addressed (Molenaar, 2004; Molenaar et al., 2009) by empirically establishing the variance structure underlying the sequence of the measurements.

The Dynamics of Daily High School Attendance Rates

One of my currently ongoing research projects is concerned with daily attendance in New York City public schools, which has recorded the daily attendance rates of all of its schools starting in 2004, and continuing up to the time of this writing. This research strictly follows a case study approach, albeit at the school level rather than at the level of individual students. Koopmans (Chap. 14) provides a detailed justification of the research agenda as well as its methodology. It will suffice here to say that the inspection of daily attendance trajectories over a longer time period allows us to discern patterns of non-randomness in temporal educational data that would remain hidden if conventional summary statistics are used to aggregate results across schools. Simply reporting mean daily attendance rates and their standard deviations is insufficient if the distribution of individual observations

Table 8.1 Demographic characteristics of School A, B, and C: Academic year 2013–2014

	School A	School B	School C
Enrollment	315	105	184
<i>Gender</i>			
% Female	50.2	77.3	59.8
<i>Ethnicity</i>			
% Asian	3.8	1.0	2.7
% Black	30.5	37.1	26.6
% Hispanic	64.1	60.2	65.2
% White	1.3	0.0	4.3
% Other	1.0	1.0	1.1
<i>Other</i>			
% English Language Learners	9.5	1.9	7.1
% Students with Special Needs	31.1	3.8	31.5
% Eligible for Free or Reduced Priced Lunch	100.0	90.5	83.2

Table 8.2 Summary statistics for the daily attendance rates in 2013–2014: School A, B, and C

	School A	School B	School C
Mean	.77	.55	.87
Standard deviation	.09	.14	.08
Minimum	.26	.17	.39
First quartile	.74	.47	.87
Median	.78	.53	.88
Third quartile	.82	.63	.90
Maximum	1.00	.99	.99

(i.e., attendance measures on any given day) is not random. A look at the attendance trajectory over time allows for an empirical verification of this assumption. A sample of approximately 180 observations in a year provides sufficient resolution to look closely at many aspects of the dynamics of school attendance as an ongoing process, and develop hypotheses about the susceptibility of these patterns to external influences (Koopmans, 2015). Below, I would like to illustrate this point by discussing the daily attendance trajectories in three New York City high schools.

Table 8.1 summarizes the basic enrollment and demographic information for the three schools in the 2013–2014 school year. It can be seen that the schools are similar in three important respects. All three are small high schools with a total enrollment ranging from 105 to 315. Moreover, the schools are demographically similar with an overwhelming majority of students being Black or Hispanic and about twice as many Hispanic as Black students. Furthermore, all three schools serve students from predominantly poor socioeconomic backgrounds, as can be seen by the high percentage of students eligible for free or reduced priced lunch.

Table 8.2 shows the traditional summary statistics for the attendance rates in these schools in the 2013–2014 school year. A total of 187 daily attendance rates were recorded in that year. The mean attendance rates vary considerably, ranging

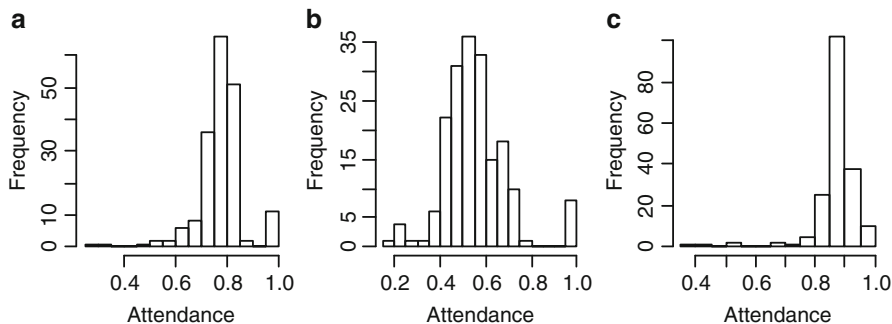


Fig. 8.2 (a, b, and c) Histograms of the daily attendance rates for the 2013–2014 School Year in School A, B, and C, respectively

from 55 % in School B to 87 % in School C. The corresponding medians are 53 and 88 %. Inspection of the other statistics in the table indicate a negatively skewed distribution in Schools A and C: the first and third quartile in those two schools are close to the median and to each other, while the minimum value in the distribution falls far below the first quartile. The quartiles are farther apart in School B than in the other two schools. The histograms shown in Fig. 8.2a, b, c and d further illustrate these features for the three schools.

Turning to the single-case dynamics, Fig. 8.3a, b, c and d show the attendance trajectories for the 2013–2014 school year in Schools A, B, and C. A straight line is superimposed in each plot to represent the average daily attendance for that school for the entire school year. In School A, daily attendance rates appear to be fairly stable across the time spectrum, except for some turbulence toward the middle of the trajectory, possibly reflecting the ramifications of inclement weather during the winter months. This turbulence is shown by the dips into lower values, but also their off-set with peaks into higher values during the same period. In School C the low-range dips are even more pronounced, though not compensated for by instances of unusually high attendance rates. It should be noted, however, that mean attendance rates are much higher on the average in School C to begin with, and very stable as well as can be seen by the low level of variability between observations.

In School B, attendance rates show less stability over time than those in the two other schools, and there is a steady decline in those rates as the school year progresses toward the middle of the year, and a slight recovery in the second half of that year. It can also be seen that in School A and B, the attendance rates become highly variable toward the end of the school year, reflecting in all likelihood the end of year celebrations and wrap-up. In School C the trajectory does not appear perturbed to the same extent toward the end of the year, but simply shows above average attendance for the last string of school days.

School A and C illustrate a stationary process, which is to say that there are no clear signs of an upward or downward trend, heteroscedasticity, or other transformations in the outlook of the series that depend on the position of the data points

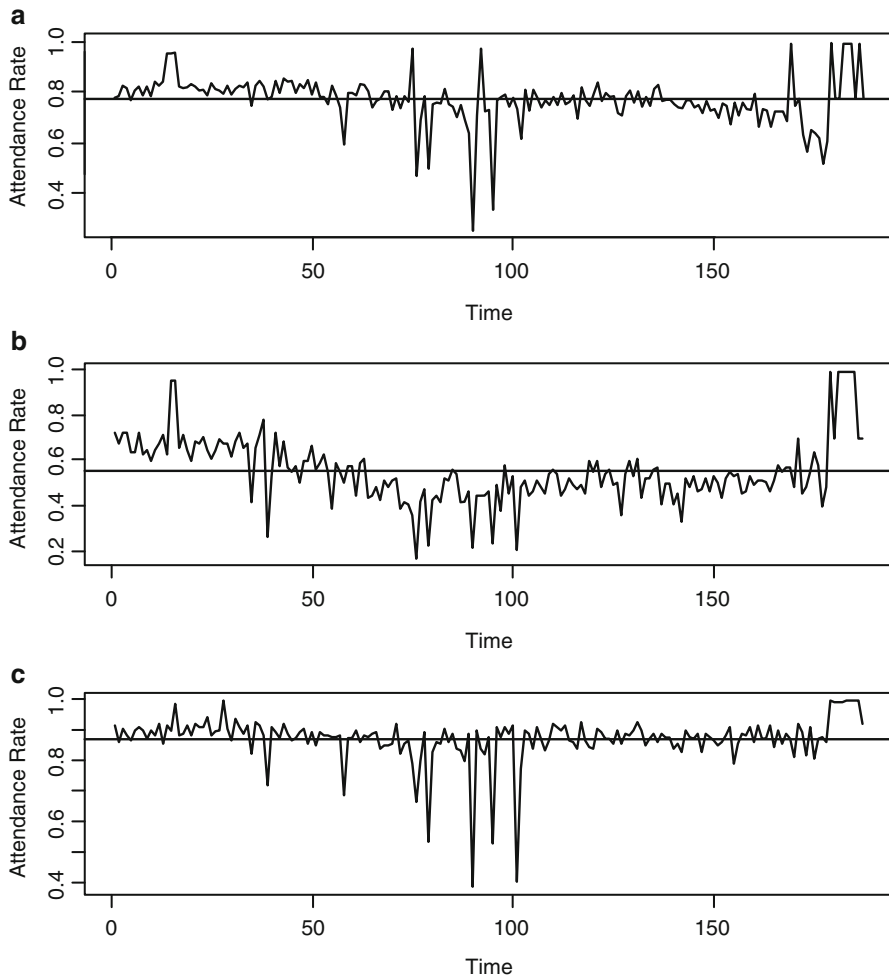


Fig. 8.3 (a, b, and c) Daily attendance rates for the 2013–2014 School Year in School A, B, and C, respectively, as a time series

on the timeline. The peaks and valleys are outlying observations in an otherwise consistent trend (and would be treated as such in the statistical analysis, see Koopmans, 2014c for an example). The trajectory shown for School B in Fig. 8.3b, on the other hand shows a mild resemblance to the Brownian motion shown in Fig. 8.1c. The mean of the series underpredicts the observed data for extended periods in the course of the school year (roughly corresponding to fall and spring), while it overpredicts in the winter months.

In all three instances discussed here, however, traditional measures of central tendency do not characterize very well what goes on in these data, and they conceal important time-dependent features. A case-by-case scrutiny of the observations and

their dependency over time provides an altogether different appreciation of daily attendance in these three schools. Educational practitioners are likely to be aware of the fluctuations in attendance rates in their classrooms and school buildings, as well as their seasonal dependence, but these conventional measures do not operationalize this aspect of the variability in those rates (Koopmans, 2015), and therefore, the study of daily attendance rates over time is a useful endeavor in its own right.

Endogenous Processes and Feedback Loops

Successful causal attribution requires that baseline conditions in the system are measured in sufficient detail to estimate the propensity toward change in the system, as well as the change processes that actually occur in response to given interventions (Koopmans, 2014b). In conventional research, the measurement of the baseline typically involves at most a limited number of pretest observations and a comparison between pretest and posttest to determine whether change has been produced by intervention. The change process is then inferred a-posteriori. A convincing description of the stability in a system with respect to certain measurement outcomes requires a reliable estimation of the intra-subject variability in those outcomes. Such estimation, in turn, requires a much larger number of measurement occasions than is typically provided in traditional pretest-posttest designs.

Moreover, given that establishing a cause-effect relationship at the aggregate level does not necessarily carry over to all individuals, the individual case may prove to be a source for instructive counter-examples. From the vantage point of complex dynamical systems, there are two interrelated aspects to causality that are not captured in conventional research designs: the *endogenous process* and the *recursive feedback loop* between different parts of the system. The endogenous process explains observed outcomes at a given point in time in terms of previous measurements of that same outcome and thus can be seen as an indicator of the adaptive behavior of the system with respect to that particular outcome (e.g., learning) irrespective of a particular impulse from sources external to the system. An educational intervention, in this context, can be seen as an *exogenous* process that may or may not have an impact on the endogenous process. The analytical framework for conducting such assessments has been in place for a long time, and has been part and parcel of the empirical behavior modification literature referred to above, that measures the impact of given interventions on the endogenous process.

The measurement of recursive feedback loops is a challenge of an altogether different magnitude, as it requires an estimation of the behavior of individual members of a given system in relation to the behavior of the system at large. While there is a rich literature to appreciate such feedback loops in theory (Koopmans, 1998; McKelvey, 2004; Minuchin & Fishman, 1981; Sawyer, 2005), the empirical literature has not yet risen to the challenge of connecting these two levels of description within a single analytical framework (but see Salem, 2013). In the linear

paradigm, there is a cross-over between levels of description in the hierarchical multilevel designs equipped to deal with nested data structures (Bryk & Raudenbush, 1992; Gelman & Hill, 2007). However, these approaches do not deal with the reciprocal nature of the influences between the behavior of systems and that of the individual members making up those systems, nor do they provide great detail about how this process plays out over time. Both aspects are central to the interests of complex dynamical scholarship because it describes the process of self-regulation through which systems maintain their structure, composition and integrity as a distinct functional unit. The analytical focus on endogenous processes does not at all preclude an analysis of external influences that may perturb the system (McDowall, McCleary, Meidinger, & Hay, 1980). In fact, the impact of those external influences on the behavior of the system is better understood if we have knowledge of the endogenous processes through which the system maintains itself.

Concluding Remarks

The single case presents an alternative perspective on the question of cause and effect in education. It can provide us with a fine-grained description of the transformations that constitute the effects of interest. In the attendance data discussed above, for example, one could speculate about a “winter effect” on the attendance trajectories, where inclement weather impacts the transportation options for students and these options may vary depending on the location of the school building. Likewise, these influences can be estimated relative to other causal factors such as parental support and effective school building leadership (Koopmans, 2015). The confirmation of such causal processes requires a triangulation of the data presented here with those from other sources. The detailed sampling of observations across the time spectrum enables us to contemplate these causal processes to begin with. They remain hidden in cross-sectional summary statistics.

Traditional summary statistics such as means and standard deviations, as well as ordinary least squares regression, make assumptions about the data that do not carry over very well to time series data. When ordered observations over time are analyzed, two issues invariably come up: error dependency between observations and the constancy of the statistical properties of the series across the entire time spectrum. The dependency between observations across the time spectrum is referred to in the time series literature as *autocorrelation*. The constancy of statistical characteristics across the time spectrum is called *stationarity* and the detection of these two features is a central part in the description and analysis of most time-ordered data (Box & Jenkins, 1970). Comparison of Fig. 8.1a and c above illustrates how drastically the appearance of outcome trajectories can differ depending on whether the data are stationary (Fig. 8.1a) or not (Fig. 8.1c). Stationary time series are not necessarily random in the sense of white noise, as they can also contain clustering and dependencies between observations i.e., autocorrelation,

that needs to be modeled as well. Koopmans (Chap. 14) provides further elaboration on that aspect of the sequential ordering of the data.

Dealing with the temporal aspects of data is common practice in education, as in traditional pretest-posttest comparisons and the use of growth modeling to describe changes in student achievement over time (Bryk & Raudenbush, 1992; Rogosa, 2004). However, these designs are not equipped to address the dynamical features of changes over time, nor are they typically used to address dynamical questions about those changes. Rogosa, Floden, and Willett (1984) provide an interesting exception to this latter point by examining the stability in teacher behavior over time. However, while very dynamical in the way it is framed, with four to six measurement occasions the study does not sample the time dependent process adequately to make meaningful inferences about the stability of teacher behavior as a temporal phenomenon. This lack representativeness of sampling across the time spectrum is characteristic of most of the cross-sectional work done in educational research.

It is far beyond the scope of the work presented here to summarize all we have learned over the years from the study of single cases in education. Suffice to say here that the case study has many times been productively used in the field, such as for example in the school district reform literature (e.g., Cuban, 2010; Reville, 2007), the developmental literature (e.g., Bassano & van Geert, 2007; Brown, 1973), or the aforementioned behavior modification studies. Sometimes, this work is explicitly concerned with hypothesized dynamical processes (e.g., Bassano & van Geert, 2007; Johnson, 2013; Laidlaw, Makovichuk, Wong, & O'Mara, 2013; Molenaar et al., 2009), but often it is not. This chapter argues that there is clear potential in the study of the single case to obtain a deeper understanding of the dynamical underpinnings of cause and effect, without the information loss that comes with the aggregation of information across cases (Rogosa, 2004). The particular contribution of statistics to this area lies in the detailed analyses of the temporal ordering of the information elements. Ethnographic research, on the other hand, can provide the thick descriptions that may one day form the basis for the development of meaningful mathematical models underlying interactive behavior (Dobbert & Kurth-Schai, 1992).

In the realm of quantitative research here is a trade-off between sampling rigor across cases between subjects and across observations within subjects. The single case design attains sampling rigor in this latter respect, and thus address different types of questions that are of interest to the field. While ethnographic research may not require this type of sampling rigor for inferential purposes, its thick descriptions are very suitable for the description of the temporal features of behavior. Miles and Huberman (1994) explain how qualitative approaches can be used to analyze the temporal features of the behavior of individuals and organizations. Both the qualitative and the quantitative single case study are well-equipped to address the question of the underlying dynamics of stability and change in greater detail, but have been used infrequently for that purpose.

In its most uncompromised reading, the idiographic approach altogether rejects the aggregation of information for purposes of statistical summary as a matter of

principle (Molenaar, 2004), as a result of which study of human development would boil down to the accumulation of all of our life stories. This would leave us without an accessible empirically based knowledge structure about what we can learn from those stories (Curran & Wirth, 2004). Rather than viewing the single case as a radical alternative to nomothetic science, it may be more productive to view it as a necessary supplement to it. The triangulation of data from large scales studies with those from single case designs can enrich those studies with added detail about causal mechanisms, and challenge the interpretation of aggregated findings through the provision of individual counterexamples (Flyvbjerg, 2006).

Perhaps it is possible one day to deal with complex cross-sectional data structures and detailed temporal information about those structures simultaneously. In the methodological field, the discussion about integrating state space techniques and structural equation modeling (Browne & Zhang, 2007; Molenaar, 2009; Molenaar, van Rijn, & Hamaker, 2007) is a highly promising development. However, this work seems to be in its early stages and does not yet provide a clear answer to the question how we can meaningfully reduce the sheer volume of the information from individual cases to manageable proportions without losing the level of resolution that is needed to study the finer details of dynamical complexity, the representativeness of observations across cases, or both.

Do we need an ergodic argument to make a case for single case designs? Perhaps this argument unnecessarily complicates a relatively straightforward justification for using such designs (Thum, 2004). Rather than framing the case for the single case in terms of underlying assumptions about data distributions, we can also promote the intrinsic interest of the particularities of the single case for the sake of learning and scholarship (Stake, 1994). The readiness to investigate the particularities of the individual case through statistical means would represent an interesting point of departure in the field of education, which lacks a time series analysis tradition at this point.

Within the field of complex dynamical systems, the increasing reliance on the investigation of complex processes through statistical means is a favorable development as well, as it takes us away from the reliance on woolly inspirational metaphors, such as “the edge of chaos” (Dodds, 2012; Goldstein, 1995; Koopmans, 2009; Waldrop, 1992) to the development of observation and measurement strategies specifically designed to identify the specific empirical referents of such hypothesized processes and constructs (Bak, 1996; Jensen, 1998; Koopmans, 2009, 2015). Such work would generate knowledge that is falsifiable as well as useful to educational practitioners and policy makers, as it gives the dynamical aspects of their thinking about education the representation it deserves in our research efforts.

References

- Allport, G. W. (1960). *Personality and social encounter: Selected essays*. Boston, MA: Beacon Press.
- Allport, G. W. (1961). *Pattern and growth in personality*. New York, NY: Holt, Rinehart & Winston.
- Bak, P. (1996). *How nature works: The science of self-organized criticality*. New York: Springer.
- Barrett, B. H., Johnston, M. M., & Pennypacker, H. S. (1986). Behavior: Its units, dimensions and measurement. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 156–200). New York: Guilford.
- 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 (Ed.), *Steps toward an ecology of mind: A revolutionary approach to man's understanding of himself* (pp. 61–71). [Reprinted from *Man*, XXXV, Article 199].
- Bloom, J. W., & Volk, T. (2007). The use of metapatterns for research into complex systems of teaching, learning, and schooling. *Complicity: An International Journal of Complexity and Education*, *4*, 45–68.
- Bogdan, R. C., & Biklen, S. K. (1982). *Qualitative research for education: An introduction to theory and methods*. Boston, MA: Allyn & Bacon.
- Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis, forecasting and control*. San Francisco, CA: Holden-Day.
- Brown, R. W. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Browne, M. W., & Zhang, G. (2007). Repeated time series models for learning data. In S. M. Boker & M. J. Wenger (Eds.), *Data analytic techniques for dynamical systems* (pp. 25–45). Mahwah, NJ: Erlbaum.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.
- Burns, R. B. (2000). *Introduction to research methods* (4th ed.). London: Sage.
- Burrell, G. (1979). *Sociological paradigms and organizational analysis*. London: Heinemann.
- Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research in teaching. In N. L. Gage (Ed.), *Handbook for research in teaching*. Chicago, IL: Rand-McNally.
- Combes, B. H., & Patterson, L. (2013, April). *Leadership in special education: Using human systems dynamics to address sticky issues*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Cook, T., & Campbell, D. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston, MA: Houghton Mifflin.
- Cuban, L. (2010). *As good as it gets: What school reform brought to Austin*. Cambridge, MA: Harvard University Press.
- Curran, P. J., & Wirth, R. J. (2004). Interindividual differences in intraindividual variation: Balancing internal and external validity. *Measurement*, *2*, 219–227.
- Dewey, J. (1929). *The sources of a science in education*. New York: Liveright.
- Dobbert, M. L., & Kurth-Schai, R. (1992). Systematic ethnography: Toward an evolutionary science of education and culture. In M. D. LeCompte, W. L. Millroy, & J. Preissle (Eds.), *The handbook of qualitative research in education* (pp. 93–159). San Diego, CA: Academic Press.
- Dodds, J. (2012, March). *Ecology and psychoanalysis at the edge of chaos*. Presented at the International Nonlinear Science Conference, Barcelona, Spain.
- Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, *12*, 219–245.

- Fromberg, D. P. (2010). How nonlinear systems inform meaning and early education. *Nonlinear Dynamics, Psychology and Life Sciences*, 14, 47–68.
- Garner, J., & Russell, D. (2014, April). *The symbolic dynamics of self-regulated learning: Exploring the application of orbital decomposition*. Paper presented at the annual meeting of the American Educational Research Association, Philadelphia, PA.
- Gelman, A., & Hill, J. (2007). *Data analysis using regression and multi-level/hierarchical models*. New York: Cambridge University Press.
- 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. (1995). The Tower of Babel in nonlinear dynamics: Toward the clarification of terms. In R. Robertson & A. Combs (Eds.), *Chaos theory in psychology and the life sciences* (pp. 39–47). Mahwah, NJ: Erlbaum.
- Gu, F., Preacher, K. J., & Ferrer, E. (2014). A state space modeling approach to mediation analysis. *Journal of Educational and Behavioral Statistics*, 39, 117–143.
- Guastello, S. J., & Gregson, R. (2011). *Nonlinear dynamical systems analysis for the behavioral sciences using real data*. New York: CRC Press.
- Guevara, P., López, L., Posh, A., & Zúñiga, R. (2014). A dynamic nonlinear model for educational systems: A simulation study for primary education. *Nonlinear Dynamics, Psychology and Life Sciences*, 18, 91–108.
- Hall, R. V., Fox, R., Willard, D., Goldsmith, L., Emerson, M., Owen, M., et al. (1971). The teacher as observer and experimenter in the modification of disputing and talking-out behaviors. *Journal of Applied Behavior Analysis*, 4, 141–149.
- Hetherington, L. (2013). Complexity thinking and methodology: The potential of ‘complex case study’ for educational research. *Complicity: An International Journal of Complexity in Education*, 10, 71–85.
- Jensen, H. J. (1998). *Self-organized criticality: Emergent complex behavior in physical and biological systems*. Cambridge: Cambridge University Press.
- Johnson, L. (2013). *A case study exploring complexity applications to qualitative inquiry*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Jörg, T., Davis, B., & Nickmans, G. (2007). Position paper: Toward a new complexity science of learning and education. *Educational Research Review*, 2, 145–156.
- 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. (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* (pp. 506–526). New York: Cambridge University Press.
- Koopmans, M. (2014a). Nonlinear change and the black box problem in educational research. *Nonlinear Dynamics, Psychology and Life Sciences*, 18, 5–22.
- Koopmans, M. (2014b). Change, self-organization and the search for causality in educational research and practice. *Complicity: An International Journal of Complexity and Education*, 11, 20–39.
- Koopmans, M. (2014c, May). *Using ARIMA modeling to analyze daily high school attendance rates*. Presented at the Modern Modeling Methods (M3) conference, Storrs, CT.
- Koopmans, M. (2015). A dynamical view of high school attendance: An assessment of short-term and long-term dependencies in five urban schools. *Nonlinear Dynamics, Psychology and Life Sciences*, 19, 65–80.
- Kramer, P. D. (2014, October 19). Why doctors need stories. *New York Times Sunday Review*, p. 1, 7.
- Laidlaw, L., Makovichuk, L., Wong, S., & O’Mara, J. (2013, April). *Complexity, pedagogy, play: On using technology within emergent learning structures with young learners*. Presented at the American Educational Research Association, San Francisco, CA.

- Lemke, J., & Sabelli, N. H. (2008). Complex systems and educational change: Toward a new research agenda. *Educational Philosophy and Theory, 40*, 118–129.
- Lewin, K. (1947). Frontiers in group dynamics. *Human Relations, 1*, 5–41.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage.
- Marion, R., Klar, H., Griffin, S., Brewer, C., Baldwin, A., & Fisk, W. (2012, March). *A network analysis of complexity dynamics in a bureaucratically pressured organization*. Presented at the international nonlinear science conference, Barcelona, Spain.
- Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry. *Educational Researcher, 33*, 3–11.
- Maxwell, J. A. (2012). The importance of qualitative research for causal explanation in education. *Qualitative Inquiry, 18*, 655–661.
- McDowall, D., McCleary, R., Meidinger, E. E., & Hay, R. A., Jr. (1980). *Interrupted time series*. Thousand Oaks, CA: Sage. Sage university series on quantitative applications in the social sciences.
- McKelvey, B. (2004). Complexity science as an order-creation science: New theory, new method. *Emergence, Complexity and Organization, 6*, 2–27.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.
- Minuchin, S., & Fishman, H. (1981). *Family therapy techniques*. Cambridge, MA: Harvard.
- Molenaar, P. C. M. (2004). A manifesto on psychology as in idiographic science: Bringing the person back into psychology, this time forever. *Measurement, 2*, 201–218.
- Molenaar, P. C. M. (2009). *State space techniques in structural equation modeling: Transformation of latent variables in and out of latent variable models*. Retrieved November 12, 2014, from <http://www.hhdev.psu.edu/media/dsg/files/StateSpaceTechniques.pdf>
- Molenaar, P. C. M., Sinclair, K. O., Rovine, M. J., Ram, N., & Corneal, S. E. (2009). Analyzing developmental processes on an individual level using non-stationary time series modeling. *Developmental Psychology, 45*, 260–271.
- Molenaar, P. C. M., van Rijn, P., & Hamaker, E. (2007). A new class of SEM model equivalences and its implications. In S. M. Boker & M. J. Wenger (Eds.), *Data analytic techniques for dynamical systems* (pp. 189–211). Mahwah, NJ: Erlbaum.
- Murnane, R. J., & Willett, J. B. (2011). *Methods matter: Improving causal inference in educational and social science research*. New York: Oxford University Press.
- 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 Educational Research. Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.
- Neef, N. A., Shade, D., & Miller, M. (1994). Assessing influential dimensions of reinforcers on choice in students with serious emotional disturbance. *Journal of Applied Behavior Analysis, 25*, 555–593.
- Nesselroade, J. R. (2004). Yes, it's time: Commentary on Molenaar's manifesto. *Measurement, 2*, 227–230.
- O'Leary, K. D., Becker, W. C., Evans, M. B., & Saundargas, R. A. (1969). A token reinforcement program in a public school: Replication and systematic analysis. *Journal of Applied Behavior Analysis, 2*, 3–13.
- Pearl, J. (2009). *Causality: Models, reasoning and inference* (2nd ed.). 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. (2014). A nonlinear dynamical systems approach to real-time teacher behavior: Differences between teachers. *Nonlinear Dynamics, Psychology, and Life Sciences, 18*, 23–46.
- Piaget, J. (1967). *Six psychological studies*. New York: Vintage.
- Raudenbush, S. W. (2005). Learning from attempts to improve schooling: The contribution of methodological diversity. *Educational Researcher, 34*, 25–31.

- Reville, S. P. (2007). *A decade of school reform: Persistence and progress in the Boston Public Schools*. Cambridge, MA: Harvard University Press.
- Rogosa, D. (2004). Some history on modeling the processes that generate the data. *Measurement*, 2, 231–234.
- Rogosa, D., Floden, R., & Willett, J. B. (1984). Assessing the stability of teacher behavior. *Journal of Educational Psychology*, 76, 1000–1027.
- Salem, P. (2013). The complexity of organizational change: Describing communication during organizational turbulence. *Nonlinear Dynamics, Psychology, and Life Sciences*, 17, 49–65.
- Sawyer, R. K. (2005). *Social emergence: Societies as complex systems*. New York: Cambridge University Press.
- Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagné, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39–85). Chicago, IL: Rand-McNally.
- Shumway, R. H., & Stoffer, D. S. (2011). *Time series and its applications: With R examples*. New York: Springer.
- Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher*, 31, 15–21.
- Smith, L. M., & Geoffrey, W. (1968). *The complexities of the urban classroom*. New York: Holt, Rinehart, & Winston.
- Sprott, J. C. (2003). *Chaos and time series analysis*. New York: Oxford University Press.
- Stake, R. E. (1994). Case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 236–247). Thousand Oaks, CA: Sage.
- Stamovlasis, D., & Sideridis, G. D. (2014). Ought-avoidance: Nonlinear effects on arousal under achievement situations. *Nonlinear Dynamics, Psychology and Life Sciences*, 18, 67–90.
- Stamovlasis, D., & Tsapalis, G. (2012). Applying catastrophe theory to an information-processing model of problem solving in science education. *Science Education*, 96, 393–410.
- Steenbeek, H., Jansen, L., & van Geert, P. (2012). Scaffolding dynamics and the emergence of problematic learning trajectories. *Learning and Individual Differences*, 22, 64–75.
- Stouffer, S. A. (1941). Notes on the case study and the unique case. *Sociometry*, 4, 349–357.
- Tani, G., Corra, U. C., Basso, L., Benda, R. N., Ugrinowitsch, H., & Choshi, K. (2014). An adaptive process model of motor learning: Insights for the teaching of motor skills. *Nonlinear Dynamics, Psychology and Life Sciences*, 18, 47–66.
- Thum, Y. M. (2004). Some additional perspectives on representing behavioral change and individual variation in psychology. *Measurement*, 2, 235–240.
- van der Maas, H. L. J., & Molenaar, P. C. (1992). Stagemwise cognitive development: An application of catastrophe theory. *Psychological Review*, 99, 395–417.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Waldrop, M. M. (1992). *Complexity: The emerging science at the edge of order and chaos*. New York: Simon & Schuster.
- Watzlawick, P., Weakland, J., & Fish, R. (1974). *Change: Principles of problem formation and problem resolution*. New York: Norton.
- White, D. G., & Levin, J. A. (2013, April). *Navigating the turbulent waters of school reform guided by complexity theory*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Willett, J. B. (1988). Questions and answers in the measurement of change. *Review of Research in Education*, 15, 345–422.
- Yin, R. K. (2000). Rival explanations as an alternative to reforms as ‘experiments.’ In L. Bickman (Ed.), *Validity and social experimentation: Donald Campbell’s legacy* (pp. 239–266). Thousand Oaks, CA: Sage.