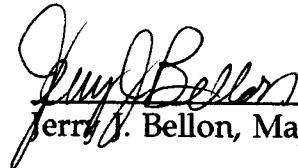
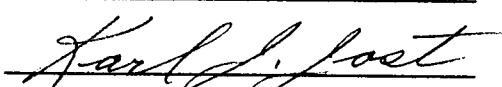


To the Graduate Council:

I am submitting herewith a dissertation written by Susan Denise Wilburn entitled "A Qualitative Study of Legislated Assessment Policy in Tennessee." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Education.


Jerry J. Bellon, Major Professor

We have read this dissertation
and recommend its acceptance:


Accepted for the Council:


Associate Vice Chancellor and
Dean of The Graduate School

**A QUALITATIVE STUDY
OF LEGISLATED ASSESSMENT
POLICY IN TENNESSEE**

**A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville**

**Susan Denise Wilburn
December 1996**

Copyright © Susan Denise Wilburn, 1996
All rights reserved

DEDICATION

This dissertation is dedicated to my husband

Bill Wilburn

ACKNOWLEDGMENTS

My deepest appreciation goes to the members of my committee: Jerry Bellon, Karl Jost, Ed Roeske and Ron Taylor. They were at times the rigor to my imagination and at other times the imagination to my rigor.

The support, encouragement and insight given by my family and friends has been immeasurable and has value that cannot be repaid in words alone.

The opportunity for self-organized learning allowed for and facilitated by the faculty will nurture and strengthen my future work in the field of education.

ABSTRACT

The general purpose of this study was to understand the impact of legislated assessment practices on student academic performance in Tennessee. Specifically, the study was to examine the Education Improvement Act (EIA) of 1992 to determine the strengths and weaknesses of its design in terms of related theory, "best" educational practices, and its potential to advance policy objectives. As a result of the study, recommendations were made to improve educational policymaking as it affects educational practice.

The conditions of this study met the primary requirements for using a case study research design. Documents and records were analyzed that preserve the historical context and decision-making processes of educational policymakers in Tennessee. Specifically, the documents and records related to the use of assessment as part of the EIA's design were reviewed. In an effort to stay as close as possible to the historical context of the enactment of the EIA, documents and records were collected from each major policy agency and then analyzed using a process coding within an open coding technique. To corroborate the evidence presented in the analysis of the data, the researcher used a data triangulation strategy, comparing concepts and categories defined in data from the legislative meetings across multiple-agency settings.

Major findings were that: (1) accountability emerged as the major theme of the EIA, (2) judicial intervention created a shift in the importance and type of influences on past legislative decision-making patterns, (3) the basic assumptions of educational policymakers support a mechanistic understanding of schooling systems, (4) the intended outcomes of the EIA were to improve education in the state and decrease political risk, (5) the

unintended outcome was to constrain educational practice to a more traditional rather than innovative practice, (6) the current assessment program strengthened the negative feedback loop in the schooling system, and (7) the EIA did not explicitly provide for a positive feedback loop.

Conclusions and recommendations of the study focused on the need for educational policy that provides for building the capacity of teachers by restoring the evaluation of student performance to the classroom level. This would require an assessment program that is formative, contextual, and highly related to the learning targeted for improvement. This would also require a commitment to professional development in the area of assessment through a dedicated percentage of the education budget.

TABLE OF CONTENTS

CHAPTER 1	1
INTRODUCTION	1
Statement of the Problem	5
Purpose	6
Significance of Study	7
Assumptions.....	8
Limitations and Delimitations	11
Methods and Procedures.....	12
 CHAPTER 2.....	 14
REVIEW OF RELATED LITERATURE.....	14
Historical Influences on Educational Policymaking in Tennessee	 14
<i>The Influence of First Policies.....</i>	<i>14</i>
<i>Educational Funding as a Historic Issue.....</i>	<i>15</i>
<i>The Influence of Politics.....</i>	<i>16</i>
The Impact of Federal Legislation on Educational Policymaking in Tennessee.....	 18
<i>Federal Aid Legislation.....</i>	<i>18</i>
<i>Accountability for Federal Aid:</i> <i>Legislated Assessment Begins.....</i>	 <i>19</i>
<i>Federal Legislation Coupled With Judicial Intervention.....</i>	<i>22</i>
<i>A Federal Template for Educational</i> <i>Policymaking in Tennessee</i>	 <i>25</i>
A Theoretical Perspective on Educational Policymaking.....	28
<i>Problems Inherent With a Linear Production-</i> <i>Function Model.....</i>	 <i>28</i>
<i>A Nonlinear Perspective on Assessing Schools</i> <i>as Social Systems.....</i>	 <i>31</i>
<i>Assessment and Changes in Learning: First-Order Change.....</i>	<i>38</i>
<i>Second-Order Change in the Teaching-Learning Process.....</i>	<i>40</i>
<i>Third-Order Change in the Teaching-Learning Process</i>	<i>41</i>
The Evolution of Policy Design in Tennessee.....	45
<i>Emergence from Tradition</i>	<i>45</i>
<i>New Possibilities.....</i>	<i>49</i>
<i>Re-thinking the Educational Policy Process.....</i>	<i>51</i>
Educational Policymaking in Tennessee: A Legislative Perspective	 53
Summary.....	55

CHAPTER 3.....	57
METHODS AND PROCEDURES.....	57
Introduction.....	57
Data Sources.....	58
Techniques for Data Collection and Analysis.....	59
<i>A Rationale for Qualitative Research.....</i>	<i>59</i>
<i>The Role of Qualitative Research in the Policy Process.....</i>	<i>60</i>
<i>The Iterative Process of Qualitative Research.....</i>	<i>62</i>
<i>Insight Leading to Interpretation.....</i>	<i>64</i>
<i>Interpretation as Positive Feedback.....</i>	<i>66</i>
Rationale for the Case Study as the Research Design.....	67
Rationale for Document Analysis as a Method of	
Data Collection and Analysis.....	72
A Rationale for Data Triangulation.....	80
Summary.....	81
 CHAPTER 4.....	 83
FORMULATION OF THE EIA.....	83
Policy Formulation: The Conditions for On-going Reform.....	84
The Initiator of the Education Reform Debate.....	88
Teachers Come to the Table.....	93
Studies, Proposals, and Resolutions:	
Analysis from Within the System.....	95
A Study of the Better School Program's Testing Program.....	99
Court, Consultant, and the Business Community	
Join the Debate: Intervening Conditions.....	100
Analysis Turns to Blame.....	111
Blame Turns to Evidence: The Performance Audit	
and the Court Decision.....	114
Evidence Turns to Plans.....	119
Plans Turn to the Testimony of Dr. William Sanders.....	126
Testimony Turns to Reactions and Preaching.....	130
The Resulting Action by the General Assembly.....	136
The Education Improvement Act:	
A Repeated Cycle of Policymaking.....	145
<i>Beginning Conditions.....</i>	<i>149</i>
<i>Context.....</i>	<i>151</i>
<i>Intervening Conditions.....</i>	<i>152</i>
<i>Action/Interaction Strategies.....</i>	<i>153</i>
<i>Unintended Outcomes.....</i>	<i>154</i>

CHAPTER 5.....	158
SUMARY, FINDINGS, CONCLUSIONS AND RECOMMENDATIONS	158
Summary	158
The Problem.....	159
The Purpose	160
Procedures	160
Findings	161
Conclusions.....	168
Recommendations	171
REFERENCES	173
VITA	190

CHAPTER 1

INTRODUCTION

In compliance with national educational policy, Tennessee has shaped and implemented educational change through a series of mandates holding teachers and local administrators accountable for the proposed learning outcomes of mandated change. Varying systems of evaluation were developed as part of these mandates. Historically, the purpose of evaluation in education was to fulfill "the requirement that educators be accountable for federal funds received under the Elementary and Secondary Education Act's (ESEA) Title I and III" (Worthen, 1995, p. 30). During the period of 1965–1975,

[Evaluation's] theories and philosophy, its proposed designs and instruments, and even its training of personnel were tilted more toward evaluating discrete, specially funded projects than toward evaluation of broad outcomes of curriculum and instruction (Worthen, p. 30).

As federal evaluation mandates were withdrawn during the 1980s, state and local agencies continued to evaluate, but with their own purposes in mind—"to shape local or state decisions about their Title I or III projects" (Worthen, 1995, pp. 30-34).

In 1984, Governor Lamar Alexander worked with the Tennessee General Assembly to formulate and enact the Comprehensive Education Reform Act (CERA), also known as the Better Schools Program (Green, Smith, & Hydorn, 1995). CERA served as Tennessee's response to the national concern for the decline of American test scores in comparison with those of economic competitors. Important components of this legislation were teacher pay based on performance, grade-level curricula and assessment

tools developed for the mastery of basic skills, and an instructional model developed for the delivery of the established curriculum and as the basis for a teacher evaluation model.

In 1988, four years after CERA, 66 small school districts in Tennessee filed a suit against Tennessee, alleging that the inequity of educational funding across districts violated students' constitutional right to equal educational opportunities.

In July 1991, a ruling in favor of TSSS [Tennessee Small School Systems] was rendered. The General Assembly was ordered to reform school funding before June 30, 1992. The State appealed, and in June of 1992, the Appeals Court reversed the trial court. On September 3, 1992, The Tennessee Supreme Court agreed to review the case. The TSSS lawsuit appeal (November 1992) was heard after the Education Improvement Act (EIA) of 1992 had been signed into law. The EIA contained a section on school funding reform called the Basic Education Program (BEP)—(Green, Smith & Hydorn, 1995, pp. 8-9).

Linked to the funding reform of the EIA was a specific accountability system to track funding and monitor student gains known as the Tennessee Value-Added Assessment System (TVAAS). TVAAS was designed to be a more specific accountability system than CERA.

Considerable controversy exists over the use of the Tennessee Comprehensive Assessment Program (TCAP) achievement test as the testing instrument for the TVAAS accountability system. The TCAP achievement test is a standardized, multiple-choice test composed of both norm-referenced items and criterion-referenced items. It is purported to measure student mastery of curricular objectives and, with the "value-added" dimension, to evaluate teaching proficiency and efficiency. There is concern that current assessment constricts teachers' instructional strategies to "teaching the test"

(Snodgrass, Baker, Xu, & Detch, 1995, pp. 19-20) at a time when parents and community and business leaders are urging educators to teach higher-order thinking skills and to produce world-class learners. "Teaching to the test" is a concern of testing experts, according to a recent RAND report. The report goes on to say:

Experts warn that if assessments are used to advance policy objectives, particularly if they involve the imposition of rewards and sanctions based on test results, negative consequences are likely to result. These may include widening the gap in educational opportunities available to different kinds of students, a narrowing of the content and skills taught, a centralization of educational decisionmaking, and the deprofessionalization of teachers (Haertel, 1989; Airasian, 1987 in McDonnell, 1994, p. 1).

There appears to be some contradiction in the Tennessee legislators' decision to equalize funding in an effort to reduce the gap in educational opportunities available to students and, then, to use assessment to advance a policy for improved performance. The selection of policy tools and the use of assessment in their design may be at cross purposes with the legislative intent to improve student performance.

Improved performance depends, in part, on the feedback teachers receive from assessment. This feedback is critical in making changes in curricula and instruction to improve student learning performance. Educational measurement expert Robert L. Linn (1993) says, "to date, measurement has done a much better job of predicting who will achieve and of describing that achievement than of helping teachers adapt instruction to enhance the learning of individual students" (p. 9).

At best, the TCAP achievement test measures content and basic skills knowledge. To know what should be done and to perform the task competently are inseparable elements in learning and in assessing student learning. In pointing out the limitations of standardized assessments, Jaeger (1993) gives an example of how a multiple-choice format for predetermined content might measure very different academic performances for two students. Yet, in the eyes of educational policymakers, it may be interpreted as a measure of the same skill or information, reinforcing the perceived effectiveness of current policy mandates.

[A]n item on a history test about the causes of the U.S. Civil War might require rote recall from an examinee who had studied the period extensively, but it might require deductive logic from an examinee who had little exposure to U.S. history in the mid-1800s, despite a good background in economics and European history (p. 501).

From this example, correct answers from both students would indicate that the teaching was effective and that the testing format encourages learning. The standardized, multiple-choice format limits the performance and the evaluation of the performance. It does so by limiting the type and amount of information in the form of feedback to students, teachers, and administrators—those most directly involved in the adjustment of content and methodology for improved student performance. Understanding how legislators select policy tools and the use of assessment in the design of these tools reveals, in turn, the legislators' understanding of the teaching-learning process and the impact their decisions concerning assessment have on the evolution of this process.

Statement of the Problem

Limiting assessment to standardized measures of performance leads to the emergence of "teacher-proof curricula" (i.e. textbooks and teaching manuals with a prescribed scope and sequence for daily lessons) that may be seen by educational policymakers as the most efficient and effective way to distribute vast amounts of information to students. "The simplest way to sum up the potential harm of our current tests is to say that we are not preparing students for real, 'messy' uses of knowledge in context—the 'doing' of a subject" (Wiggins, 1993, p. 207). Whether students are being prepared for college or work, a prescribed scope and sequence for daily lessons often limits student participation to listening and responding to teacher comments and questions—rather than active problem-solving based on relevant content. In covering the prescribed curriculum that will be tested with a standardized instrument, the lecture mode of teaching often is perceived as the most efficient instructional method, again, limiting feedback about student performance.

If a predictable outcome is the single goal of assessment, then the fluctuations that naturally result from informational changes in the teaching-learning process are necessarily damped through the linear transformation of input/output interactions. In trying to maintain expectations represented by a linear model of performance, an unintended outcome is that the teaching-learning process is controlled to maintain the characteristic of predictable gain scores. This may mean isolating the process so that the interactions are limited as much as possible to data sets that "on average" are mastered at a given age or grade level as measured by a standardized test. This isolation may take place via a low-level information exchange between teacher and

student such as lecture, low-level thinking skills, or textbooks with simple concepts, grammar, and activities. This results in minimal exploration for increased understanding and personal or shared meaning.

On the other hand, if the purpose of assessment is to improve student performance, and if teachers are to be held accountable for that improvement, then it follows that assessment should be congruent with what students should know and be able to do and that teachers should be afforded the feedback from performance assessment to modify their instruction to facilitate such improvement.

The problem with the use of current assessment in holding teachers accountable for improved student performance is the limiting use of a linear model to measure the worth or merit of a nonlinear process—the teaching-learning process. If legislators continue to view the teacher or school effect as an isolatable, independent variable constituting a linear relationship with student performance on decontextualized measures of achievement, current assessment practice will constrain the evolutionary nature of the teaching-learning process by limiting the process to teaching the test.

Purpose

The general purpose of this study is to understand the impact of legislated assessment practices on student academic performance in Tennessee. Specifically, the study will examine the EIA Act of 1992 to determine the strengths and weaknesses of its design in terms of related theory, “best” educational practices, and its potential to advance policy objectives. As a result of the study, recommendations will be made to improve legislative policymaking as it affects educational practice. This study will offer answers to the following questions.

- How does the Tennessee State Legislature select a policy tool when making educational reform decisions? What role does assessment play in the design of the policy tool?
- What are the intended and unintended outcomes of legislated assessment practices in Tennessee?
- How can assessment as a component of the policy tool be developed to inform educational decision-making from the classroom to the legislative level?

Significance of Study

The current controversy over the use of standardized assessment for the purpose of measuring student improvement or teacher accountability in Tennessee may provide what Kingdon (in Lowham, 1995) calls a policy window. Policy windows occur "when three sets of thoughts or actions flowing through the bureaucracy, namely, the Problem Stream, the Alternate Stream, and the Political Stream, [coalesce]" (p. 109). The problem stream flows from the failure of current assessment practices to allow for the evolution of the nonlinear process of teaching and learning. The alternate stream is constituted by current grass roots restructuring efforts such as National Council of Teachers of Mathematics (NCTM) development of high standards for students and the authentic assessments to evaluate performance based on these standards. The political stream's source is found in the ongoing development of public policy based on current beliefs concerning the purpose of education as well as the purpose of educational assessment.

These streams coalesce in Tennessee with the State Board of Education's call for a review of the student assessment program. In response to a recommendation issued by the Comptroller of the Treasury, Office of Education Accountability (Snodgrass, Baker, Xu, & Detch, April 1995) and

lobbied for by the Tennessee Education Association, an independent panel was hired by the state to review the technical aspects of the value-added system. This independent review moves the focus of policymaking from questions concerned with the development and implementation of policies mandated by the EIA of 1992 to questions concerned with the selection of policy tools used to achieve policy goals and objectives. How do legislators select the policy tools they use to implement policy? What are the outcomes of these decisions? What purpose does and should assessment play in the design of these policy tools? According to Rist (1994), "few guiding data are available" to help policymakers in the selection of policy tools as "research to help policymakers in this regard is extremely sparse" (p. 553). This study will initiate the development of guidelines for careful selection of policy tools for educational improvement based on related theory and best practice.

Assumptions

The assumptions that form the basis of this study are drawn from the new scientific theories of chaos and complexity (Gleick, 1987; Lewin, 1992; Waldrop, 1992). The scientific ideas of chaos and complexity provide insights about the task of assessing classroom interactions. These insights challenge the current notions of causality—that events occur in a linear, predictable pattern in which the output of a given system is proportional to the input, or in the case of education, that student learning outcomes are proportional to teacher effect.

The studies of chaos and complexity present the complexity of biological and social systems not as closed systems or as the result of linear, predictable events arising from outside control. Rather, such systems are viewed as nonlinear, open systems in which "the complexity of ecosystems

and societies is the result of a network of transactions and communications between individual organisms" (Cohen and Stewart, 1994, p. 220).

There is evidence (Varela, 1979; Maturana and Varela, 1980; Priesmeyer, 1992; Senge, 1990; Sterling, 1992; Mainzer, 1994; Goldstein, 1994) to indicate that learning, thinking, and policymaking are nonlinear processes characterized by the tendency of a system to move in a certain pattern depending on the initial conditions of the system. These patterns are referred to as "attractor states" (Mainzer, 1994, p. 129-130). In the case of educational systems, attractor states can be thought of as the clustering of values that help shape particular world views that, in turn, become the context or culture for policymaking at the systemic levels of individuals and groups (Hall, 1995, p. 131).

Exposure to new ideas or unexpected problems in society often causes fluctuations in understanding and a mixing of theoretical perspectives. In education, we engage in heated debate—the clashing and changing of worldviews—over product versus process, standardized versus performance-based assessment, and mandated policies versus capacity-building policies. Oftentimes we retreat to the safety of repeating our past history or current mythology (Berliner & Biddle, 1995) because we are unsure how to measure the nonlinearity of the teaching-learning process to evaluate the impact of change in the educational context or change in the larger ecology. This may be due to policymaking based on a global theory that ignores local practices *or* based on local practice without regard for the theoretical connectedness of individual, social, and political contexts.

As theory and practice are inexorably linked, it is imperative that cross-scaling (Kugler & Shaw, 1990) be increased between theories and assumptions

within and among the multiple contexts of education. For it is with varied and increased interactions between theorists and practitioners that the organization of educational systems occurs at higher levels of order and the consequence of public education moves toward improved teaching and learning. What and how we evaluate the learning-teaching process will be, "within broad limits, to *pose* the relevant issues to be addressed at each moment of our life. They are not pre-given, but *enacted* or *brought forth* from a background, and what counts as relevant is what our common sense sanctions as such, always in a contextual way" (Varela, 1992, p. 250). What is *brought forth* will, as always, depend on our choices, which will depend on what we value within the parameters of the individual, social, and political contexts—and will be expressed in how we evaluate human performance. It may be the interdisciplinary theories of chaos and complexity that offer us new ways of seeing and knowing and choosing—new ways to resolve conflict brought about by change in the structure of education and its coupling to the larger ecology.

The discussion above leads to the following explicit assumptions of this study:

1. The teaching-learning process is a nonlinear versus a linear process.
2. Current linear assessment is limiting the use of best practices in Tennessee.
3. Policy tools are selected by Tennessee legislators based on past experience and the desire for predictable outcomes.
4. Current mandates are insufficient policy tools for the implementation of educational change as they do little to build capacity for schooling systems to organize at higher levels of order

based on the understanding and experience of those implementing the policy.

Limitations and Delimitations

This study will be focused on policy tools and policy decision-making procedures for education in the state of Tennessee. Findings and conclusions may or may not be generalizable to other contexts. The study will include a brief history of the selection of policy tools and the role of assessment to advance policy in Tennessee; however, since current assessment is a part of the EIA, the study will focus specifically on policy issues and design of the EIA.

The methodology for conducting this study is primarily a document analysis. Though it is an important source of information concerning policy decisions and decision making processes, there are some limitations to document analysis.

Program documents are subject to a variety of measurement errors. They may be incomplete or inaccurate. They may be selective in that only certain aspects of a program (that is, positive) are documented. Files and records are often highly variable in quality, with great detail in some cases and virtually nothing for other programmatic components (Patton, 1990, p. 245).

A variety of document sources will be used to validate and cross-check findings of the study.

The profession of educational evaluation under the direction of Dan Stufflebeam and the Joint Committee on Standards for Educational Evaluation (1981, 1994) has developed a set of standards that serve as guidelines for the development of evaluation practices. These standards will

set the parameters for the conduct and metaevaluation of this study concerning current and proposed assessment practices in Tennessee.

Methods and Procedures

Primarily, this study will be conducted as a document analysis (Denzin, 1978; Glaser & Strauss, 1967; Taylor, 1994; Hodder, 1994; Patton, 1990) to determine patterns in the selection and development of the policy tools used for planned educational change in Tennessee and the role of educational assessment in the process. Educational reports and records of student academic performance will be examined for trends in comparison to policy tool selection and development. A variety of document resources will be used to cross-check particular biases of any one source.

Documents to be collected will include federal, state, and local policies from the time period of 1965 to 1992 (ESEA, P.L. 94-142, CERA, EIA). In addition, minutes, memoranda, and reports from the State Legislature, State Board of Education, State Department of Education, Office of Education Accountability, Tennessee Education Association, Value-Added Research and Assessment Center, the Institute for Assessment and Evaluation, State Testing Office, the Commissioner's Assessment Task Force and the independent review panel, and newspaper accounts of assessment policy discussions will be analyzed to examine policy tools and their corresponding assessments along four dimensions within each of three policy cycles—policy formulation, policy implementation, and policy accountability (Rist, 1994).

The dimensions for the analysis are as follows:

- Resource intensiveness—the complexity of the operations, the relative costliness of different options, and the degree of administrative burden that different tools place on organization.
- Targeting—the capacity of the policy tool to be aimed at particular populations, problems, or institutions for whom the tool is primarily intended.
- Institutional capacity—the ability of the institution to deliver on the tasks that have been delegated to it.
- Political risk—concerns of unanticipated risk, chances of failure, timing, and sustainability of the policy initiative (Rist, 1994, pp. 554–555).

CHAPTER 2

A REVIEW OF RELATED LITERATURE

Historical Influences on Educational Policymaking in Tennessee

The Influence of First Policies

In the beginning, Tennessee did not readily embrace the ideas and principles of public education, owing its distrust to the ideas that education for all was dangerous and that education was a parental matter (Warf, no date). "Tennessee's first constitution (1796) had been silent on the question of education; thus, by implication it left the task of educating youth to private initiative by individuals or religious groups" (Warf, no date, p. 1). Changes in public distrust and apathy came about through the dedication of impassioned individuals winning the support of the people and pressuring the legislature to enact educational policy. It was the Act of 1873 that established an organizational framework for a system of public education in the state.

The act, in brief, provides for a state superintendent of public instruction, county superintendents, and district school directors with the state superintendent having central responsibility for insuring that the state school laws are faithfully executed. The legislation also stipulates that the superintendent see to the collection and dissemination of statistical and other information related to the schools (Warf, p. 8).

Since the Act of 1873, the policy tools selected for the implementation of educational change in Tennessee historically have been legislative mandates. This act was, in effect, a "reenactment" of the education mandate formulated by Gov. William G. 'Parson' Brownlow and his "carpetbag" legislature of 1867-68, repealed by the 1869-70 legislature of native Tennesseans (Holt, 1938, p. 3). In 1873, the dies were cast for additional "state school laws" and the beginnings of a state system for assessment of education.

Though educational improvement was made after the Act of 1873, it was not until the turn of the century that the barriers of public distrust of government intervention in educational matters and public apathy due to harsh economic conditions were broken. Through a united effort of educational forces to take the need for public education to the people of the state and thereby influence the lawmakers, legislation during the education campaigns of 1903–1913 resulted in the following principles:

- 1) The county, rather than the district, should be the unit of local school administration.
- 2) The state owes each child within its bounds the right to a common school education, and has the right to dictate such policies as are deemed necessary to guarantee this right.
- 3) The state should afford substantial financial aid to the public school system in all of its branches.
- 4) The purpose of state aid is to encourage and assist the counties in the support of schools, rather than to shift the burden of school support from the county to the state (Holt, 1938, pp. 257–259).

In 1913, the legislature approved a measure applying 33 1/3 percent of the state's general fund to public education. Because of the subsequent practice of "earmarking" funds for specific purposes and removing them from the general fund category, "the guarantee that financial prosperity in the state would always be accompanied by proportionate financial prosperity in the schools" was eliminated (Holt, 1938, p. 260).

Educational Funding as a Historic Issue

The practice of earmarking the state general fund for specific purposes coupled with the political activity of the educational forces in the state kept

the issue of funding a public system of education on the forefront of the legislature's political agenda.

During the period 1903–1913 the main purpose of state support for public schools was to serve as an incentive for better local support. Accordingly, each state school appropriation was accompanied by a provision which must be met by the locality. . . . During the period 1913–1922 the conception was rapidly developing that the state should assume a larger share of the burden of supporting the public schools (Holt, 1938, pp. 326–327).

One reason for increased interest in the state's fiscal responsibility was the need for a more equal distribution of funds across the counties of the State.

While the issue of equalization was addressed by the legislature as early as 1909, it was not until 1925 that "Tennessee's school programs for Grades 1–8 were funded on a proposed equalization concept" (Green, Smith, & Hydorn, 1995, p. 34). The issue of funding equalization across the state continued to fuel legislative debate and action, resulting in litigation and the passage of the Educational Improvement Act (EIA) of 1992. "For the first time in Tennessee history, the EIA provided for a separate fund for education revenue at the state level" (Green, Smith, & Hydorn, 1995, p. 36).

The Influence of Politics

Holt's analysis of the 1913–1923 period of education in Tennessee reveals a time of implementation of the laws enacted from 1903–1913. With implementation and interpretation of the school laws, subsequent legislation was enacted that would allow for the alliance of politicians, the State Board of Education, and the State Superintendent¹ (Holt, pp. 283–288). This trend was

¹ The title State Superintendent of Public Instruction was changed to Commissioner of Education in 1923 (Holt, p. 339).

to continue and to grow stronger through the period of 1923–1933 with legislative success for school programs dependent on the alignment of educational policymakers. Generally school programs were supported when aligned with the political faction that represented the majority and were defeated when the political faction represented the minority (Holt, p. 385).

In 1933, the state's educational associations including the State Teachers' Association, the Public School Officers Association, and the eastern, middle and western regional associations suffered politically when

the political faction to which the state's educational associations had tied their fortunes was publicly discredited and politically rejected. The existence of the associations and the cause for which they stood were threatened. In meeting the crisis, Tennessee's educational leaders effected a far-reaching reorganization. They created a single, unified state educational association, developed more effective campaign techniques, expanded association membership and function, employed a full-time executive-secretary, began publication of a monthly educational journal, began to democratize their procedures, and evolved as the Tennessee Education Association, one of the most powerful non-legal organizations of the state (Smith, 1952, pp. 7–8).

After the Act of 1873 and before the organization of the Tennessee Education Association (TEA), the organization that influenced legislative decision making in education was the Public School Officers Association, made up primarily of county superintendents. The State Teachers Association had lapsed into dormancy. Holt (1938) cites three possible reasons for this: 1) the 1873 Act was comprehensive, 2) the education bill compared favorably with those passed in other southern states, and 3) there was a hesitancy on the part of school administrators to push education onto the political platform (pp. 23–24). This period of relative satisfaction (1873–1902)

combined with convention programs designed to the interests of administrators showed a marked lack of participation on the part of classroom teachers (Smith, pp. 6-7). It was the Public School Officers Association, not the State Teachers Association, that was active in influencing the legislature during the educational campaigns of 1903-1913. With the reorganization and the creation of TEA, there was a unification of educational membership across the state. Holt (1938) describes the reorganized group as

a membership of over five thousand, composed of teachers and administrators from every level in the public school system and representing every county in the state. This group possessed not only the authority to express the will of the teachers, but also the power to impress upon vote-seeking politicians the importance of the group's will (p. 411).

That tradition of powerful political organization has continued to the present day. Former Governor (later United States Secretary of Education) "Lamar Alexander, expresses equal respect—if not admiration—for his main lobbying nemesis as governor, the Tennessee Education Association. 'It took me a year and a half of mustering all the forces I could muster to accomplish what I thought was a very simple idea,' said Alexander, recalling his battle with TEA over teacher merit pay" (Humphrey, Knoxville News-Sentinel, 1994).

The Impact of Federal Legislation on Educational Policymaking in Tennessee

Federal Aid Legislation

Education in Tennessee again suffered from inadequate funding during the Depression and World Wars. "In 1933, appropriations for higher education were cut 66-2/3 percent, secondary schools 28 percent, and elementary schools 19-1/2 percent. Throughout the period of Depression and

World War II, retrenchment or hold-the-line fiscal philosophy prevailed in state government" (Warf, no date, p. 19). Though the Tennessee General Assembly approved a resolution in 1945 for the "Tennessee congressional delegation to support the various bills regarding federal aid in the national Congress" (Warf, p. 24), the introduction of legislation for the purpose of federal aid for education failed until the passage of the Elementary and Secondary Education Act (ESEA) in 1965. Questions of aid to parochial schools and no aid to segregated schools had defeated a consensus for federal aid until the issue was changed to one of aid for the poor in the development of the Great Society (McLaughlin, 1975; Eidenberg & Morey, 1969). Now perceived as a solution to social inequity, federal aid came with "the requirement that educators be accountable for federal funds received under the Elementary and Secondary Education Act's (ESEA) Title I and III" (Worthen, 1995, p. 30) to ensure their destination. The development of a system of accountability was as heated a debate as that surrounding the appropriation of federal aid (McLaughlin, 1975).

Accountability for Federal Aid: Legislated Assessment Begins

While assessment of student performance began as early as 1845 in Boston to "evaluate the effectiveness of a school or instructional program" (Madaus, Stufflebeam, Scriven, 1994, p. 24), assessment as a system of accountability originated at the federal level and as a concern of New York Senator Robert Kennedy for political accountability and the possibility for self-organization at the local level (McLaughlin, 1975, p. vii).

Kennedy was concerned that the schools would not use Title I funds to develop special and successful programs for the disadvantaged, and hoped that project reports would provide

parents with the information (and thus the power) they needed to ensure that Title I dollars were spent in the effective interest of poor children (McLaughlin, 1975, p. vii).

Simultaneous with Kennedy's push for political accountability was the introduction of systems analysis to the Department of Health, Education, and Welfare by William Gorham (McLaughlin, 1975, p. viii). Gorham's expectation was that "data generated by Title I reporting system could be displayed in cost-benefit terms, thereby leading to more effective local practice and more efficient federal decision making" (McLaughlin, p. viii). The desire for cost efficiency won out in developing systems of accountability in education leading to educational assessment and a research paradigm that were counterproductive to the intended purpose of ESEA, that was to meet the needs of disadvantaged children (McLaughlin, 1975). The idea of information in systems thinking at the time was a reduction of disorder in a system through the use of feedback to correct error and "that problems ultimately reduce to the evaluation of the efficiency of alternative means for a designated set of objectives" (Ackoff in Checkland, 1995, p. 4). In the ESEA system of accountability, information about students' academic performance was measured using standardized tests. Academic performance or output was related to monetary input using production-function models of efficiency.

In the enactment of ESEA, a basic assumption of policymakers was that with more money and mandated evaluation, educators would know what to do with the money and new information from assessment to develop programs to meet the needs of disadvantaged children (McLaughlin, 1975).

An outcry from educators during the formulation and proposed implementation of such an accountability system warned policymakers of the following potential drawbacks. Those drawbacks included:

- lack of experience with the implementation of federal programs,
- no precedent for mandated evaluation,
- resistance to the perceived beginning of federal control—"first step in the federal prescription of a national curriculum" (McLaughlin, 1975, p. 9),
- limited knowledge of compensatory strategies in education,
- teaching to the test,
- deprofessionalism of teaching,
- a waste of scarce resources,
- incompatibility with best educational practice,
- validity of standardized tests were unproven (McLaughlin, 1975, pp. 8-11).

To calm the opposing schoolmen and meet the requirements of Senator Kennedy's criteria for accountability, the original assessment component of the ESEA was "loosely worded" resulting in a loose interpretation and implementation of evaluation procedures. This ultimately resulted in amendments to the act requiring studies of program results and effectiveness (McLaughlin, 1975, p. 17, 33). In the summary of her case study of ESEA, McLaughlin (1975) describes the legacy this federal legislation was to leave for future evaluation policy:

The history of Title I evaluation indicates that the success of evaluation—both its conduct and use—is primarily contingent upon the extent to which the evaluation effort is congruent with the dominant incentives at the local and state level. Furthermore, the experience of Title I suggests that the local

perception of federal initiatives and commitments as inherently unstable, combined with a basic defensiveness about achievement measures, will probably continue to frustrate federal attempts to secure objective, reliable program outcome measures. The Title I evaluation history demonstrates both how resistant the educational policy system is to assessment, and that a number of obstacles to objective evaluation are inherent in the system itself (p. vii).

In dealing with the resistance of the educational policy system, the political strategy for legislative decision making in education began as one of "compliance and assistance" in 1965, but changed to one of "compliance-dominant" with the implementation of ESEA (Elmore & McLaughlin, 1982). According to Elmore and McLaughlin (1982):

Partly by inertial drift, partly by intentional action . . . the federal posture toward states and localities has changed since 1965 from a relatively sophisticated mix of compliance and assistance objectives to a compliance-dominant strategy. The effectiveness of federal policy toward disadvantaged children has come to be judged increasingly in terms of the targeting and flow of federal funds—things which can be readily measured and monitored—and less in terms of the competence, innovativeness, and responsiveness of state and local educational organizations (p. 170).

Federal Legislation Coupled with Judicial Intervention

Concern for handicapped children was snagged by the tail hook of Great Society legislation—the concern for the equality of educational opportunity—with the formulation and passage of the Education for All Handicapped Children Act (EAHCA) in 1975 (Alexander, 1992, p. 359). In 1975, EAHCA, also known as Public Law 94-142 mandated that "children with disabilities could no longer be denied a free, appropriate education" (Hayes, 1992, p. 6).

Two important aspects of this federal legislation were: (1) the intervention of the courts in determining the legal right of handicapped children to an appropriate education and (2) the specified procedures for the implementation of the act at the local level monitored by the state education agencies (Elmore & McLaughlin, 1982). Procedural due process, individualized education plans, and periodic evaluations established by court decisions allowed the federal government to intervene in local education with maximal directives and minimal assistance. "[F]ederal policy was to be used to 'leverage' state and local resources toward handicapped students. The measure of success was to be compliance with procedural requirements" (Elmore & McLaughlin, 1982, pp. 170-171). P.L. 94-142 was not replete with suggested improvements to educational practice or theory in meeting the needs of handicapped children. These improvements would have to be invented at the local level under the procedural constraints of the law (Hargrove in Elmore & McLaughlin, 1982, p. 171). Echoes of the educators' rebuttal to the enactment of ESEA were heard with the passage of P.L. 94-142—renewed concern with an increased federal role in the states' rights regarding education and a lack of knowledge in meeting the needs of another disadvantaged group of students. Again, the effect was mandated change without consideration for the capacity of the institutions which would implement the change.

To implement P.L. 94-142, school systems began to include children with disabilities in the regular education programs of art, music and physical education while offering special education programs to remediate learning problems for academic areas. Special education training for teachers involved the study of specific handicapping conditions and the learning strategies

necessary for successful academic performance. This was different from the general education teacher training to teach for subject matter mastery. Special education training improved students' performance in the special education classroom, but fell short in preparing them for the pace and teaching styles of the general education classroom. By 1990 more than 4.5 million children nationwide were being served by special education programs.

In 1990 a new law (Individuals with Disabilities Education Act) emphasized that "to the maximum extent appropriate, handicapped children . . . are to be educated with children who are not handicapped"—an original premise of P.L. 94-142. Efforts to implement the new law emphasized the term "inclusion" and remain a part of the nationwide restructuring of education (Villa, 1992). The new law forced special education students back into regular education programs with consultation from special education teachers. The system of maintaining special education classrooms for some students and mainstreaming with consultation for others failed. Only 57 percent of the students in special education received a diploma or certificate of graduation, and only 49 percent of out-of-school youth aged 15–20 were employed between 1 and 2 years after high school (Hayes, p. 8).

In Tennessee and other states, current efforts in inclusion—the placement of special education students in regular classroom settings—go beyond consultation from the special education teachers. Special and general education teachers are co-teaching in the general education classrooms to provide services to all students. This is an attempt at the state and local level to meet the federal mandate for compliance with procedures for assessment and placement of handicapped children while implementing best practices based on current research, new educational theory, and practice.

A Federal Template for Educational Policymaking in Tennessee

By 1975, the context for educational policymaking was set by federal template for educational decision making at the state level—the enactment of compliance policies with inducements for compliance. In compliance with national educational policy, Tennessee has shaped and implemented educational change through a series of mandates holding teachers and local administrators accountable for the proposed learning outcomes of mandated change. Varying systems of evaluation were developed as part of these mandates. As federal evaluation mandates were withdrawn during the 1980s, state and local agencies continued to evaluate, but with their own purposes in mind—“to shape local or state decisions about their Title I or III projects” (Worthen, 1995, pp. 30-34).

In 1984, Governor Lamar Alexander worked with the Tennessee General Assembly to formulate and enact the Comprehensive Education Reform Act (CERA), also known as the Better Schools Program (Green, Smith, & Hydorn, 1995). CERA served as Tennessee’s response to the national concern for the decline of American test scores in comparison with those of economic competitors. Important components of this legislation were teacher pay based on performance, grade-level curricula and assessment tools developed for the mastery of basic skills, and an instructional model developed for the delivery of the established curriculum and as the basis for a teacher evaluation model. Tennessee took the compliance strategies of ESEA and P.L. 94-142 to a new level—past the state and local education agencies—to the classrooms. Procedures for compliance mirrored those of the federal mandates—compliance-based mandates with inducements for compliance and sanctions for noncompliance and the use of production-function models

to determine the cost efficiency and effectiveness of educational programming. With the enactment of CERA, procedures for teacher instruction and the evaluation of teacher instruction were developed and implemented.

In 1988, four years after CERA, 66 small school districts in Tennessee filed a suit against the state of Tennessee, alleging that the inequity of educational funding across districts violated students' constitutional right to equal educational opportunities. The State Legislature passed the Education Improvement Act (EIA) in 1992 before the final court decision was given on the case. "The EIA contained a section on school funding reform called the Basic Education Program" (Green, Smith & Hydorn, 1995, pp. 8-9). Linked to the funding reform of the EIA was a specific accountability system to track funding and monitor student gains known as the Tennessee Value-Added Assessment System (TVAAS). TVAAS was designed to be a more specific accountability system than that of CERA, using the Tennessee Comprehensive Assessment Program (TCAP) achievement test as the testing instrument for the TVAAS accountability system.

The TCAP achievement test is a standardized, multiple-choice test composed of both norm-referenced items and criterion-referenced items. It is purported to measure student mastery of curricular objectives and, with the "value-added" dimension, to evaluate teaching proficiency and efficiency. There is, however, concern that current assessment constricts teachers' instructional strategies to "teaching the test" (Snodgrass, Baker, Xu, & Detch, 1995, pp. 19-20) at a time when parents, community and business leaders are urging educators to teach higher-order thinking skills and to produce world-

class learners. "Teaching to the test" is a concern of testing experts according to a recent RAND report:

Experts warn that if assessments are used to advance policy objectives, particularly if they involve the imposition of rewards and sanctions based on test results, negative consequences are likely to result. These may include widening the gap in educational opportunities available to different kinds of students, a narrowing of the content and skills taught, a centralization of educational decisionmaking, and the deprofessionalization of teachers (Haertel, 1989; Airasian, 1987 in McDonnell, 1994, p. 1).

There appears to be some contradiction in the Tennessee legislators' decision to equalize funding in an effort to reduce the gap in educational opportunities available to students and, then, to use assessment to advance a policy for improved performance. Selection of policy tools and the use of assessment in their design may be at cross purposes with the legislative intent to improve student performance.

Improved educational performance depends, in part, on the feedback teachers receive from assessment. This feedback is critical in making changes in curricula and instruction to improve student learning performance. The standardized, multiple-choice format limits the performance and the evaluation of the performance. It does so by limiting the type and amount of information in the form of feedback to students, teachers, and administrators—those most directly involved in the adjustment of content and methodology for improved student performance as outlined in the state's curricular goals. Understanding how legislators select policy tools and the use of assessment in the design of these tools reveals, in turn, the legislators' understanding of the teaching-learning process and the impact assessment

has on the evolution of this process. Concerning the impact of who controls assessment, Linda Darling-Hammond (1993) says:

the question of who controls assessment is one of the major dimensions of the current debate about assessment reform. Even the most challenging and thought-provoking performance-based assessments will fail to transform schools if they are externally mandated and delivered. If some significant portion of the assessment process does not support teachers, students, and parents in their efforts to define themselves as a learning community, then the possibilities for organizational change and improvement will once again be wrested away from schools. The engine for school change—the catalyst for a community's political and educational development—will have been removed once again from the local school arena, where it must reside if it is to be effective (p. 760).

A Theoretical Perspective on Educational Policymaking

Problems Inherent with a Linear, Production-Function Model

The development of educational policy to improve student learning and encourage creative teaching will need to move away from the production-function model (Alexander, 1995) and toward a model “designed to support the programs and strategies required to meet [student performance] goals ” (Odden, 1992, p. 14). This requires alternate policy designs that provide resources to support human resource development [by] “focus[ing] resources holistically and build[ing] sufficient teacher capacity” to achieve student performance goals (Marsh and Sevilla, 1992, p. 125).

In his chapter on education production functions (input-output analysis), Alexander (1995) compares the application of production-function models in industrial and educational settings. He makes a compelling argument that the application of production-function models in their current states in educational settings “can be accurate only by chance. Its explanatory

power is problematic, and its predictive ability is without validation" (p. 352).

Alexander continues his argument against the continued use of current production-function models as the analysis from such models "is based on a standardized process derived from fully articulated and detailed production inputs and techniques that allow very little variation. . . [while on the other hand] education inputs and techniques gradually evolve as teachers reshape and modify instruction" (p. 356) creating a great deal of variance in inputs and techniques. This is not the only factor that calls into question the validity of a model based on linear, causal relationships between inputs and outputs. Since there is no exact measure for other educational inputs such as family background, peer influence, initial endowments, school factors, and achievement measures, variance in student achievement based on change in one or more of these inputs does not "represent causal relationships in education" (Alexander, 1995, p. 359).

Improvement in assessment models that track student improvement over time using multiple assessments, such as writing assessments, portfolios, and projects, not only allows but also encourages variance in the independent variables (or inputs) of educational settings. When comparisons between and among these different measures are made, patterns in student data are looked for in relation to curricular, instructional, or operational decision-making. As patterns in student data are discerned, qualitative investigations at the state, district, school, and classroom level would be necessary to explain fluctuations in the data. Current assessment that continues to view teacher or school effect as an isolatable, independent variable and constituting a linear relationship with student performance on decontextualized measures of achievement (Gardner, 1991) constrains the

evolutionary nature of the teaching-learning process by limiting the process to teaching the test.

[T]he insistence on controlling input so as to ensure the quality of output discourages the free exploration that results in disequilibrium. Without allowing disequilibrium . . . , creativity and the possibility for innovation is seriously hampered (Sterling, 1992, p. 160).

If the purpose of assessment is improved student performance and teachers are to be held accountable for that improvement, then, it makes sense that assessment should be congruent with what students should know and be able to do and teachers should be afforded the feedback from authentic assessment to modify their instruction to facilitate student improvement.

The current controversy over the use of standardized assessment for the purpose of measuring student improvement or teacher efficiency and proficiency may provide us with what Kingdon (in Lowham, 1995) calls a "policy window". This occur[s] when three sets of thoughts or actions flowing through the bureaucracy, namely, the Problem Stream, the Alternate Stream, and the Political Stream, [coalesce]" (p. 109). The problem stream flows from the failure of current assessment practices to allow for the evolution of teaching and learning as a dynamical process. The alternate stream is the current grass roots restructuring efforts such as National Council of Teachers of Mathematics (NCTM) development of high standards for students and the authentic assessments to measure performance on these standards. The political stream is the on-going debate concerning the purpose of education in the development of a better society. For the coalescence of these streams to occur, education needs a research methodology that seeks to explain rather

than predict. Such a methodology is on the horizon in the new sciences of chaos and complexity.

A Nonlinear Perspective on Assessing Schools as Social Systems

The scientific ideas of chaos and complexity (Gleick, 1987; Lewin, 1992; Waldrop, 1992) may give some insight into how to achieve the task of assessing classroom interactions, but will challenge current notions of causality—that events occur in a linear, predictable pattern in which the output of a given system is proportional to the input, or in the case of education, that teacher effect is proportional to student learning outcomes. The assessment issue viewed from the theoretical perspective of the sciences of complexity reveals evidence that the current assessment design based on a single instrument falls short in evaluating the dynamical processes of teaching and learning. Authenticity of assessment indicates a concern with evaluating the actual teaching-learning experience, that is, assessing the interactions within the processes of teaching and learning that facilitate understanding and meaning for those involved in the process.

Schooling systems, like other dynamical systems, “[are] connected by feedback to everything else” (Briggs, 1992, p. 59). Dynamical systems are “systems whose state changes over time. Systems such as these are multifaceted, complex, and interdependent” (Briggs, 1992, p. 15). In discussing psychometric problems of assessing change, Hanson (1994) says one of the most serious problems is

that most available assessment instruments were designed to measure static traits, rather than developmental changes (Hanson, 1982; Mines, 1985). Assessment instruments that measure static traits typically have high degrees of stability; that is, highly similar scores are obtained by students over time. In

addition, most measures of static traits assume that the underlying structure of the construct being measured does not change with time. Hence, these assessment instruments may not detect student characteristics that change (p. 588).

This might lead one to believe that assessing learning and teaching is a complex nonlinear process requiring measurement in which changes in the system become iterated, or recycled into the measurement process itself to determine the stability of the system (Bryk & Raudenbush, 1987).

The studies of chaos and complexity do not present the complexity of biological and social systems as closed systems or as the result of linear, predictable events arising from outside control. Rather, such systems are viewed as nonlinear, open systems in which "the complexity of ecosystems and societies is the result of a network of transactions and communications between individual organisms" (Cohen & Stewart, 1994, p. 220). These transactions and communications within and between living or open systems produce a phenomenon known as self-organization. In social systems,

(self-organization) occurs when political interactions and dialogues between members of a group produce coherent behavior, despite the absence of formal hierarchy within that group or authority imposed from outside it. . . . Most managers are not used to the idea that a system, that is, a set of interactions, can control itself. For them, control requires some individual to take charge (Stacey, 1992, p. 6).

The learning-teaching process as a nonlinear process is evident in the characteristics it shares with other recognized nonlinear biological and social processes.

There is evidence (Varela, 1979; Maturana & Varela, 1980; Priesmeyer, 1992; Senge, 1990; Sterling, 1992; Mainzer, 1994; Goldstein, 1994;) to indicate

that learning, thinking, and decision making are nonlinear processes characterized by attractor states or the tendency of a system to move in a certain pattern depending on the initial conditions of the system (Mainzer, 1994, p. 129–130). In the case of social systems such as schools and school systems, attractor states can be thought of as the clustering of values that help shape particular world views which, in turn, become the context or culture for decision making at the systemic levels of individuals and groups (Hall, 1995, p. 131). The context or culture is created by the interactions of individuals (micro-level) and the interactions of sub-systems (macro-level) in which shared values help create a “relational space” (Horn, 1995) over time. The stronger the attraction to central themes or common beliefs, the greater the resistance to new ideas emanating from within or from outside the given culture. Resistance to change within any given culture is the degree of attraction to a particular set of interactions believed to ensure the continuation of the culture and, thereby, the continuation of the interactions that define the individuals or the sub-system. Goldstein (1994) refers to this resistance to change as the system’s behavior in congruence “with what is possible at that phase in the [system’s] development” (p. 62). Goldstein (1994) describes this resistance to change as the development and continuation of equilibrium-seeking processes:

Equilibrium-seeking processes insure that all of the behavior in the system conforms to the attractor. They keep the system isolated from environmental variations that threaten to change the system’s operation, and they keep any creative impulse toward new behaviors in check. They mask nonlinear growth potential in the system so that the system appears to be a linear process. Furthermore, equilibrium-seeking processes strive to keep new information out of the system because it has the potential to disrupt current operations. In other words,

equilibrium-seeking processes serve to maintain stability at all costs (p. 66).

Increased or different interactions among the elements of the system or between systems often push the system away from equilibrium or the status quo and toward higher levels of complexity. "Changes in context then, signify a transition out of the equilibrium attractor altogether and into more complex attractors" (Goldstein, 1994, p. 66). It follows, then, that for systems to move beyond the concepts of predictability, centralization, and one-way causality "the far-from-equilibrium challenge must be designed to interrupt the interactions between beliefs, expectations, actions, and results" (Goldstein, 1994, p. 139) of current policymakers. In educational systems, the measurement or assessment applied to the implementation of educational policies creates the parameters for the values, the beliefs, expectations, actions, and results of the interactions within the system.

Ballantine (1993) presents not only the inner working of a classroom or school but all of education as an open system (See Figure 1) in an effort to help visualize the elements of an education system in interactive and complex sets of relationships.

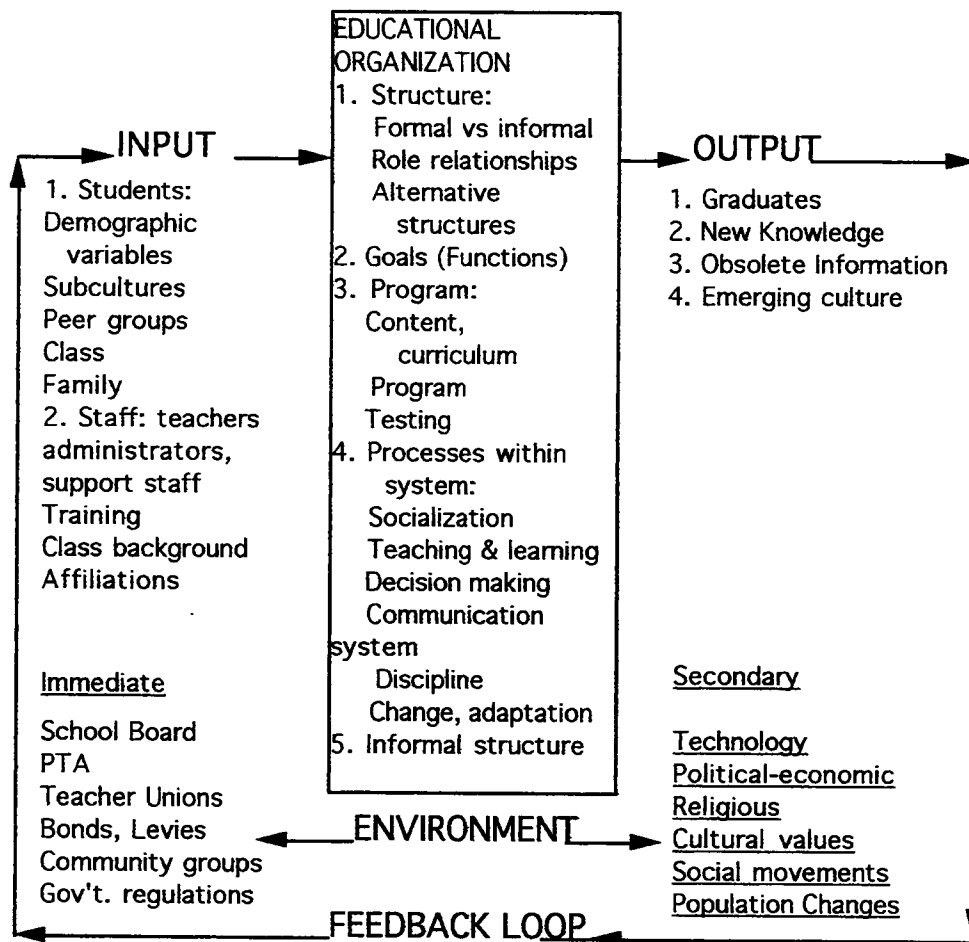


FIGURE 1 Systems model of education (Ballantine, 1993, p. 15).

Change in one element of the system affects the interactions and relationships throughout the system and calls for creative as well as adaptive behavior to integrate change while maintaining the organization of the system.

Lewin (1992) states in his book, *Complexity*, "If you have populations that interact, and their fitness depends on that interaction, you will see periods of stasis punctuated by periods of change" (p. 21). Change is characteristic of all living systems, both biological and social. Manageable change for growth and development is created in systems by using feedback, both negative and positive. Negative feedback occurs when the system is

monitored for error and the error corrected to maintain stable equilibrium. "Negative feedback is the type of feedback most common to social planning. Negative feedback holds a system within pre-specified boundaries" (Michaels, 1995, p. 23). Positive feedback occurs when there is an amplification of a fluctuation or change within the system (sometimes referred to as noise within the system) moving it away from equilibrium. "Positive feedback is the iteration of information. By iteration we mean that the output of the system is fed back into the system at the start of its next action" (Michaels, 1995, p. 24). "One of the most important discoveries of chaology has been that positive feedback can cause complex, even chaotic behavior concealed inside orderly systems to unfold, and that negative feedback can grow inside an otherwise chaotic system, suddenly organizing it and making it stable" (Briggs, 1992, p. 119). Stacey (1992) amplifies this premise when he states:

. . . . Scientists have discovered that when nonlinear feedback systems are pushed far from equilibrium, they follow a common sequence of steps in which they may move from one state, through chaos, to unpredictable states of new order. These new states of order are not comfortable equilibrium ones in which the system easily stays until disturbed again but dissipative structures that are difficult to sustain because they require continual inputs of energy if they are to survive. The common sequence of steps is as follows. At each point of transition, systems driven far from equilibrium move through patterns of instability in which previous symmetry or order is broken, thus confronting the system with choices at critical points. Through a process of spontaneous self-organization, a form of communication and cooperation among the components of the system, new order may be produced. Such phase transitions do not necessarily have predictable outcomes; the possibility is there in some cases for the system to make choices that produce new and unexpected behavior. Chance may be involved in moving from one state to another. This is definitely not a possibility in or near to equilibrium states, where outcomes are predictable (p. 82-83).

Both positive and negative feedback loops that carry information concerning the collective interactions of students-in-environments to all parts of the educational system would seem to be loosely connected due to the lack of political interaction and dialogue among all elements within the system and to elements of the larger social system. Given the educational tradition of social equilibrium, curriculum and instruction formulas have been linear functions of knowledge for an accumulation of facts necessary to maintain the status quo of the dominant social force rather than "information feeding back on itself and changing in the process" forming new patterns of order (Wheatley, 1994, p. 125). Without the possibility of new patterns of order, rigid parameters for the types and numbers of interactions within the educational system will constrain the evolvment to higher levels of complex thinking that would provide creative solutions to current societal problems.

The error in a linear approach is the assumption that the social interactions remain constant and that knowledge is based on a given reality rather than being continually created to solve current problems. Students bring nonlinearity into the curriculum and instruction equation with their different configurings of the world—emerging patterns of meaning based on the qualitative differences of personal insight, interactions, reflection and change in their thinking and/or actions. The close observation and description of students working on problems and issues in a variety of settings gives teachers more accurate insight into both teachers' and students' learning strategies. Through questions, recommended reading, mentoring, and curricula centered around problems and issues taken from students'

environments, learning algorithms can be expanded by the recursive feedback of student enaction².

This approach contrasts sharply with 'scientific' educational and psychology research which, for example, uses unidimensional numbers derived from quantitative measures of intelligence and home background to make linear predictions concerning similar numbers obtained from standardized tests of educational achievement (Cziko, 1989, p. 24).

Assessment and Changes in Learning: First-Order Change

The measurement of nonlinear processes at particular points in space and time may reveal different patterns of learning, thinking or decision-making as fixed point, limit cycle, or strange attractors³. For example, if learning is measured at a point when a student is gathering data to establish a mental repertoire—not only maintaining but also increasing factual knowledge, the pattern may appear as a fixed point attractor symbolized as a straight line—the number of years in school and the number of facts learned should be proportional. A senior in high school should have more factual knowledge in his repertoire than a kindergarten student. McWhinney (1992) describes this as first order change which "consists of reformations [over time]

² In his essay "Whence Perceptual Meaning? A Cartography of Current Ideas," Varela (1992) proposes "the greatest ability of all living cognition is, within broad limits, to *pose* the relevant issues to be addressed at each moment of our life. They are not pre-given, but *enacted* or *brought forth* from a background, and what counts as relevant is what our common sense sanctions as such, always in a contextual way" (p. 250).

³ Emerging patterns in the dynamical behavior of nonlinear systems are characterized by attractors (Cohen & Stewart, 1994, p. 207). Attractors come in three types: steady states (fixed point attractor), periodic cycles (limit cycle attractors) and nonperiodic cycles also known as "strange attractors" (Gleick, 1987). Attractors are the gravitational centers for systems exhibiting various levels of organization or development. A unique feature of biological and social systems is that they are characterized by all three attractor states, fixed point, limit cycle and strange.

that occur with no change in the meaning of the context. In studying learning, we are concerned with memorization and rote learning. . . ." (p. 58).

In assessment, if predictability of a linear system is the goal, then the fluctuations a nonlinear process exhibits would necessarily be damped and information ignored through linear transformation of input/output interactions. In trying to maintain expectations represented by a linear transformation, the nonlinear process is controlled to maintain the characteristic of predictability. This may mean isolating the process so that the interactions are limited as much as possible to data sets that "on average" are mastered at a given age-level as measured by a standardized test. This isolation may take place via a low-level information exchange between teacher and student such as lecture, low-level thinking skills, or textbooks with simple concepts, grammar, and activities—minimal exploration for increased understanding and personal or shared meaning.

This isolation of the learning-teaching process is antithetical to what Caine and Caine (1994) present as the acquisition of natural knowledge that "fundamentally changes the way our self is organized, including the way in which we perceive ourselves. In meaningful learning, the brain itself is reorganized, and that has consequences for the perceived self" (p. 137). These authors go on to say:

Many fundamental shifts, large and small, are generated within the learner as new knowledge links with what is familiar or meaningful. Moreover, these shifts are neither directed by the teacher nor controlled by the learner. Acquisition of natural knowledge is the result of the entire process. When immersion is orchestrated properly, including the fact that the learner is intrinsically motivated, the learner will have spontaneous and often unnoticeable shifts or flashes of insight—those "aha's" that represent felt meaning. The emergence of the "aha" is frequently preceded by periods of uncertainty or ambiguity and

hard, intrinsically motivated work requiring delay of other forms of gratification. The uncertainty is natural because the relationship between the information gleaned and parts of the self are constantly being rearranged, enlarged, and reorganized. Karl Pribram (1987) calls this aspect of learning active uncertainty. We suggest that this process is critical to helping the brain make maximum connections and that it can proceed smoothly only in an atmosphere of general safety and sufficient challenge (p. 137-138).

Within the atmosphere of uncertainty expressed through students' and teachers' questioning, interactions push the teaching-learning process beyond predictability. At best, teachers can provide a learning environment that encourages personal insight. However, the insights students have in answering questions about their environment and selves-within-that-environment are not proportional to the teachers' input in creating the learning environment. Likewise, the teaching of concepts and skills does not ensure a proportionate degree of student understanding or personal meaning. "Nonlinear systems are those in which there is an unmistakable influence of the parts on one another" (Goldstein, 1994, p. 111). That influence, however, is not necessarily proportional.

Second-Order Change in the Teaching-Learning Process

As students encounter organizational schemes to classify, label, or restructure existing concepts within their mental repertoires, they may experience some cognitive conflict, thus perturbing the equilibrium of their learning or thinking as it applies to their living day to day. Knowing and doing (Maturana & Varela, 1987) are out of balance. McWhinney (1992) refers to this imbalance as second-order change involving "the creation or change of a context" (p. 58) and having a range that includes "adapting a class of ideas to

include a new example, to forming a society, [within which] such changes participate in the basic processes of concept formation, adaptation, and propagation" (p. 59). If the student integrates the change of context into his existing repertoire relatively easily, his learning may return to the fixed point attractor or stable equilibrium. But if the change in context perturbs his learning sufficiently, a limit cycle attractor may characterize his learning in which competing ideas or concepts push and pull his stable reality.

Third-Order Change in the Teaching-Learning Process

When the learner's mental system is within a limit cycle attractor or oscillating among competing ideas and is perturbed further by additional insights, context changes, reflection, or increased interactions, it will be pushed far from equilibrium. This occurs when the learner joins a different subculture, invents a new theory, adapts his or her core values, or has "an experience in which a person suddenly realizes the arbitrary nature of his or her own paradigm" (Berman in McWhinney, 1992, p. 60). The learner, then, understands the rules for constructing a given reality but also the rules for constructing the rules of a given reality. What this allows the learner to do is enter the "space of the possible" (Cohen & Stewart, 1994). McWhinney calls this third-order change or a meta-praxis that "will give the freedom and space in which to choose the realities required for resolving conflict" (p. 61). To move into the space of the possible enables the learner to wander through numerous microstates of data to construct relevant problems and issues and therein understanding and meaning of the constructs. This process creates the parameters represented by a strange attractor—for "learning to learn" (Bateson, 1972). Within these parameters, the learner and the learning

context coevolve to higher levels of complexity in which new problems and issues are constructed and this, in turn, recreates the parameters for the learning. In this sense, learning is a self-organizing process—recursive in that values established in the learning are fed back into the learning process.

Measuring this process for any other reason than to map the learning in order to “describe the tracks and marks left by the passage of dynamical activity” (Briggs, 1992, p. 22) is to introduce external constraints on the learning-teaching process that are grounded in values from other contexts beyond the learner’s or the teacher’s. For the evolution of learning to reach higher levels of complexity, change and flexibility are required to maintain a balance between the apparent chaos of meaningless data and the emergent order created in the arrangement of data to construct meaningful living.

With this emphasis, “no variable is deliberately maximized and the ethics of the situation is ‘karmic,’ that is, it obeys a law of nonlinear cause and effect” (Berman, 1981, p. 257).

Köck (1980) describes cognition not as the “processing of well-defined objective ‘information’ from outside or inside, or its categorizing, analysis, storage and retrieval, but the permanent process of self-production through self-transformation, and the securing of the possibility of the process through adaptive interaction” (p. 100). Cognition as self-organized learning requires flexible parameters for the exchange of data. This exchange enables new connections within the brain that, in turn, may create new insights concerning previous knowledge and actions. To acknowledge the importance of flexible parameters for the creation of student understanding and personal meaning is to acknowledge the dependence on the initial conditions of the learning-teaching system. This, in turn, “requires that we break with

traditional statistical approaches that accommodate any concept of error” (Priesmeyer, 1992, p. 7) or constrain the exchange of data, as this limits the information available to the system for adaptive and creative interaction. As Priesmeyer (1992) explains:

We generally believe that this world operates with some degree of randomness; therefore, we have developed statistical tools that simplify activity into central themes and, thus, have discarded the inaccuracy of our computations as meaningless error. We compute means and regression lines to find the common themes. We then compute standard error and perform analysis of variance, never realizing that intricate patterns of behavior may exist in those “errors.” We even apply time series analysis and consolidate seasonal indexes, failing to recognize the functional relationships each observation has to its previous one. Worse yet, we focus on what we have explained and ignore the fact that we have left a large percentage of the behavior unexplained(p. 13).

With scores from standardized assessments, we are settling for predictable behavior instead of more complex patterns of learning and teaching because lessons become individual events to be imitated and measured rather than a continually recursive process of connecting new natural knowledge to previous understanding for increased understanding and meaningful change. Caine and Caine (1994) suggest four relevant indicators of performance and understanding that should be used to guide evaluation rather than using standardized tests only:

- The ability to use the language of the discipline or subject in complex situations and in social interaction.
- The ability to perform appropriately in unanticipated situations.
- The ability to solve real problems using the skills and concepts.
- The ability to show, explain, or teach the idea or skill to another person who has a real need to know (p. 166).

These indicators are more evident in writing assessments and portfolios assessments and are congruent with Romberg and Wilson's (1995) discussion of authentic assessment in mathematics. They make the argument that the development of assessment follows from the values of the educational system's culture that create a world view of learning and teaching:

If one considers mathematics to be a static, linearly ordered set of discrete facts, then the logical choice for a valid assessment system is the traditional standardized achievement test. On the other hand, if one views mathematics as a dynamic set of interconnected, humanly constructed ideas, then the assessment system must allow students to engage in rich activities that include problem solving, reasoning, communications, and making connections (p. 4).

In developing assessment models within a constructivist world view, "we can no longer use an atomistic model for assessment. We must assess the level of complexity of student understanding, not just the number of facts that students can pick out of a multiple-choice test" (Wilson in Romberg & Wilson, 1995, p. 6). If we view authentic assessment as contextual, it

must provide information on what the learner knows or does not know and on the developmental changes in such learning. Repeated measures of appropriate learning indicators must be made to obtain a robust picture of the learner's knowledge. . . . At least three different assessment mediums . . . ought to be used to obtain an integrated picture of the learner (Lajoie, 1995, p. 28).

In education, the policy-making context changes depending on the perceived purpose of schooling. As long as public education is defined as a linear process for the purpose of gathering facts rather than integrating knowledge for the purpose of creative living, schools will continue to participate in what Rifkin (1980) says is "little more than a twelve to sixteen

year training program for the Newtonian world view" placing emphasis on "quantities, distance, and location but rarely on qualities, relationships, or context" (p. 262).

The Evolution of Policy Design in Tennessee

Emergence from Tradition

Educational policy in Tennessee is designed from a tradition of legislated mandates with rewards for compliance and sanctions for non-compliance. This tradition began with the Act of 1873 to develop a public system of schooling in the state and was reinforced with federal legislation as the state became familiar with the legislated educational process through administering federal programs of ESEA and P.L. 94-142. McDonnell and Elmore (1987) have identified the assumptions underlying mandates as

(1) the required action is something all individuals and agencies should be expected to do, regardless of their differing capacities, and (2) the required actions would not occur, or would not occur with the frequency or consistency specified by the policy, in the absence of explicit prescription (p. 14).

These assumptions are consistent with the thinking of policymakers in 1873—that "the state owes each child within its bounds the right to a common school education, and has the right to dictate such policies as are deemed necessary to guarantee this right" (Holt, 1938, p. 258). These were also the basic assumptions of the federal government in 1965 and 1975 when policy for the education of the disadvantaged and the handicapped was enacted.

Inducements to achieve policy goals have been used by both the federal and state governments. The assumptions underlying inducements are:

(1) in the absence of additional money, one would not expect certain valued things to be produced, or to be produced with the frequency or consistency prescribed by policy, and (2) individuals and agencies vary in their ability to produce things of value, and the transfer of money is one way to elicit performance (McDonnell & Elmore, 1987, p. 15).

Initially, state aid in Tennessee was designed to stimulate local funding. As recorded by Holt (1938), "the purpose of state aid is to encourage and assist the counties in the support of schools, rather than to shift the burden of school support from the county to the state" (p. 259). ESEA and P.L. 94-142, on the other hand, carried inducements to the state to comply with federal rules and regulations for educating their targeted populations.

Assessment or the evaluation of educational programs is often the link between policy and practice, between mandates for programs and inducements for compliance with those programs. The door was opened for assessment with the passage of the Act of 1873—"the legislation also stipulates that the superintendent see to the collection and dissemination of statistical and other information related to the schools" (Warf, no date, p. 8). The collection of data to evaluate educational programs was strengthened with the assessment components of ESEA and P.L. 94-142. The original requirement for evaluation of educational programs was based on two assumptions: (1) that evaluation would be " 'self-executing'—that a requirement to establish procedures for the 'objective' measure of program 'effectiveness' would generate reliable, objective, and useful program reports" and (2) that evaluation would be " 'self-winding'—that the input of evaluation data about program activities and outcomes into the Title I policy

setting would lead to policy choices or program modifications based on these 'facts' " (McLaughlin, 1975, p. 11).

The idea of feedback was present in these assumptions but was ambiguous and based on the assumption that the capacity for "self-executing" and "self-winding" was present in the educational system. Inconclusive research about Title I projects, the influence of systems theory or management science (with its introduction to the idea of negative feedback), and an increased emphasis on standardized testing moved the potential of educational systems at the local level away from increased self-organization and toward a closed system in which the management of inputs and outputs would be closely monitored by outside agencies. Increased accountability for funds in compliance with specific procedures would define the use of assessment by 1975 with the passage of P.L. 94-142.

During the period of the 1960s to the early 1980s, the effect of standardized testing on instruction was an increased emphasis on basic skills in reading and mathematics—"recognition of facts, work analysis, mathematics computation skills, routine algorithmic problem solving" (O'Day & Smith, 1993, p. 257-258). To ensure success on these tests, teaching methods and textbook content were aligned with the tasks of the test (O'Day & Smith, 1993; Cohen, McLaughlin, & Talbert, 1993). At about the time the gap between disadvantaged student and middle-class student performance on standardized testing began to narrow, the winds of reform changed to stir public and policymakers' concern "about the relative underachievement of U. S. students and their lack of improvement over the past two decades" (O'Day & Smith, 1993, p. 258). Madaus (1988) cites this passage from the national report on education, *A Nation at Risk* (1983), to illustrate an

increased emphasis in using standardized testing to measure student competency:

Standardized tests of achievement . . . should be administered at major transition points from one level of schooling to another and particularly from high school to college or work. The purposes of these tests would be to: (a) certify the student's credentials; (b) identify the need for remedial intervention; and (c) identify the opportunity for advanced or accelerated work. The tests should be administered as part of a nationwide (but not Federal) system of State and local standardized tests (p. 111).

Standardized tests became the gauge for keeping the educational system operating in such a way that classrooms stayed within pre-set curricular and instructional limits, ensuring the stability of federal and state mandated policies.

With the enactment of educational policy since 1965, state and federal legislators have asked educators to improve the academic achievement of the poor, improve the academic achievement of the handicapped, ensure the global economic competitiveness of American students by developing higher-level thinking skills, and ensure equal opportunity for all children to learn. The measure of success in meeting these mandates was/is to be improved scores on standardized tests. The underlying economic production-function theory has "shaped and propelled the drive for accountability in education. . . Since there are no profits to serve as indicators of whether or not schools are doing a good job, test scores have been assigned that function instead" (Raywid in Berliner & Biddle, 1995, p. 194). The unintended outcome of such mandated policies that use standardized tests as the gauge for compliance has been to reduce "educational outcomes to testable products and deemphasize the role of the school in other areas, such as preparing young people for civic

participation, encouraging their personal development, and helping them master higher-level intellectual skills" (Raywid in Berliner & Biddle, 1995, p. 194).

New Possibilities

The question educational policy analysts (McDonnell, 1994; Fuhrman, 1993; Gifford & O'Connor, 1992; Odden, 1991; Cohen, McLaughlin, Talbert, 1993; & McDonnell & Elmore, 1987) are researching is: What might policymakers do to make "policy itself much more educational for educators?" (Cohen, McLaughlin, Talbert, 1993, p. 251) A beginning place for answering the question is:

To make such policies much more educational for educators, the processes in which policy is made and enacted would have to be opened up so that they created many more opportunities to learn. And those opportunities would have to be designed so that they embodied the sorts of teaching and learning that reformers wish to promote for classrooms (Cohen, McLaughlin, & Talbert, 1993, p. 251).

To accomplish such a policy design, policymakers would need to integrate capacity-building and systemic-change (McDonnell & Elmore, 1987) components into the development, implementation, assessment, and redesign cycles of educational policy. The professional development of teachers and the distribution of authority in the educational system have the potential to break the policy-making pattern of increased use and specificity of mandates. Mandates may narrow curricula and instructional strategies leading to poorer student performance that, in turn, leads to increased specificity of mandates. Building capacity through reform legislation would allow for the "mutual adaptation" (McLaughlin in Odden, 1991) and

"coevolverment" (Holland in Waldrop, 1992; Kauffman, 1995) of policymakers and those who implement the policy. Assessment as information for adaptive dynamical systems involving different levels of the policy system would consist of both macro-level analyses and micro-level analyses (McLaughlin, 1991; Senge, 1990; Dellar, 1992).

Macro analyses operate at the level of the system. They stress regularities of process and organizational structures as stable outlines of the policy process and frame individual action in terms of position in a relational network. Micro analyses, conversely, operate at the individual level. They interpret organizational action as the problematic and often unpredictable outcome of autonomous actors, motivated by self-interest. . . . Strategies for analysis and evaluation might become self-consciously multi-staged, developmental, and iterative, keying questions and methodologies to the point in the process under study, to the needs of the decision-makers, and establishing a regularized system of feedback to actors at all levels of the system. Policy analysis and evaluation thus becomes both as contingent and as regularized as the process it seeks to understand and inform (McLaughlin, 1991, p. 194-195).

The emergence of this type of "self-reflexive praxis—creating a process that would itself use the praxis to achieve implementation" (McWhinney, 1992, p. 243) will influence not only student learning but also the methods and content teachers use in their classrooms. This praxis will restore the dimensions of dialogue and negotiated meaning to learning environments.

If assessment is understood as the process of evaluating change over time in relation to the organization of an educational system, what standardized assessment practices within current policy design do is to ignore not only particular interactions within a learning context but also the structural coupling of the education system to its environment. Assessment that reveals the interaction of student and text only, and in only one way (i.e.

multiple-choice), actually gives more information about the tradition of assessment in the state than about the interactions of teachers and learners within a particular environment. Emphasizing standardization in the performance of the learner and reliability and validity in the evaluation of the performance may as Wolcott (1990) says distract researchers from the purpose in human research to “understand a social world we are continuously in the process of constructing” (p. 147). This study will initiate the development of guidelines for careful selection of policy tools for educational improvement based on related theory and best practice.

Re-thinking the Educational Policy Process

Kronenberg (1995) suggests adding a new metaphor in understanding the phases of the policy process, the cloud, and a new phase of the policy process, issue transformation. He discusses past metaphors for policy. The machine metaphor connotes control and mechanical equilibrium—policy becomes a tool to “serve our needs” and “can be evaluated in terms of standards we apply to machines: efficiency and effectiveness in their use of resources and other inputs” (p. 254). His critique of this metaphor is based on what is ignored in understanding policy as a machine—the complex processes of goal development, the dynamical nature of the learning context, and the idea of self-organized learning. The other popular metaphor for policy is organism, one that breathes life into the machine, and, thus, the locus of control is still viewed as outside the social system it affects. “The idea here is that the success (or survival) of a given organization is *contingent* on its ability to fit itself to the environment of other organizations and social structures in which it is embedded” (Kronenberg, 1995, p. 255). This idea

depends on the notion of a given reality which conflicts with the current theory of constructivism that implies policymakers impose their interpretations of reality on the policy process, shaping the policy environment while being shaped by it through numerous and varied interactions.

Finally, Kronenberg (1995) selects the cloud for its characteristics of fuzzy boundaries, multiple internal states, changeable patterns of shape and texture. The cloud metaphor allows for the conceptualization of positive feedback or what Kronenberg call "issue transformation" (p. 260) in the policy process. Positive feedback occurs when there is an amplification of a fluctuation or change within the system (sometimes referred to as noise within the system) moving it away from equilibrium. "Positive feedback is the iteration of information. By iteration we mean that the output of the system is fed back into the system at the start of its next action" (Michaels, 1995, p. 23). Issue transformation is:

the ill-bounded, formative set of behaviors and interactions that may lead to changes in the political support for current policy and to the consideration and emergence of support for other policy arguments and approaches, and even to redefinitions of the 'problem' to be solved (Kronenberg, 1995, p. 260).

If Tennessee legislators view educational policy as mechanistic, operating with negative feedback to control the behavior of teachers and administrators, for the intended output of improved scores on standardized measures, the possibility for the new solutions to current educational problems, such as the issue of accountability, will constrain the system within a tradition of mandates and inducements. What will be ignored are the possibilities for building teachers' capacity and creating systemic change that

would allow for the coevolution of teaching and learning as well as the coevolution of the educational and political systems of the state.

Educational Policymaking in Tennessee: A Legislative Perspective

TEA has been the only lobbyist and the strongest policymaker for education.

Special interest groups play a significant role. Lobbyists can provide useful information, since we can't be experts, but lobbyists are more and more controlling what happens in Nashville.

They [lobbyists] influence the hell out of us. They are a driving force that oils the machine. (Legislators' comments in Keese, 1990)

In her study of educational decision-making by the Tennessee State Legislature, Keese (1990) found that legislators made decisions concerning educational policy based on personal values and experience. "Of lesser concern is knowing what effects a law will have and listening to expert information" and of least importance to Tennessee Legislators was policymaking that was consistent "with existing law and procedure, maintaining established legal principles, or providing for orderly and manageable change" (Keese, 1990, p. 75). From interviews with the legislators, Keese concluded that legislators depend on the State Attorney General, staff lawyers, and the courts to make decisions about the legality of the policies they enact (p. 78).

In reference to influence groups on legislative policy, Keese stated that "participants identified educational lobbyists as the most important outside decision-making resources. Least important outside resources were college

and university representatives and governor" (p. 80). Though influenced by these outside groups, legislators "see themselves working with fellow legislators and staff in negotiations and study of an issue rather than reacting to demands from education agencies, lobbyists, or constituents" (p. 106). Inside resources most frequently identified as influential were "committee members, knowledgeable colleagues, and bill sponsor" (p. 82). The decision style most used to shape educational policy by legislators was the analysis of issues rather than bargaining (p. 105). Keese found one exception to these "patterns of decision orientations" for all groups of Tennessee Legislators according to their positions, party, experience, age, and region. "Contrasting with all other group rankings was the 50 or older group who chose a decision style of bargaining more often than one of analysis" (p. 107).

Though Keese found that "educational issues in a specific bill may affect the decision style or the decision mechanism legislators use" (p. 109), her study did not include discussion of legislators' design of policy instruments based on their values, experience, reference groups, or analysis of issues and, therefore, their underlying assumptions about teaching and learning. From her study, insight is given into how legislators making educational policy are shaped and how they intend to shape policy. What still needs to be understood is how legislative policy shapes educational practice that, in turn, shapes future legislative policy. The underlying assumptions of the educational policymakers are figural in defining the outcomes of current policy and in evaluating current policy in terms of best educational practice in comparison with current educational research.

Summary

Education was not an assumed state's right in Tennessee—no policy provisions were provided for in the state's first constitution. Education was promoted and campaigned for by dedicated and aggressive schoolmen. The passage of the Act of 1873 established a public system of education. Since that time, the priority for improvements in education wavered with economic conditions, public apathy, educators' involvement, and finally the alignment of educators and legislators. The issue of funding education in the state was and is forever in the forefront of the legislative agenda. With the introduction of federal aid and increased state funding to schools, mandated systems of accountability were developed and implemented. Linked to the latest funding reform effort of the legislature (EIA) was an accountability system to track funding and monitor student achievement gains known as TVAAS. TVAAS was designed to be a more specific accountability system than CERA, Governor Alexander's Education Improvement Act. The TCAP achievement test was the testing instrument for the TVAAS accountability system.

The problem with the use of current assessment in holding teachers accountable for improved student performance is the limiting use of its linear model to measure the worth or merit of a nonlinear process—the teaching-learning process. If legislators continue to view the teacher or school effect as an isolatable, independent variable constituting a linear relationship with student performance on decontextualized measures of achievement, current assessment practice will constrain the evolutionary nature of the teaching-learning process by limiting the process to teaching the test.

Policymakers share an opportunity to shape the parameters for interactions from which local communities may shape and sustain change. Parameters form the boundary conditions wherein school systems may evolve along a series of changes that are identified as most meaningful by the system itself. These parameters provide for interactivity and lead to adaptation and self-organization, both essential conditions for the continued transformation through changes that characterize the living social organization. The ability to make available the appropriate level of interactions and potential perturbations that conserve the adaptive and self-organizing capacities necessary for continued growth lie within the art and science of managing and facilitating this kind of self-transformative organization (Horn & Wilburn, 1994).

The sciences of complexity may offer a methodology for understanding the development of educational policy—from the politicians' choice of issues and the interactions involved in policy development to the complex process of policy implementation and assessment. The consequence will, as always, depend on our choices, which will depend on what we value within the parameters of the individual, social, and political contexts—and will be expressed in how we evaluate human performance. It may be the interdisciplinary theories of chaos and complexity within a "science of qualities" (Goodwin, 1994) that offer us new ways of designing policy—new ways to resolve conflict brought about by change in the structure of education and its coupling to a larger ecology.

CHAPTER 3

METHODS AND PROCEDURES

Introduction

The general purpose of this study is to understand the impact of legislated assessment practices on student academic performance in Tennessee. Specifically, the study will examine the Education Improvement Act (EIA) of 1992 to determine the strengths and weaknesses of its design in terms of related theory, “best” educational practices, and potential to advance policy objectives. As a result of the study, recommendations will be made to improve legislative policymaking as it affects educational practice. This study will offer answers to the following questions:

- How does the Tennessee State Legislature select a policy tool when making educational reform decisions? What role does assessment play in the design of the policy tool?
- What are the intended and unintended outcomes of legislated assessment practices in Tennessee?
- How can assessment as a component of the policy tool be developed to inform educational decision-making from the classroom to the legislative level?

The design for this study was based on the following conditions:

1. The nature of the research questions is exploratory and explanatory in nature, and they therefore, “deal with operational links needing to be traced over time, rather than mere frequencies or incidence” (Yin, 1994, p. 6).
2. The research questions are being asked about a contemporary event “over which the investigator has little or no control” (Yin, 1994, p. 9).

The conditions of this study met the primary requirements for doing a case study that utilized analyses of documents and records that preserve the

understandings and processes of educational policymakers in Tennessee, specifically the documents related to the development, implementation, and evaluation of the Educational Improvement Act (EIA) of 1992 and the use of assessment as part of that mandate's design.

Data Sources

To determine patterns or themes in the selection and development of the policy tools used for planned educational change in Tennessee and the role of educational assessment in the process, documents relevant to the policy making process were secured from the major policy making agencies as well as from archival records of student performance. A variety of document resources were used to cross-check particular biases of any one source. In addition, informal follow-up interviews with participants in the policy decision-making process were conducted as questions emerged from a thorough study of the documents.

Documents were collected from 1987 to 1996, reflecting a time period that began with initial reviews of the Comprehensive Education Reform Act of 1984 and ended with the outside evaluations of the Tennessee Value-Added Assessment System (TVAAS). The documents collected reflect the interactions of the major policy actors involved with EIA: House, Senate, and Joint Oversight Education Committees; House, Senate, and Joint Education Committees; the Governor; the Commissioner of Education; the Executive Director of the Tennessee State Board of Education; the Executive Secretary of the Tennessee Education Association; the State Department of Education; and the Tennessee Association of Business.

Techniques for Data Collection and Analysis

A Rationale for Qualitative Research

In both human language and human behavior, patterns emerge, whether on the scale of encounters or societies (Lofland, 1976). Varela (1979) explains these patterns emerging from both a tradition and from the constraints of our biological structure.

We have been raised in the midst of a tradition and with a biological structure that we cannot escape or pretend not to have. Thus we do have, by necessity, a world of shared regularities that we cannot alter at whim. In fact, the act of understanding is basically beyond our will, precisely because the autonomy of the social and biological system we are in goes beyond our skull, because our evolution makes us part of a social aggregate and a natural aggregate, which have an autonomy compatible with, but not reducible to, our autonomy as biological individuals (pp. 275–276).

As researchers endeavor to explain patterns of human behavior in social contexts such as education, research methods emerge from constraints and traditions also. In the best of Cartesian dualities, the debate over appropriate methods becomes the quantitative/qualitative distinction (Glassner and Moreno, 1989). Quantitative methods have evolved from the traditions and constraints of the physical sciences emphasizing an objective reality. Qualitative methods, on the other hand, have evolved from the social sciences emphasizing a socially constructed reality.

Wolcott (1990) says, "I do not go about trying to discover a ready-made world; rather, I seek to understand a social world we are continuously in the process of constructing" (p. 147). The probability of our understanding the possible meanings of human interaction at any one point in time are

determined by the constraints on our observational procedures and the values of our cultures. In terms of the new sciences of chaos and complexity, DePryck (1993) says:

My hypothesis is that probability can be described as the interaction between a range, or distribution, of entropic probabilities [a tendency toward disorder] and a range of constraining probabilities [a tendency toward order]. Together, they define a field of possible events, actions, or sentences that depend upon the specific level of complexity. The interaction between them at a specific moment in space-time, but continued through space-time, creates an interference pattern that functions as a "strange attractor," organizing the events at that level of complexity (p. 158).

It is at the edge of understanding where order and disorder exist simultaneously that the fractal dimension of the strange attractors we call meaning is evident. Its spiraling shape in levels of greater and greater complexity records the "tracks and marks left by the passage of dynamical activity" (Briggs, 1992, p. 22). In educational systems, the accountability measure or assessment applied to the implementation of educational policies creates the parameters for the interactions within the system. If educators and policymakers are to make meaningful changes in the policies that guide or constrain the dynamical activities of teaching and learning, understanding the consequences of policy decision making and feeding that knowledge back into the policy cycle is imperative.

The Role of Qualitative Research in the Policy Process

Policy making is multidimensional and multifaceted. . . . It is a process that evolves through cycles, with each cycle more or less bounded, more or less constrained by time, funds, political support, and other events. It is also a process that circles back on

itself, iterates the same decision issue time and again, and often does not come to closure. (Rist, 1994)

Of critical importance to the iterative nature of the policy process is the use of policy tools for the implementation and evaluation of policy objectives. The selection of policy tools reveals both the tradition or culture of the policy makers and the underlying assumptions about the issue to be resolved. If transformation of an issue is to occur and the parameters for political interactions are to be expanded, policy tools must allow for positive and negative feedback. As stated in Chapter 2 issue transformation is:

the ill-bounded, formative set of behaviors and interactions that may lead to changes in the political support for current policy and to the consideration and emergence of support for other policy arguments and approaches, and even to redefinitions of the 'problem' to be solved (Kronenberg, 1995, p. 260).

Redefinitions will occur when the consequences of policy decisions are fed back into the policy cycle in a way that interrupts the interactions of current policymakers by challenging their beliefs, expectations, actions, and perceived results in using particular policy tools to achieve their objectives. This perturbation, or interruption, to the tradition of policymaking occurs when the basic assumptions underlying the use of a particular policy tool are examined. To question basic assumptions, qualitative research might be viewed as having a reflective or what others (Janowitz, 1971; Patton, 1988; Weiss, 1988 in Rist, 1994) call an enlightenment function. Viewing policy research as reflective inquiry

suggests that policy researchers work with policy makers and their staffs over time to create a contextual understanding about an issue, build linkages that will exist over time, and strive constantly to educate about new developments and research findings in the area (Rist, 1994, p. 547)

According to Rist (1994), there is little research to guide policymakers in the selection of policy tools. "Research to help policy makers in this regard is extremely sparse. Policy makers decide either based on past experience with a tool. . . . or because they have a clear proclivity for or against a particular tool. . . ." (Rist, 1994, p. 553). By looking at the selected policy tools in relation to the targeted population, resource intensiveness, institutional capacity, and political risk, the facilitation or constraint on teaching and learning may be explored. Again, Rist (1994) indicates there is little research being done on this topic.

Qualitative research into the domain of institutional constraints and how it is that these constraints play out in the relation of the organization to the fulfillment of its mission is not now being done. . . . the issue of institutional capacity and what does or does not hinder the ability of the organization to achieve its stated objectives is important to understand explicitly. If policy makers establish the parameters around an organization to the degree that it can never clearly achieve its goal. . . ., then there is a built-in level of failure that ought not be ignored and for which the institutions should not be held accountable (p. 555).

The Iterative Process of Qualitative Research

Qualitative research as an iterative process (Dey, 1993) begins with some notion of connectedness, ensues with preparing questions to examine the context from the participants' points of view, involves the temporary suspension of the researcher's own set of beliefs, guides dialogue with the participants, and ends with emerging patterns grounded in the data collected. These emerging patterns are then re-examined or reflected upon both for integrity with the data and for possible reinterpretations. This iterative process requires data that are "valid, real, rich, deep, thick." These data are

collected primarily in the natural settings of those being observed, through in-depth interviewing and participant observation. As "contemporary practices are historically embedded" (Tuchman, 1994, p. 310), document analysis provides a "behind-the-scenes look at program processes and how those came into being" (Patton, 1990, p. 234).

Thus this epistemology of participation sees man in continuity with the natural world, in which knowledge comes into being in autonomous units through an interwoven mesh of frozen histories, like a castle of cards—structured, yet bootstrapping its content and solidity from within. Conversely it sees nature as human history, where every factual statement has a hermeneutic from which it derives and which contains its possibilities. The successor to objectivism is not subjectivism, by way of negation, but rather the full appreciation of participation. . . . (Varela, 1979, p. 276).

In analyzing data, the qualitative researcher uses an iterative process of interpretation and reinterpretation to reveal something akin to Salinger's "semantic geometry" of meaning as experienced in a particular context. Whether in an individual, social, or political context, semantic geometry contains the rules by which humans give their lives languaged shape—iterative rules that result in shapes as complex as those of trees, clouds, coastlines, and narratives. Through methods involving observation, dialogue, interpretation, and negotiation, the researcher hopes to understand, from the informants' points of view, the rules of interaction that give meaning to the behavior of an individual or a group. From the ground of possible interactions within the context under study, the researcher makes figural those rules describing a particular experience and its pattern of meaning. The researcher's questions serve as feedback that amplifies

fluctuations or distinctions in a person's or group's actions or thinking through the process of reflecting on a particular event.

Insight Leading to Interpretation

Insight guides the qualitative researcher not only to study a particular context, but also to study the research process itself. Insight allows for the recognition of new connections by the exploration of possible states of fit between the data and meaning. One particular element of the research context that becomes more figural for the qualitative researcher may blind him to other possible states of fit within the data. His or her questions may then become focused on the figural element to the exclusion of the connections among other elements within the research context. This is also an explanation for multiple realities, the phenomenon of figure-ground relationships based on how the researcher or knower constructs his or her study and, therefore, determines the possibilities of what he or she may know.

As an organizing concept of qualitative research, insight allows (1) for the creation of theory from the data and so theoretically can only be challenged by re-examination of the original data, thereby making theory more valid; (2) for the development of shared meaning—as the goal of the research is to sensitize concepts so that the participants, as well as the researcher, are able to “grasp the reference in terms of one's own experience” (Glaser & Strauss, 1967, p. 39); and (3) the “on-going process of fitting together the activities” (Blumer, 1969, p. 7) of human groups. Insight creates the possibility—not the probability—of finding a pattern in the interactions of all the elements of a particular context.

In deriving theory from the data, qualitative researchers do not claim generalizability as a major goal, but seek to understand interaction in a particular, natural setting. If a qualitative researcher's initial insight does not fit the data, all is not lost; for the ill-fit, the weak connection can be the basis for the next research question leading to new insights. Insight is not a controlled variable, but is a significant part of every research finding because the researcher is seeking meaning for himself as well as trying to understand meanings for, between, and among the participants of his study. As a qualitative researcher, "you are not putting together a puzzle, whose picture you already know. You are constructing a picture that takes shape as you collect and examine the parts" (Bogdan & Bilken, 1982, p. 29). When this constructivist process takes place through the interpretation of documents, the artifacts of dynamical activity, Varela (1979) echoes Gadamer (1960) in stating that:

There is no methodological escape from dealing with the elusiveness of understanding, and this makes it very evident that whatever is informative in a conversation is intrinsically codependent and interpretational. Whatever is said in order to fix and objectify the nature of a conversation's content is said from a perspective, from a tradition, and is always open to question, to revision, to disagreement. This is not failure or weakness, but the heart of the process (p. 269).

Qualitative research is self-similar to the process it is used to understand—interaction. In order to get to meaning, the research must take on the fractal shape of a tree or coastline, branching out for details, rather than comparing inputs and outputs in linear fashion. Along the fractal dimension, in which meaning is expressed by a semantic geometry, the shortest distance between the researcher's questions and understanding of

particular interactions is dialogue or the artifacts of dialogue. The shortest distance between understanding and meaning of those interactions is interpretation or the articulation of the intention of the actors.

Interpretation as Positive Feedback

The interpretive process creates friction to slow the "inertia of tradition" (Jaynes, 1976) by interfering with established meaning(s) in a particular context. The researcher's presence amplifies some existing interactions while ignoring others within the context under study. "The unexpected interference will start a chain of positive feedback, leading to a dramatic reorganization, a fundamental redistribution of the original probabilities" of meaning (DePryck, 1993, p. 160). Questions posed by the researcher determine the degree of interference with existing coherence among the participants' shared experience. New metaphors emerge to explain new patterns in the data, reshaping the framework for understanding through interpretation.

Theories take root in the attempt to understand or explain, and without metaphor, there would be no models or theories. Models and theories. . . . become more familiar and less rarefied when seen as part and parcel of the universal attempt to transfer meaning from the familiar and comprehensible to the remote and perplexing. Metaphors that evolve into models or theories serve not only to direct research by creating a symbolic language that provides the framework for the collection and interpretation of data, but as a way of isolating a dimension of the question to be examined that is not visible without the aid of the metaphor (Kliebard, 1992).

Interpretation of interactions rather than the description of the interactions creates the asymmetry of meaning-making within the experienced social or natural aggregate. As researchers reinterpret the role of

schooling in the next century, parameters of classrooms, and the pathways of knowledge, assessment plays the pivotal role in retaining relevant frames of reference from the past while establishing new metaphors in emerging education theory. In fine-tuning the tools with which we evaluate our knowing, we may "move from an abstract to a fully embodied view of knowledge," in which "facts and values become inseparable. To know is to evaluate through our living, in a creative circularity" (Varela, 1992, p. 260).

Rationale for the Case Study as the Research Design

....[C]ase studies feature: descriptions that are complex, holistic, and involving a myriad of not highly isolated variables; data that are likely to be gathered at least partly by personalistic observation; and a writing style that is informal, perhaps narrative, possibly with verbatim quotation, illustration, and even allusion and metaphor. Comparisons are implicit rather than explicit. Themes and hypotheses may be important, but they remain subordinate to the understanding of the case (Stake, 1978, p. 7).

The theoretical propositions that will guide the data collection and analysis have been discussed at length in Chapter 2. Based on Yin's (1994) definition of a case study, it is an appropriate design for researching the EIA. His definition is as follows:

1. A case study is an empirical inquiry that
 - investigates a contemporary phenomenon within its real-life context, especially when
 - the boundaries between phenomenon and context are not clearly evident.
2. The case study inquiry
 - copes with the technically distinctive situation in which there will be many more variables of interest than data points, and as one result

- relies on multiple sources of evidence, with data needing to converge in a triangulating fashion, and as another result
- benefits from the prior development of theoretical propositions to guide data collection and analysis (p. 13).

The case study is an appropriate design for this study because an understanding of a contemporary piece of educational legislation, the EIA of 1992, within a historical context forms the boundaries of the study. This piece of legislation is a unique phenomenon for a case study for two reasons. 1) The EIA is an evolutionary piece of legislation born of a distinct tradition and bound by its assumptions of teaching and learning. 2) It is a controversial piece of legislation because of its specific design and use of student assessment for system, school, and teacher accountability. The case study design allows for the exploration and possible explanation of multiple units of analysis (Yin, 1994). In this study, for example, the primary unit of analysis is the EIA, but other units of analysis embedded in the case are policy tool selection, legislated assessment policy, intended outcomes of policy, and unintended outcomes of policy.

The use of multiple sources of data is a characteristic of the case study well suited to the design of this study. The documents used in this study record the dialogue, decisions, and actions of the major policy actors throughout the policy process. Therefore, a need arises to triangulate the data for comparison of similarities and differences in each agency's records or "agency perspectives" (Denzin, 1978) of interactions, concerns, intentions, assumptions and understanding about the issue of educational improvement. "Observers must realize that the reality of any archival producing agency is a reality representing that agency's interpretations of

what has occurred, what is to be reported, and what is to be saved for future generations" (Denzin, 1978, p. 265). Through collection and analysis of documentary evidence and archival records, "the potential problems of construct validity also can be addressed, because the multiple sources of evidence essentially provide multiple measures of the same phenomena" (Yin, 1994, p. 92).

As the EIA is a legislated mandate, the researcher began collecting documents and records at the Tennessee State Library and Archives (TSLA) with the assistance of the legislative history librarians. Audio tapes of the House, Senate, and Joint Education Oversight Committee meetings (1987-1992) as well as the House, Senate, and Joint Education Committee meetings (1987-1992) provided the testimony of key policy actors before the committees, questions and answers between those giving testimony and individual legislators, and discussion of relevant issues between legislators and between the Senate and House committees. These tapes recorded, verbatim, the interactions taking place within the committee meetings. The audio tapes were also a source for identifying other documents of importance such as the Governor's papers, amendments to the education bill, commissioned reports, and handouts from those giving testimony—these documents were discussed by those being tape recorded. The tapes were previewed at the TSLA and those containing discussion of teaching and learning, accountability, and assessment were selected and fully transcribed. The TSLA was also a source for news releases, memos, letters of correspondence, speeches, and reports of major policy agencies relevant to the enactment of the EIA. As these documents are part of Tennessee's historical record, they could not be

removed from the premises of TSLA, and, therefore, were read and annotated at the library.

Data collection proceeded by securing entry into each of the major agencies involved in the development of the EIA. Phone calls to the records managers of each agency were made to request assistance and permission to review relevant documents. These calls were made in advance of the researcher's visit to each site in order to 1) establish a communication link with the records manager, 2) inform key personnel of the purpose of the research, and 3) prearrange assistance from the records manager to ensure a thorough collection process. Records managers were asked to supply the researcher with the agency's minutes, reports, and correspondence from the years 1987 to 1992. These documents were read and notes taken on discussion and actions relevant to the education act.

Records managers in all cases were knowledgeable about the EIA and the records management system, and were extremely helpful in providing all documents requested—on several occasions providing documents that had not been requested but were later deemed relevant to the study. In addition to the TSLA, the agencies visited were the Tennessee State Board of Education, the Tennessee Department of Education, the Tennessee Education Association, the Tennessee State Testing and Evaluation Center, the Value-Added Research and Assessment Center, and the Comptroller of the Treasury Office (Office of Education Accountability).

In addition to the documents and records of the major policy agencies, reports from the Tennessee Advisory Commission on Intergovernmental Relations, electric journal articles concerning the debate over TVAAS, and more than 275 newspaper articles from the Nashville *Tennessean* and the

Nashville *Banner*, both newspapers of record in the state capital, were available for review and annotating at the University of Tennessee–Knoxville. These articles (from 1987–1992) provided a public review of the policymaking process resulting in the EIA. Documents related to the *Tennessee Small School Systems v McWherter* lawsuit were obtained from Lewis Donelson, attorney for the plaintiffs in the *Tennessee Small Schools Systems v McWherter*, and the University of Tennessee–Knoxville law library.

Annotation involved the transcription and researcher's impressions of recorded dialogue or printed material that could not be copied or borrowed from the agencies visited. Annotations were related to the topics of public apathy, past legislated policy, legal considerations, educational funding, educational accountability, teaching, learning, assessment, policy tools, expected policy outcomes—all topics that “either connected with previous literature or suggested patterns emerging from the analysis of previous data” (Marshall & Rossman, 1989, p. 112). Notes and documents were designated and filed by the name of the agency from which they were obtained and subfiled according to topic. Filing was fine-tuned throughout the data collection and data analysis process.

By collecting data from multiple agencies to the point of redundancy (the emergence of patterns or themes), the researcher was able to develop a holistic description of the complex policymaking process resulting in the EIA. Through the data available in the documents and records of each agency, the words and deeds representing the process as it happened from 1987 to 1992 were preserved, therefore preserving the relationships of the variables not so easily isolated in the development and implementation of the education act.

It was from this holistic description that the researcher began the interpretive process of data analysis.

Rationale for Document Analysis as a Method of Data Collection and Analysis

. . . .[C]ontemporary practices are historically imbedded. To use today's theoretical concepts: Social meanings are recursive (Giddens, 1984). The past continues to speak to the present. All that we take for granted as "natural" is a product of both historical and contemporary processes. Our task as social scientists is to interpret those multifaceted meanings, including their interactions with one another. How to choose a point of view that serves as an interpretive frame is quite another matter (Tuchman, 1994, p. 310).

After the enactment of the EIA of 1992, the greatest resistance to changes made in the educational system was to the use of the TVAAS in measuring district, school, and teacher effect on student academic gain. Goldstein (1994) refers to this resistance to change as the system's behavior in congruence "with what is possible at that phase in the [system's] development" (p. 62). This resistance raises some questions. What had been the underlying assumptions concerning the educational system that lead policymakers to the particular design of the EIA? Why was this particular design met with such resistance on the part of educators? Were the changes specified in the EIA incongruent with the possibilities for the systems to change?

At first glance, the answers to these questions seem simple to locate: ask the participants. However, the answers for policymakers would most likely be in defense of their policies; for teachers, the answers would most likely be in defense of their current practices. From a historical perspective, the use of a production-function model by legislators for resource allocation

in proportion to expected student performance is repeated from Elementary and Secondary Education Act (ESEA) to P.L. 94-142 to Comprehensive Education Reform Act (CERA) to EIA. The drawbacks to the various configurations of the production-function model have been forewarned repeatedly as well. Currently, a task force under the direction of Tennessee Education Commissioner Jane Walters is reviewing the entire state testing system, a component of the accountability system that is the heart of the EIA (Mayshark, *KNS*, May, 5, 1996). Before history repeats itself and a more specific accountability system is devised or, perhaps, abolished, research must serve a reflective or enlightenment function—to “create a contextual understanding” (Rist, 1994, p. 547) of the assessment and accountability issues and their relation to the nonlinear process of teaching and learning. To choose the interpretive frame for doing this research is quite another matter (Tuchman, 1994).

The process and dialogue of the interactions leading to the enactment of the EIA are embedded in the documents and records of the different policy agencies or actors involved. “Such texts are of importance for qualitative research because, in general terms, access can be easy and low cost, because the information provided may differ from and may not be available in spoken form, and because texts endure and thus give historical insight” (Hodder, 1994, p. 393). According to Reinharz (1992), cultural artifacts, such as documents and records, “have two distinctive properties. First, they possess a naturalistic, ‘found’ quality because they are not created for the purpose of study. Second, . . . [c]ultural artifacts are not affected by the process of studying them as people typically are” (p. 147). In the case of the EIA, the “concrete texts differ from the abstract structures of language in that they are

written to do something. . . . [and are] embedded within social and ideological systems" (Hodder, 1994, p. 394). In this study, the documents and records were used to interpret the meaning of legislative intent in light of current resistance to the assessment and accountability measures.

Through the analysis of the documents, the researcher became a nonintrusive participant-observer of the dynamical activity of policy actors in the natural setting and time of their action. Documents and records revealed the design process of policy tools used to resolve the educational problems of inadequate funding and poor student performance. Such issues involved the worldviews of lawmakers, judges, business leaders, teachers, parents, students, and others. "The study of material culture [texts and artifacts] is thus of importance for the qualitative researchers who wish to explore multiple and conflicting voices, differing and interacting interpretations" (Hodder, 1994, p. 395).

The discontinuity between the policy actors' realities is revealed as the researcher identifies emerging patterns of actions and their resulting rules of interaction. The patterns of actions and rules of interaction within their realities are not known, necessarily, to the participants at the time of their occurrence. To bring to consciousness how the assumptions of one reality become the constraints of a second reality, "the interpreter of material culture works between past and present or between different examples of material culture, making analogies between them" (Hodder, 1994, p. 398). Changes in the rules of interaction create discontinuity in existing policy design and problem solving. The researcher must identify the resulting conflict created from the simultaneous existence of past and present policies. McWhinney (1992) recognizes that the "basic mechanisms of problem solving create

changes in a given reality according to the rules of a second reality" (p. vi). How do the new rules of intended policy affect the given or current reality of those responsible for implementing policy? It is the rules of the second or intended reality that the researcher must understand and interpret if the policy process is to become iterative and not simply repetitive. Tuchman (1994) summarizes the researcher's use of document analysis as follows:

. . . . the assumptions of an era (an *époque*) are both inscribed and embedded in (documentary or lived) texts. Texts are to be analyzed as parts of webs or systems of signification that may be viewed as "a set of language systems." Because language systems are characteristic of an era (place, class, or situation), one can analyze any particular text in relationship to other texts; that is, as part of a *structure* of meaning. Indeed, the analyst's (researcher's) task is to elucidate that structure (p. 315).

The concepts representing the properties and dimensions (Strauss & Corbin, 1990) of policy actions over time and in relation to the EIA, the relationships between and among these concepts, and the interpretation of the relationships form the structure of meaning in this study. To elucidate that structure the researcher used a specific analysis strategy of "process coding" within the general strategy of "open coding" (Strauss & Corbin, 1990).

While reading and annotating the data collected, the researcher began the process of "open coding: the process of breaking down, examining, comparing, conceptualizing, and categorizing data" (Strauss & Corbin, 1990, p. 61). The answers to researcher's initial questions about the phenomenon, the EIA, provided the initial labels for concepts that describe the properties and dimensions of the phenomenon. Strauss & Corbin (1990) describe this process as "taking apart an observation, a sentence, a paragraph, and giving each discrete incident, idea, or event, a name, something that stands for or

represents a phenomenon" (p. 63). For example, one of the initial research questions was: "How does the Tennessee State Legislature select a policy tool when making educational reform decisions?" One concept to be explored was policy tool. Using McDonnell & Elmore's (1987) classifications of mandates, inducements, capacity-building, and system-changing, the researcher labeled the EIA as a mandate. Possible properties of any policy tool are as follows:

- Resource intensiveness—the complexity of the operations, the relative costliness of different options, and the degree of administrative burden that different tools place on the organization,
- Targeting—the capacity of the policy tool to be aimed at particular populations, problems, or institutions for whom the tool is primarily intended,
- Institutional capacity—the ability of the institution to deliver on the tasks that have been delegated to it, and
- Political risk—concerns of unanticipated risk, chances of failure, timing, . . . and sustainability of the policy initiative (Rist, 1994, pp. 554–555).

Graphically the pattern for this example of open-coding looked as follows:

Concept	Property	Dimensional Range
Policy tool	Resource intensiveness	Complex-----simple High cost-----low cost More burden--less burden
	Targeting	High capacity---low capacity
	Institutional capacity	High ability----low ability
	Political risk	High risk-----low risk

As each discrete part was labeled and relabeled, it was fitted with other discrete parts for comparison. As these parts were compared for similarities or differences, they were grouped or categorized according to properties and dimensions of those properties (Strauss & Corbin, 1990). How the researcher labeled the properties and dimensions elucidated different patterns in the data—different labels yielded different categories, in turn, yielding different structures of meaning. It was, therefore, essential that the researcher reread and requestion the data to the point of redundancy—to the point that new concepts or categories developed throughout the analysis process either fit within or called for the collapse of existing concepts and categories. It was at this point that the categories of data were compared and combined into major categories and moved closer to the structure of meaning embedded in this particular context.

The understanding derived from the development of categories informed the researcher's new questions, amplifying some distinctions over others in the data. According to Dey (1993):

Categorization of the data requires a dialectic between categories and data. Generating and developing categories is a process in which one moves backwards and forwards between the two. It is this interaction of category and data which is crucial to the generation of a category set (p. 117–118).

Moving back and forth between categories and data organized the data in increasing levels of complex relations, moving the data analysis from an insightful description of context and intention to qualitative “in-formation” (Varela, 1979) concerning action and consequence. This process of moving from description to in-formation or interpretation occurs through the linking

of categories. At this point in the study, the researcher used process coding to continue the analysis.

Process coding is a particular way of linking categories of data into "sequences of action/interaction as they pertain to the management of , control over, or response to, a phenomenon" (Strauss & Corbin, 1990, p. 143). Process coding was an appropriate linking scheme for this study as it was concerned with a) conditional changes in the use of assessment in legislated educational policy over time, b) the reaction of those implementing policy, c) the consequences that resulted from the changes in policy, and d) "how those consequences became part of the conditions influencing the next action/interactional sequence" (Strauss & Corbin, 1990, p. 143). Linking categories as a process gave the researcher a basis for the interpretation of interactions.

Interpretation of interactions rather than only the description of interactions allows for the asymmetry of meaning-making within the social or natural aggregate of experience. Interpretation connotes Maturana's question: What is the process generating the observed phenomena, rather than, what is the phenomenon as an entity or characteristic? (Mingers, 1995, p. 76). Description connotes an observer but not the observing, whereas interpretation connotes the involvement of the observer enacting the description—providing the coherence or regularity between the description and the history of interactions being described. If the result of a qualitative study is the naming or the description of the interactions , there is a danger of reifying present interactions without understanding the "structural coupling" (Maturana and Varela, 1980) of the phenomenon to its context. It is a unique

connectedness Maturana and Varela (1980) refer to when referring to structural coupling:

In the history of interactions of a composite unity [phenomenon] in its medium [context], both unity and medium operate in each interaction as independent systems that, by triggering in each other a structural change, select in each other a structural change (p. xx).

In 1987, Maturana and Varela modified this definition to read: "We speak of structural coupling whenever there is a history of recurrent interactions leading to the structural congruence between two (or more) systems" (p. 75). It is the structural coupling that allows the system to transform itself to higher levels of complexity in response to change while maintaining its identity as a system. The meaning of change in interactions triggered by perturbations from the environment or from within the system is without validation unless this structural coupling is understood.

The structural coupling in this study lies in the connections between legislated policy and the teaching-learning process, the connections between the policy design and policy implementation, the connections between the policy implementation and policy accountability, and the connections between policy accountability and the transformation of policy issues. These connections represent the interactions involved in the structural coupling of evolving political systems to evolving systems of education. For example, policymakers select policy tools based on assumptions about an issue formed through tradition and experience. The policy tool selected, whether a mandate with incentives and sanctions or a system-changing policy, will have an impact on the curricula and instructional techniques a teacher uses in the classroom, in turn, having an impact on student learning. The

improvement or lack of improvement in student learning determines whether or not teachers and administrators are doing their jobs. The degree and quality of student learning in relation to the expectations of policymakers will have an impact on future policy tool selection to achieve expected outcomes.

Using a process coding within an open coding strategy to analyze the data, the researcher was able to interpret the meaning of policymakers' interactions historically embedded in the documents and records from the different policymaking agencies. The actual analysis of the data is the content of Chapter 4.

A Rationale for Data Triangulation

Though documents are an important source of information concerning policymaking processes, there are some limitations to document analysis. Patton (1990) delineates these as follows:

Program documents are subject to a variety of measurement errors. They may be incomplete or inaccurate. They may be selective in that only certain aspects of a program (that is, positive) are documented. Files and records are often highly variable in quality, with great detail in some cases and virtually nothing for other programmatic components (p. 245).

To corroborate the evidence and interpretation of the legislative proceedings and to insure a thorough investigation of the data, the researcher used a data triangulation strategy. First, data were analyzed from the audio-taped legislative meetings in which each agency was represented and gave testimony. Concepts, categories, and relationships between categories were determined. Second, separate data analyses were completed for each political

agency involved in the production and implementation of the EIA. Each separate analysis was based on data collected from the minutes of meetings, reports, records, and correspondence issued by individual agencies such as the Tennessee State Board of Education or the Tennessee Education Association. The data from each individual agency may or may not have become part of the data collected from the legislative meetings, depending on the interactions of the policymakers. Finally, each analysis was compared with the others, in an effort to compare concepts, categories, and relationships between categories for similarities and differences. According to Denzin (1978) this strategy of data triangulation is basically "the use of dissimilar comparison groups as a sampling strategy" (p. 301). He continues his explanation:

By selecting dissimilar settings in a systematic fashion, investigators can discover what their concepts (as designators of units of reality) have in common across settings. Similarly, the unique features of these concepts will be discovered in their situated context (p. 301).

This data triangulation strategy allowed for the "consistency in overall patterns of data from different sources and reasonable explanations for differences in data from various sources contributing significantly to the overall credibility of the findings presented" (Patton, 1987, p. 162), thereby increasing the credibility of the study.

Summary

The conditions of this study met the primary requirements for using a case study research design. Documents and records were analyzed that preserve the historical context and decision-making processes of educational policymakers in Tennessee. Specifically, the documents and records related to

the development, implementation, and evaluation of the EIA of 1992 and the use of assessment as part of that mandate's design were reviewed. In an effort to stay as close as possible to the historical context of the enactment of the EIA, documents and records were collected from each major policy agency and then analyzed using a process coding within an open coding technique. To corroborate the evidence presented in the analysis of the data, the researcher used a data triangulation strategy, comparing concepts and categories defined in data from the legislative meetings across multiple-agency settings.

CHAPTER 4

THE FORMULATION OF THE EIA

The Education Improvement Act (EIA) of 1992 is an artifact of dynamic but not dynamical activity, of debate but not dialogue, in which the issue of continuing education reform in Tennessee became accountability rather than the teaching/learning process. Policymaking participants joined the debate with the intent of improving student performance. In the end, each argument was too weak to perturb the system to a higher level of complexity by breaking with the tradition of mandated educational policy to allow for the self-organization of the schooling system at the local level. The winners in this debate were those who ruled in favor of "efficiency," "effectiveness," and "an improved final product—a student prepared for the 21st century workforce."

Undertaking the task with the best of intentions, policymakers failed nonetheless to confront their basic assumptions about teaching, learning, and improvement, leading to participants' defending their current positions and resisting change. Weighing the political risk of alienating constituents with increased taxes or new taxes against the threat of an impending judicial ruling that state funding methods for education were unconstitutional, the Tennessee General Assembly chose to maintain as much of the existing system as possible with the least cost to the constituents by enacting the EIA. When caught between political risk and court order, policymakers relied on past experience of using mandates to create "an accountability system second to none in the nation." The resulting legislation gave the appearance of

changing the system to allow for self-organization for genuine deep-rooted reform, but, in fact, did not.

A chronology of the debate followed by an analysis of the debate and its resulting mandate comprise the contents of this chapter. Findings presented will begin with the conditions for reform and proceed with the interactions of those participating in the selection of policy strategies during the formulation stage of the EIA. A discussion of the resulting action of the policy formulation stage will conclude the chronology.

The EIA is composed of four major components: accountability, management, funding, and curriculum. The analysis of the findings will focus on the accountability sections of the EIA and will include 1) a dimensional profile of accountability during the policy formulation stage related to the properties of resource intensiveness, targeting, institutional capacity, and political risk; 2) a comparison of the EIA with other educational mandates for similarities and differences. In Chapter 5, conclusions, implications, and recommendations concerning the findings include a discussion of issues and underlying assumptions as a basis for the EIA, specifically assumptions underlying the use of assessment in the overall design of accountability.

Policy Formulation: The Conditions for On-going Reform

It was, to borrow a phrase from Dickens, the best of times and the worst of times for continued education reform in Tennessee. Education reform nationwide was a top priority, with George Bush dubbing himself as the "Education President" for election in 1988. The National Governors Association, rallied by former Tennessee Governor Lamar Alexander to focus

on improving education, was continuing to work in that direction. In Tennessee, some of Alexander's Better Schools Program goals were being partially met with some improvement in standardized test scores (Snodgrass, 1990). And, importantly, the reform movement in Tennessee had awakened to the public's awareness of problems in public education (*Nashville Banner*,⁴ 6/7/88). Still, there was discontent with education in Tennessee.

Discontent around the policymaking table took the form of a need for more money by educators. Tennessee ranked 37th in high school graduation rate, 45th in expenditures per pupil, 44th in pupil/teacher ratio, 9th in federal funding, 30th in state funding, 25th in local funding, and 37th in teachers salaries (Office of Planning, Budget, and Evaluation, 1989). To reduce class size, to make kindergarten mandatory, to improve the attendance and graduation rates, to introduce technology, to retain good teachers and attract new teachers, to revise the evaluation system, and to prepare students for economic competitiveness in a global market in a state with high illiteracy and drop-out rates, more money was needed. Though the Comprehensive Education Reform Act (CERA) provided more than \$350 million for education reform in 1984, Tennessee Education Association (TEA) Executive Secretary Cavit Cheshier defended the need for additional money:

There will not be any significant improvements in K-12 education in Tennessee in 1991 or any subsequent year unless there's a substantial revision of financial support of education. As utilities go up, it takes more money. As the cost of textbooks go up, it takes more money. As more handicapped children come to our schools, it requires more money. If more programs are placed upon the schools to implement, it takes more money. And I submit to you that unless we want to pay for it, we

⁴ The *Nashville Banner* and *Tennessean* are two daily newspapers of record in Nashville, Tennessee. In this paper they will be referred to as the *Banner* and the *Tennessean*.

probably won't have it (Legislative minutes, House Education Committee, 4/8/91).

For the legislators, discontent took the form of anxiety from political risk in asking their constituents for more money on the one hand and avoiding a court order for a new state education funding formula on the other. The legislature was gearing up to play a major role in the coming education reform and would look to the business community for support. Senator Ben Atchley said:

I've read a lot where different people have stated throughout the country that education will not reform itself from within. It's gonna come from without, the people that use the product of the system, and I'm glad to see the business community stepping forward (Legislative minutes, Senate Education Committee, 3/6/91).

For the business community, schools were the link to increased productivity, improved quality, global economic competitiveness and, ultimately, more profit. For business leaders, discontent took the form of dissatisfaction with the poor management of schools and a lack of accountability for the final product—students prepared for the workforce in the 21st century. Education needed to adopt the corporate model for running schools with the superintendent as the chief executive officer appointed by the local board of education. John Parish, speaking for the Tennessee Business Roundtable, described the accountability the business community wanted:

I firmly believe that appointed superintendents is absolutely the right thing to do. . . In this plan, I think it's the strongest part. Like it's been said, it's the part that doesn't cost anything, but it's

a thing that can make a huge difference in the quality of our schools (Legislative minutes, Senate Education Committee, 3/6/91).

And for the judiciary, which invited itself to a seat at the policymaking table when the Tennessee Small School Systems filed a Davidson County Chancery Court lawsuit against the state on July 26, 1988, discontent took the form of weighing the evidence concerning the funding system for education in Tennessee to determine its constitutionality.

Economic conditions during this period were sluggish and would later worsen. The state was facing a \$46 million revenue shortfall. According to Finance Commissioner David Manning, sales tax growth was below budget estimates "primarily due to a slump in sales that [had] been seen across the nation" (*Tennessean*, 11/15/88). This revenue shortfall presented a significant problem because recent changes in welfare laws under the federal Catastrophic Health Care Act and the Welfare Reform Act would require some major state appropriations over which the state had virtually no control (*Tennessean*, 11/15/88). Stan Chervin, director of research for the Tennessee Department of Revenue, explained that the state's tax structure had not been generating enough revenue to keep up with inflation because "consumers are spending an increasingly large share of their money on services, which are exempt from sales tax" (*Tennessean*, 10/18/90).

When Governor McWherter proposed a state income tax to move away from an increase in sales tax to fund further education reform, there was an outcry from the Tennesseans for Limited Taxation Association, which told the legislature,

In the country in general, we just had a \$500 billion federal tax increase, postage rates went up, the savings and loan bail out is not yet concluded, we don't know what the Gulf War is going to cost, and we are in the middle of the recession still and when you add to that pot a proposed state income tax, which is a massive tax increase, that is why legislators are having trouble with constituents not trusting what they're going to do (Legislative minutes, House Education Committee, 4/8/91).

Citizens had come to the table discontented with proposed tax increases. They felt much of the needed reform could be accomplished without raising taxes. In further comments to the House Education Committee (4/8/91), the spokesman for the Tennesseans for Limited Taxation said that when legislators asked for increased taxes, citizens have to give something up—a vacation, wearing a pair of shoes for a longer period of time. Historically, increasing taxes for educational improvement in Tennessee had never fared well during poor economic times.

The Initiator of the Education Reform Debate

Ned Ray McWherter saw education reform from two perspectives: legislative and executive. In 1984, as Speaker of the Tennessee House of Representatives, he signed the CERA under the administration of then Governor Lamar Alexander. In 1992, as governor of Tennessee, McWherter signed the EIA into law. When McWherter took office in 1987, he had a legislative perspective of the past education reform efforts of Lamar Alexander who during his campaign for governor (1978) “walked 1,022 miles across Tennessee to talk and listen to citizens. [Alexander’s] administration had education as its top priority and Alexander’s Better Schools Program and the career ladder pay plan for teachers drew national attention” (Darnell, 1995–1996, p. 463).

CERA's accountability system was an evaluation process for educators designed around a direct instructional model called the Tennessee Instructional Model (TIM). "TIM gave educators state-wide a common language when discussing teaching and learning instructional tactics and strategies" (Green, Smith, & Hydorn, 1995, p. 19). Based on an educator's years of experience and evaluation performance, he or she could apply for the Career Ladder Program for incentive pay and the option of working under an extended contract—more hours for additional pay. It was the Career Ladder Program and the Extended Contract Program that created the most controversy in implementing CERA.

In 1984 the Tennessee General Assembly had funded CERA with \$350 million new dollars (Wissner, *Tennessean*, 9/29/90). By fiscal 1987, the State Department of Education was spending \$1.4 billion and employing a staff of 1,305 (Snodgrass, 1990, p. i) to implement CERA. Despite this transfusion of money, there was discontent with the progress in accomplishing the goals for improved performance in Tennessee's schools. The Comptroller's Performance Audit report (Snodgrass, 1990) stated that "some of the measures indicate the state as a whole has improved. For almost every measure, some systems have not improved at all and some have worsened" (p. iii).

A history of mandates had failed to bring about the intended outcome of accountability for improved student performance compared with the other industrialized nations of the world. This accountability was the intended outcome of mandated evaluation. The Elementary and Secondary Education Act (ESEA) of 1965, the Education for All Handicapped Children Act (P. L. 94-142) of 1975, and Tennessee's own CERA of 1984 had become increasingly

more specific as to the use of evaluation to bring about improved student performance.

The result of this specificity was to remove the evaluation process further and further from the center of the teaching/learning process by putting the emphasis on "effective procedures, including provision for appropriate objective measurements of educational achievement, [that] will be adopted for evaluating at least annually the effectiveness of the programs in meeting the special educational needs of educationally deprived children" (*U. S. Statutes at Large*, 1965, p. 31). CERA, then, gave specific instructions for effective evaluation procedures. It specified the purpose of evaluation to:

... study the use of student progress or achievement as measured by standardized testing or appropriate measures, as an indicator of successful teaching and effective schools and to review periodically the standards and criteria used for teacher and principal evaluation in view of the findings resulting for such study (*Tennessee Public Acts*, 1984, p. 31).

CERA specified when to administer the test, what grade levels would be tested, who would develop other relevant measures and educational goals, and how much improvement was expected by the General Assembly:

... it is the legislative intent that nationally normed achievement tests be selected for administration at three grade levels in the 1984-1985 school year and that within five years, the statewide average test scores achieved on such tests by Tennessee students be higher than the national average for comparable grade groups in each subject (*Tennessee Public Acts*, 1984, p. 64),

... 15 percent increase in the number of students mastering each skill in reading and math as measured on the Basic Skills criterion-referenced tests in grades 3, 6, and 8 (*Tennessee Public Acts*, 1984, pp. 63-64),

... it is the legislative intent that the State Board of Education develop other measures, goals, and/or benchmarks and submit to special joint committee of the General Assembly (*Tennessee Public Acts*, 1984, p. 64).

With past reform in mind, when McWherter took over as governor, he had a vision for future education reform in Tennessee: "to give every young person in Tennessee an equal opportunity to compete with the best students anywhere in the world" (*Tennessean*, 1/27/90). In 1987, Gov. McWherter began the work of improved education reform by embarking on a nine month "Listening to Tennessee" tour across the state. Over the next five years, the concerns and ideas of "anyone who has any relationship to our schools" (*Banner*, 11/3/89) were developed to form the next major education reform legislation known as the EIA of 1992. The first four years were spent in building consensus among the major policy players. The next year was spent putting the pieces of the program in place and garnering political support. The plan for educational improvement was McWherter's top platform priority in seeking re-election in 1991. The final bill was signed into law on March 11, 1992.

The educators heard from on the listening tour spoke of the need for more dollars, fewer state regulations, more time and greater flexibility in implementing state-mandated programs, a voice in curriculum matters, and an evaluation process for teachers that would allow more local input and that would provide more guidance for teacher improvement. Echoes of past educators concerned with the evaluation system of the ESEA were heard in a plea from teachers to superintendents for "freedom to be creative in the classroom and to be freed from the necessity of teaching to the test" (Smith, 1987, p. 9-10). A majority of local educators believed the Basic Skills program

(the testing program of CERA) to be "flawed by too much structure, too much paperwork, and too much testing. The general belief [was] that creativity in the classroom [had] been negatively impacted by the program" (Smith, 1987, p. 12). On the other hand, business leadership and a large portion of the public at large were concerned about accountability in public education and the inequitable state funding formula in use at that time. Policymakers were faced once again with the problem of ensuring that children get equal educational opportunities without stepping on the professional toes of educators.

After the Listening Tour was ended and the shaping of the plan began, Governor McWherter set the tone for the conversation at the policymaking table with the following remarks:

I want policies to relate to standards and goals and accountability. . . . And I think I'm in a unique position, as I can say to the scholars, to the State School Board, to the Commissioner of the Department of Education, to the teachers in Tennessee: "you develop the program and we'll get the expertise to help you. You develop the program. Show us what the program will do before you ask us for the money. If you do that, if you will develop the program that gives us standards, goals, accountability and a quality education system and an equal education system for all boys and girls, then I'm prepared to get on the line to do what's necessary to make the funds available" (Governor's Papers, 9/25/89, State Library and Archives).

With respect to the request for more money, the governor went on to say:

I will ask for no additional funds in the name of education that do not go into education. In the 18 years that I served the General Assembly, I observed tax increases from time to time under different administrations and the taxes. . . the revenue increases. . . would all be sold in the name of education. And in most cases that I can remember, thousands and thousands and thousands of dollars got into other state programs other than in

the name of what they were raised for. When I ask for revenue for education in K-12 or higher education. . . and when I ask for accountability. . . I will be accountable to see that those dollars go into education. Accountability will reach all of us (Governor's Papers, 9/25/89, State Library and Archives).

Accountability was a point of conversation that became the theme of debate before it ever got to dialogue in which basic assumptions about teaching and learning were stated explicitly and the implications of those assumptions discussed in terms of the development of educators and students alike. The debate lined up along the issue of accountability for an improved end product as described by Rep. David Copeland as "the caliber of the student coming out" like the employee of a business is "what can he produce" (Legislative minutes, Joint Education Oversight Committee, 10/30/90).

Teachers Come to the Table

Teachers felt justified in coming to the table with requests for more beginning money because Governor McWherter had made the \$18,500 salary promise a part of his gubernatorial campaign and the state's teachers association had delivered their support and votes and wanted the governor to make good on his promise (TEA News, 1986). In a workforce of 46,000, more than half—25,000 teachers— would be eligible to retire by 1993 and the current salary was \$14, 760. Teachers were working in a state in which one out of four students dropped out before graduating from high school and in which more than half of the adults over age 25 did not have a high school diploma in 70 of 95 counties. Terry Weeks, a Murfreesboro teacher, in remarks made upon receiving the national Teacher of the Year award, said, "Economic prosperity is ultimately linked to an educated citizenry and to

good schools. . . the task for policymakers is to provide the training and resources to enable all schools to become effective schools" (*Banner*, 4/24/87).

Teachers were also unhappy with their evaluation system to receive merit pay. When 95 percent of the TEA membership was questioned about the Career Ladder Program, they responded that it was unfair and ineffective (*Banner*, 1/24/87). TEA's membership voted to seek legislative changes in the program, including alternative methods of evaluation for the upper levels of the ladder and more local input in the evaluation process rather than the current use of outside state evaluators (*Banner*, 1/26/87). The proposed bill "would allow teachers to earn merit raises by successfully completing teacher training courses or workshops. It would not require that teachers be evaluated [by the state's evaluators]" (*Banner*, 3/24/87). TEA's suggested changes to the Career Ladder Program were controversial among reformers and teachers alike and provoked intense opinions in the media over the next two years.

We know the Career Ladder does nothing of significance to improve teaching in Tennessee. It is a travesty both to school children and taxpayers to promote a program which was designed primarily for political reasons using the name of "education reform" (*Banner*, 1/29/87).

It's very difficult to measure teaching ability. This one [career ladder evaluation] is very narrow. Many wonderful, creative teachers did not make it. Some teachers I consider mediocre did make it. So many important things are not measured (Teacher quoted, *Banner*, 3/23/87).

We wonder why TEA bothered to require the seminars. Why don't we just pay the bonuses to any teacher who sends in proof of TEA membership, two box tops from his or her favorite cereal and 50¢ for postage and handling (*Banner*, 3/8/89).

Studies, Proposals, and Resolutions: Analysis from Within the System

In April of 1987, the Senate approved a resolution to revise the Career Ladder. In May of 1987, a study of the Career Ladder Program and the Extended Contract Program began under the direction of Dr. Jerry Bellon of the University of Tennessee-Knoxville. Dr. Bellon coordinated the research of a project team that included members from both the State Department of Education and the State Board of Education. In addition, a 17-member task force of local educators assisted the project team in completing surveys and a review of the literature. Commissioner Smith discussed preliminary results of the study with the legislative Joint Education Oversight Committee (12/17/87). Smith said, "In large measure this state's Better Schools program is working. And the two areas in which problems exist—the Career Ladder and the Basic Skills First—are today in line for corrective action."

At the same committee hearing, Dr. Brent Poulton, Executive Director of the Tennessee Board of Education, picked up the discussion while answering questions concerning the development of evaluation processes for other educational personnel besides teachers and principals. His response to that question was as follows:

You have to create a system, an evaluation system, that essentially works off a standard, uniform job description and you get into situations like assistant superintendent, for example, where the nature of the work is so varied that it really doesn't lend itself to the kind of evaluation system that we have now. We owe it to that kind of a situation to come up with an alternative approach to evaluate if indeed we want to allow everybody an opportunity to be in a career ladder program . . . if merit pay has value for one group, I think it has value through the system and we ought to do it. But it is a process that takes a lot of time. . . and it has to be something we can live with, something that won't result in all kinds of court challenges or

appeals. And that's the cautionary note (Legislative minutes, Joint Education Oversight Committee, 12/17/87).

In addition to the Career Ladder, Sen. Leonard Dunavant followed Dr. Poulton by introducing a resolution that called for a study to be completed by July 1, 1988 concerning

the use of student progress or achievement as measured by standardized testing or other appropriate measures as an indicator of successful teaching and effective schools and to review periodically the standards and criteria used for teacher and principal evaluation in view of the findings resulting from such study, and whereas the use of student progress or achievement as a factor in the evaluation of teachers has not yet received the attention which it requires, and whereas ultimately the advancement of students from one level of knowledge or skill to a higher level of knowledge or skill is the purpose of education and is the reasonable expectation of parents, taxpayers and the business community for the massive commitment of money and effort which is made in the name of education reform. . . be it resolved that the state evaluation commissioner include in evaluation criteria for teachers appropriate indicators of student progress. . . (Legislative minutes, Joint Education Oversight Committee, 12/17/87).

The resolution inspired some debate among committee members about the valid use of standardized testing. The following is an excerpt from that debate:

Sen. Robert Rochelle: . . . the product [student progress] is still a valid factor in considering the progress of the teacher. Just as a lawyer is judged partially among other things by how many times he wins a lawsuit. And dentists are judged among a lot of other things by whether the fillings they put in stay in. That's a product. . . recognizing again that it is very difficult to determine the progress of the student, it is still a relevant factor to be looked at in a teacher's performance.

Rep. Leslie Winningham: . . . If you are going to give an adequate evaluation of a teacher based on student performance you are almost going to have to determine the movement of that student 24 hours a day—everything he does or she does and we are going back to where we were three years ago in our evaluation process. We ought to be moving ahead instead of backwards (Legislative minutes, Joint Education Committee, 12/17/87).

The resolution was approved. Certainly the shift from CERA's accountability for process to accountability for results was on the rise in legislative thinking.

In his report to the commissioner concerning the results of extensive surveys across the state and in-depth interviews with 600 teachers and 100 principals about Career Ladder and Extended Contract, Dr. Bellon (1988) stated,

Learning is a complex process that depends on the best and most insightful efforts of teachers as well as students who have a predisposition or desire to learn. Teachers are more likely to make a commitment to professional development activities if they believe these activities will improve their effectiveness with students so that student learning will be enhanced (p. 3).

Education Commissioner Smith reported the results of the study with one of the major concerns being that teachers

welcome evaluation activities that help them improve their teaching performance. They perceive, however, that the current evaluation system does not provide the feedback necessary to help them improve and in too many cases actually inhibits effective instruction. . . . Another important concern is that current programs [extended contracts] are too often driven by teacher and school needs rather than by students' needs. (Tennessee Department of Education News Release, 1/26/88).

Proposed recommendations for proposed changes in the state evaluation system for Career Ladder Level II and III certification based on the Bellon study were as follows:

- Involving the principal or other local administrator as a member of the evaluation team. This option would be mutually agreed upon by the person being evaluated and the local administrator. The local administrator would conduct one of the three evaluation visits and take part in the summative conference.
- Giving Career Ladder candidates more timely feedback on areas of strength and factors needing improvement.
- Collapsing the time frame for conducting state evaluations.
- Keeping to a minimum the amount of paperwork Career Ladder candidates are required to submit during the evaluation process.
- Separating extended contracts allocations from the Career Ladder program to give school systems more flexibility in meeting locally identified student needs.
- Establishing an incentive grants program to help school systems address high priority instructional needs.
- Sponsoring research programs in selected Tennessee school systems to identify factors contributing to and detracting from the quality of school life (Tennessee Department of Education, 1/26/88).

TEA reported Dr. Bellon's remarks in the February 1988 issue of the *TEA News* as follows:

[Bellon] called the career ladder evaluation "demeaning" because it "attempts to reduce teaching—a complex process—to seven or eight simple principles. . . We have never found a state mandate that has improved education. It may have improved test scores but not learning" (p. 4-5).

A study done by French, Malo, and Rakow (1987) came to the conclusion that "the system works as it was intended to work in the technical sense" (Tennessee State Board of Education minutes, 5/29/87). And, in 1990, the "Career Ladder Validity Study" done by an outside evaluation team reported that the "Career Ladder staff had done a stellar job of conceptualizing, developing, and implementing a model teacher evaluation system that allows for valid inferences about teaching quality" (Tennessee State Board of Education minutes, 9/19/90). The State Board of Education's response was their work to "make some substantial changes in the evaluation system that are not inconsistent with the findings" of these subsequent studies and said that while the validity study

concludes that the procedures incorporated in the current evaluation system are valid, it does not speak directly to the issues of efficiency or participant perception. It is our belief that an equally valid evaluation system can be created which will be much less intrusive, easier to administer, and appreciated by a greater number of those for which it is designed to evaluate (Tennessee State Board of Education minutes, 9/19/90).

A Study of the Better School Program's Testing Program

A study of the state-mandated student testing program that had paralleled the Bellon study was one conducted by the Department of Education and coordinated by Tom Cannon, Assistant Commissioner for School Success (Smith, 1988). The proposals based on the study were to 1) combine the Basic Skills First criterion-referenced test with norm-referenced test items for each grade level from second through eighth grade, 2) increase collaboration between the curriculum staff and the testing staff in the development of the test items, and 3) implement a writing assessment by 1990. Cannon said,

We believe this is an optimal testing program. It doubles the information available to the school systems and the state, gives continuity to test data generated from year to year and is designed to make the testing program more efficient and more useful to educators. By working with the curriculum consultants we are trying to tailor the test to fit the curriculum. The curriculum ought to be driving the test rather than the test driving the curriculum. We also are recommending that the criterion-referenced test be developed by general subject area (domain) instead of by isolated objectives. This will relieve teachers from feeling they must teach to the test. (in Smith, 1/27/88)

The proposals were approved by 50 local educators serving kindergarten through twelfth grades comprising a state Testing Advisory Council. Much like the debate during the formulation of the ESEA, the proposals for change in the testing program would become the common language of legislators, businessmen, educators, and others—"efficiency," "effectiveness," and "coherence between the test and the curriculum." The debate quickly turned from "learning is a complex process and how do we improve the quality of school life to facilitate that process" to "how do we make the testing program and schools, in general, more efficient."

Court, Consultant, and the Business Community Join the Debate:

Intervening Conditions

Educational funding in Tennessee had been under study for a number of years when the Kentucky Supreme Court ruled in Rose v. The Council for Better Education, Inc. (790 S.W. 2d 186) that Kentucky's entire school system was unconstitutional. The Kentucky Supreme Court ruled that:

Kentucky's system of common schools is underfunded and inadequate; is fraught with inequalities and inequities throughout the 168 local school districts; is ranked nationally in the lower 20-25% in virtually every category that is used to evaluate educational performance; and is not uniform among the districts in educational opportunities (Alexander, 1992, p. 31).

Dr. Brent Poulton was to review the Kentucky case for the Education Oversight Committee of the General Assembly. In his review, he compared the Kentucky case to a similar lawsuit pending in Tennessee, Tennessee Small School Systems v. McWherter (851 S.W. 2d 139), in which the small systems claimed that state funding statutes creating disparities were unconstitutional, violating both the education clause and the equal protection clause of the state constitution. In response to a question from committee members about why state mandates for personnel had not been adequately funded in the past, Poulton replied:

Merely going back over the past few years and looking at actions either by the State Board of Education or the General Assembly which increased the need for personnel and identify for whatever reason systems where enough money wasn't available to fully fund those positions really amounted to tinkering with an overwhelmingly archaic system that if we're really serious about having good schools in the State then we ought to go to the heart of the problem and change the system (Legislative minutes, House Education Oversight Committee, 7/12/89).

Poulton then went on to explain data collected in past funding reports completed by the State Board of Education and sent to the governor and the General Assembly. One report included the study of the Tennessee Foundation Program (TFP), "created in 1924 as a means for insuring adequate funding in Tennessee's public schools" (Weeks, 1988, p. 1). This study cited three problems with the TFP:

- There is no link between the changes in appropriations and changes in the costs of delivering programs and services at the local level.
- The amount of funding provided is too little to ensure adequate funding of a basic educational program in all systems.
- It does not ensure the children in Tennessee equal access to quality educational resources (Weeks, 1988, p. 3-4).

The study goes on to define a basic education program (later to be called the Basic Education Plan or the BEP) in terms of "essential elements for school operation" such as personnel including teachers and custodians and non-personnel such as transportation and textbooks, program areas, staffing ratios, and expenditure levels and to develop a new formula to fund these elements equitably. There is no mention in this definition of students nor the essential elements of a learning program. How much would it cost to provide all the elements in the BEP? "An estimate of cost was produced by applying the program and cost specifications to 1986-87 school census data. . . it would cost approximately \$1.587 billion" (Weeks, 1988, p. 7). In repeated attempts since 1909, policymakers tried to develop a plan to equalize funding across the counties of the State. Now the defendants in Tennessee lawsuit were highlighting that failure.

The State Attorney General, Charles W. Burson, defending the State said,

There are no constitutional rights to any specific allocation of funds for public education. All that is constitutionally required is some legislative scheme for funding public education that has a rational basis. These requirements have clearly been met. Even if students do not have computers or new textbooks as the complainant alleges, there is no indication that they are denied access to the fundamentals of education (*Tennessean*, 11/27/88).

Cavit Cheshier, Executive Director of the TEA, responded that the defense's brief

reflects a narrow view of what is required for a basic level of education in today's society. Computers, up-to-date textbooks, modern vocational training equipment and reasonably sized classrooms are essentials for educating children and preparing them to live in an increasingly complex, technological society (*Tennessean*, 11/27/88).

Clearly, the opinions of what constitutes a basic education was determined by the assumptions of the beholder, whether you were in the classroom—textbooks, computers, and, the right ratio of students to teachers; the funding agency—personnel and operations; or the State—minimum requirements of a system. In 1873, that had meant a governance framework and the “collection and dissemination of statistical and other information related to the schools” (Warf, p. 8). Since 1873, a lot of basics had been hung on that framework.

####

Donald Thomas, an educational consultant, was hired by Governor McWherter to assist in developing a funding formula that would narrow the disparities between counties across the state. Thomas identified inadequate funding as the overriding problem in Tennessee but doubted that the current method the State was using would violate the Tennessee Constitution (Wilson, *Banner*, 2/28/90). “There is no illegality. All it says in the constitution is that the state shall set up a system of free public schools. It has done that.” Thomas went on to say that the provision for a “system” of free public schools had become part of the constitution only a few years before. “It’s under miscellaneous. That blows my mind—education in an after thought.” Under Article XI, entitled “Miscellaneous Provisions,” Section 12, the Tennessee Constitution states:

The State of Tennessee recognizes the inherent value of education and encourages its support. The General Assembly shall provide for the maintenance, support, and eligibility standards of a system of free public schools. The General Assembly may establish and support such post-secondary educational institutions, including public institutions of higher learning, as it determines (Darnell, 1995-96, p. 511).

The education section appears right after the "personal property exemptions" section and right before the "protection and preservation of game and fish" section of the constitution.

Thomas, who had directed the Division of Public Accountability in the State Department of Education in South Carolina and was a nationally recognized figure in education reform efforts, made the following broad recommendations to the governor in his 20-page, detailed report, "Tennessee Education for the 21st Century (11/90)." Recommendations related to establishing the 21st century classroom include the following:

To create 21st century classrooms in all Tennessee schools, it is necessary to establish the most effective conditions for teaching and learning. These include:

1. High expectations for all students with appropriate, rigorous programs of study and classrooms staffed by competent professionals.
2. Support programs and specialized assistance to students with special needs.
3. State-of-the-art teaching practices supported by the most effective technology.

At the same time, schools must be given greater flexibility in serving an increasingly diverse body of students. School principals and teachers must be given greater opportunity to make decisions at the school level. Parents must be enticed into becoming actively involved in their children's education. As schools attain greater responsibility, they must also assume

greater accountability for results, measured by a broad array of instruments (p. 3).

Recommendations related to accountability and school governance include the following:

To support the restructuring movement in Tennessee, it is necessary for the state to establish accountability, a stronger governance system, and decentralized decision-making. Accountability for school results should be established against standards which can reasonably be achieved. Those in leadership positions should provide direction under conditions which are free from undue influence. At the school level, principals and teachers should have professional responsibility for selecting instructional materials, deciding on appropriate methodology, and making decisions about the use of school funds.

The challenge to policy makers is to concentrate on standards to be achieved and not methods for achieving them, to hold people accountable for results and to take action when results are not attained, to encourage and support the separation of the public will (what schools should teach) from the professional prerogatives (how schools should be organized and how teachers should teach) (p. 8).

Recommendations related to adequate and sustained funding for education include the following:

If the restructuring program is to work in Tennessee, the legislature must provide adequate financial support for all children, regardless of where they live. Schools must have funds to operate a basic education program, meet the needs of special groups of children, establish realistic staffing ratios, employ and pay competent personnel, purchase appropriate instructional materials, transport students, and obtain needed equipment. School districts must also be able to renovate old buildings and build new ones when needed.

The financial program in the state should be a partnership between the state and local districts. The state should, of course, carry the major burden. The financial program should be built on stable tax sources and have revenues which are placed in a dedicated trust fund for education. The legislation should also provide opportunities for individuals, agencies, and foundations to make donations to educational efforts in the state. Such funds should be placed in a state education fund for initiating innovative practices and rewarding exemplary performance.

The challenge for policy makers is to establish adequate financial support on a continuing basis—one that goes beyond the basic education program and incorporates all the components of an appropriate and adequate state educational system (p. 15).

In his final note, Thomas said:

The implementation of these recommendations will require that the State Department of Education and the State Board of Education be strengthened in their accountability functions; in their ability to collect and synthesize information; and in their training, leadership, and staff development functions. One should examine the degree to which their current resources and structures support their ability to take on these responsibilities (p. 18).

Many of these recommendations find their place in the EIA of 1992.

With the recommendations, the specter of systems analysis, a management concept introduced to the Department of Health, Education and Welfare by William Gorham in the 1960s and used in the accountability system of the ESEA, whispered around the education reform table, "Data generated by [a state] reporting system could be displayed in cost-benefit terms, thereby leading to more effective local practice and more efficient [state] decision making" (McLaughlin, 1975, p. viii). In the ESEA system of accountability, information about students' academic performance was measured using standardized tests. Academic performance or output was related to monetary input using production-function models of efficiency.

###

Dan Frierson, a Chattanooga businessman representing the Tennessee Association of Business as chairman of the education committee, spoke to the Senate Education Committee (3/6/91) concerning the support of the association for increased costs to continue education reform if the elements of a corporate management system were put into place. Sen. Ray Albright, Chairman of the Education Committee, reiterated the elements when he said:

If the chair might, and I might be the only one that got this opinion, but I found six points in your presentation and I find your statement "that without a sufficient system of management in place, the cost of these [reform efforts] cannot be justified. I anticipate that that's a pretty strong statement. . . The corporate management system, what we call the corporate model, is the elected school board, appointed superintendent, principal's not being part of the bargaining unity, value-added testing for our students, . . . sufficient rewards and sanctions for those who perform and those who do not perform, eliminating the Chancery Court appeal, an audit by a professional outside the system and recommended the Comptroller of the State of Tennessee. . . These are what need to be, in your opinion, a management in place if this bill is to have support of your organization or of your education committee. I interpret that from your remarks sir. (Legislative minutes, Senate Education Committee, 3/6/91)

Frierson responded, "Mr. Chairman you've gone to the heart of it."

Another spokesperson from the business community, John Parish, echoed "support for reform in return for accountability for results" and that accountability would be based on the corporate model. Parish, a Tullahoma businessman, speaking for the Tennessee Business Roundtable told the Senate Education Committee:

Tennessee is just one of just a few states that still selects its chief school administrators in political popularity contests. . . I also know that some of your local constituents have complained of

our depriving our citizens the cherished right to vote when you remove the office of the local school superintendents from the ballot. . . [But if appointed by the local school board, superintendents] would then be able to focus all their skills and energy on their professional duties and not be forced to make decisions based on political expediencies. . . I think [accountability] is by far the key ingredient [to education reform]. A good step in that direction is to create a system that holds the school board accountable for policy. The school superintendent for execution, that is, carrying out the policy. Teachers who teach and students who learn. If you do nothing else but change the way we select school superintendents, you will have taken a good step in the right direction in improving education in Tennessee. Also, I might add, that this costs nothing, no new revenues are needed and no new revenues are required. . . And I can assure you [that if you fail to pass this part] the support for [education reform] by the business community would be in serious doubt, as well as, the support to pay for it (Legislative minutes, Senate Education Committee, 3/6/91).

####

The Kentucky Supreme Court was to lay the "sole responsibility" for equal educational opportunities for all children squarely on the shoulders of the Kentucky General Assembly to increase funding for education to alleviate disparities and to build a new system of adequate education for all children based on seven basic "capacities:"

- sufficient oral and written communication skills to enable students to function in a complex and rapidly changing civilization;
- sufficient knowledge of economic, social, and political systems to enable the student to make informed choices;
- sufficient understanding of governmental processes to enable the student to understand the issues that affect his or her community, state, and nation;
- sufficient self-knowledge and knowledge of his or her mental and physical wellness;

- sufficient grounding in the arts to enable each student to appreciate his or her cultural and historical heritage;
- sufficient training or preparation for advanced training in either academic or vocational fields so as to enable each child to choose and pursue life work intelligently; and,
- sufficient levels of academic or vocational skills to enable public school students to compete favorably with their counterparts in surrounding states, in academics or in the job market (Alexander, 1992, p. 32).

These capacities were very similar to the curricular goals listed in the State Board of Education's *Rules, Regulations, and Minimum Standards for the Governance of Public Schools in the State of Tennessee*. Book 4, Goals of Education and Curriculum Frameworks, (May, 1988). The Tennessee General Assembly felt the hot breath of court-ordered reform down their legislative necks and began looking for answers as to how the situation had gotten to the point of a lawsuit rather than just doing what needed to be done. Dr. Brent Poulton, Executive Director of the State Board of Education, and Rep. Paul Starnes brought some insight in the following exchange about the social studies curriculum and students' lack of understanding of their government. Reform was a philosophical question.

Dr. Poulton: . . . this just happens to be the year that the social studies curriculum has been reviewed and . . . in effect [the State Board of Education] deferred accepting the proposed framework because on the one hand they were looking at performance indicators that said our students weren't doing well enough and at the same time we're looking at a framework and a schedule of courses that was for the most part the same as before. . . And then I look back upstream at all the time that was supposed to be spent on aspects of that and ask myself why is it that so many students don't understand, seem to understand, these ideas when there was supposed to be instruction in the fourth grade

and the fifth grade and so forth. And why is it we don't get better results from our existing commitment? I think that question has to be answered first before we say we're going to insist that everybody take a government course before they can graduate from high school.

Rep. Starnes: We can also apply that same analysis to the teaching of grammar. We start teaching grammar early on and we teach it through grammar school and through middle school. . . we get them their senior year and they still, in many cases, do not know proper grammar from "kanip kanop." And yet that's something we drill. English is one of those basics, till death do us part, that as students you've got to have.

Pouton: I agree and it's exactly the same point. I just can't bring myself to suggest that we ought to have a grammar requirement in high school. Surely there ought to be a way that you can teach children grammar prior to them reaching high school. Yes, you're right we do have students who obviously don't fully understand when they get to high school. So I want to go back upstream and see what's been done before and in what ways we can make improvement before they get to high school.

Starnes: . . . To go back to what you were saying about what we should put in the social studies curriculum. . . the point that I would like to make is that we can learn probably what is expected as a national norm from questions that you would find on standardized tests. Just as a guide. . . the point I'm trying to make is we have not, in fact, as educators, kept up with the curriculum offerings that should, in fact, take place. We're always trying to play catch-up. Now we've got low scores in social studies and we're trying to figure out what to offer. Should we offer geography? Do we need to offer world history? Should we offer contemporary history? Do we need to understand China better? Do we need to understand the Soviet Union? Do we need to know what's going on in South and Central America? And we don't teach those things.

Poulton: The answer's yes to all of those. That's the dilemma that we're in. I don't know whether it was set out in the beginning or we set out a system to accommodate it. Or maybe it's just possible to deduce what our philosophy must be in what we've done. We have a philosophy that in high school there's a very narrow core of required courses. The state says that you shall take this—this narrow core of courses. It amounts to about

a fourth of the opportunity that the student might have. But beyond that, courses that are offered are largely dependent on what a local school system wants to offer. . . Our philosophy is to offer a cafeteria style program as opposed to other countries having a very narrow, very defined core of courses that may be academics or may be something else. We ask ourselves, "What do we want? How do we want to proceed? Do we want to narrow the curriculum or do we want to sustain the existing philosophy of a cafeteria style? (Legislative minutes, House Education Oversight Committee, 7/12/89).

Analysis Turns to Blame

The first place the General Assembly looked for answers to questions about the lack of improved performance was the State Board of Education. Rep. Shelby Rhinehart laid the fault of funding disparities at the feet of Dr. Poulton:

You are charged, the State Board of Education is charged with recommending the funding. But in my research, I cannot find in any year that you all have recommended the state increase funds to any system in order to fund these [state mandates]. When you have done that, of course, you're charged with coming up with the formula to do it. You have not made the formula right that would cover those things. You just plugged the same amount of money you had before which made more and more disparities. It seems to me that the Board of Education is a lot of the fault in this.

Dr. Poulton replied:

Well, as a matter of fact, the report that we're required to submit to you January 1 of each year titled, "The Funding Needs for the Public Schools," identified exactly what we need to get the job done. And we presented that to you. . . If you look at the recommendations we made in the funding needs report in 1989, if by chance we would have done that in Tennessee, not only would there be less disparity among the school systems in funding but there would be adequate resources. . .

Rep. Rhinehart then gave an example of how the current formula was creating disparities between his county and another county and that the funding formula was "probably the most asinine piece of work the General Assembly ever did" (Legislative minutes, House Education Oversight Committee, 7/12/89). Poulton agreed that the formula was not meeting the state's needs and that it should be replaced.

Reminiscent of something a frustrated parent might say to a headstrong child, Sen. Milton Hamilton said to Nelson Andrews, Chairman of the State Board of Education:

When Lieutenant Gov. (John) Wilder came to me and said, "Hamilton we need to sit down and we need to put something in place to where we can get some recommendations from an independent source that can tell the legislature what we need for public education," we created you. And the reason we created you was. . . to create those recommendations. Now if you don't, this is what we're coming to. We're going to continue to have the governors, legislatures, the corporate heads. . . [making] the emotional, immediate decisions that are going to have to be made for public education.

Chairman Andrews replied,

I can remember one of the first things that happened when I got this job. . . Gov. Alexander was there and he said, "The legislature has done its job. The legislature has passed the Better Schools program and you're not going to hear from us for four or five years. Don't worry about the legislature." The next year we had 125 [education] bills that came out of this legislature. And we've had at least that many ever since. . . If you have any trust at all in the nine people that you put on your Board of Education and if you understand what we've said do, it won't be 100 percent right, then let's get about that (Legislative minutes, Senate Education Committee, 2/7/90).

While the General Assembly and the State Board were blaming each other for not fulfilling their assigned responsibilities, there also were

problems between the State Board of Education and the State Department of Education concerning assigned responsibilities. When the department and board were created in 1925, both the commissioner of education and the board members were appointed by the governor—a point of contention in the early part of the century as a move by the governor to allow for the alliance of politicians (Holt, 1938, pp. 283–288). The State Board of Education maintained that it had “no enforcement power to ensure the department adequately implements the law or the board’s rules, policies, and standards” (Snodgrass, 1990, p. 7). Roles and responsibilities had become blurred with the State Department of Education taking a more active role in policy formulation, “in part because of legislative demands for specialized programs to meet perceived constituent demands” and the State Board of Education in policy implementation “as members attempt to ensure that the policies they have adopted are actually carried out” (Snodgrass, Baker, & Detch, 1994, p. 19).

The public accused the General Assembly of being more interested in politics than progress. “Three years ago the state’s prison system was in court. Now it is the schools. If Tennessee had more leadership in the General Assembly, it wouldn’t have to spend so much time pleading its case before a judge” (*Tennessean*, 7/13/88). It was also reported in the *Tennessean* “that many state officials privately hope [the lawsuit] will be lost so the courts, rather than politicians, can be blamed for tax increases that will be necessary to equalize education funding across Tennessee” (11/27/88).

The funding formula consultant, Donald Thomas, said that in the current education system “it is difficult to hold any agency accountable for education results” because there was no clear line of authority. Thomas said,

You've got an elected superintendent who can blame the board. You've got a board that can blame the county commission. You've got a county commission that can blame both the superintendent and the board. What you've got to do is concentrate the dollars under the autonomy of the local school board. You've got to give the board the authority to hire the superintendent, give him a term contract and hold him accountable. Then if the results don't come, you know where the accountability lies—with the school board (*Tennessean*, 9/24/90).

Before blame turned to evidence in the forms of a performance audit by the Comptroller's Office and the Chancery Court's decision, everyone from students to the governor was blamed for the poor shape of education in Tennessee and the only thing that everyone could agree on was that accountability had to be the "key" or the "heart" of any new education legislation. TEA President, Ann Robertson, challenged policymakers' basic assumption that teachers are ultimately responsible for improved student performance when she said, "Accountability is an interesting word. Generally, when it is used in connection with teachers, it is synonymous with blame" (*Tennessean*, 11/3/89). In November of 1988 the State Board of Education approved the BEP to improve education funding in Tennessee, with Poulton saying, "The plan should ensure that children will receive a basic education no matter where they live. As it now stands, quality of education is based partly on each city or county's ability to supplement funding that educators say is inadequate" (*Banner*, 11/18/88).

Blame Turns to Evidence: The Performance Audit and the Court Decision

In February of 1990, W. R. Snodgrass, Comptroller of the Treasury, submitted a performance audit to the General Assembly. The report was

"intended to aid the Joint Government Operations Committee in its review to determine whether the department and the board should be continued, abolished, or restructured" (Snodgrass, 1990, attached letter). The audit would make official problems and the rumors of problems at the top of the accountability hierarchy. The performance audit produced the following major conclusions:

- 1) The department does not periodically inspect schools to ensure that school systems meet the State Board of Education's minimum rules" (p. i).

Though not required by the rules and regulations of the State Board of Education, the department did not follow up on reported or ambiguously reported deficiencies in school systems meeting compliance with State Board standards.

- 2) The department approval ratings are apparently inappropriate classifications because they do not necessarily indicate the degree of compliance with state rules. As a result, the ratings do not inform the department of problem systems, do not provide school systems with an incentive to strive for full approval, and do not give the public accurate information on its school" (p. i-ii).

Again, a lack of specificity in the State Board of Education's rules was cited in the report as a "contributing factor."

- 3) The department takes minimal action on identified deficiencies (p. ii).

An interesting, additional finding of the audit in this area was that "although authorized by law to withhold state funds from noncomplying school systems, the department rarely does" (p. ii).

- 4) The department has not developed sufficient information systems to carry out its regulatory functions. . . Such information is necessary to monitor the status of the deficiencies, to decide on corrective action, and to evaluate the department's effectiveness. . . (p. iii).

All of these concepts are based on the use of negative feedback for monitoring and holding the system to set standards.

- 5) The department, with the exception of the Office of Special Education, does not have a formal complaint-handling system to ensure that complaints are resolved and to determine possible deficiencies in school systems" (p. iii).

- 6) More than half of the statutory goals of the Better Schools program (passed in 1984) will be met by the April 1989 statutory deadline. Of the 39 specific educational performance measures, 23 (59 percent) will be met but 16 (41 percent) will not be met" (p. iii).

This particular finding (6) would have an impact on the development of the EIA in terms of the accountability indicators used to show progress in meeting goals. The goal to decrease the dropout rate was not met. The report indicated that because of the measurement the department used, dropout rates were either overestimated or underestimated (p. 34). In meeting the goal to reduce proficiency test failure, 44 percent of the school systems in the state "did not improve their pass rates for both parts—mathematics and language arts— of the exam from 1985 to 1988—35 percent actually decreased (p. 37).

To determine readiness for college entry, student performance above the national average on the American College Test (ACT) and the Scholastic Aptitude Test (SAT) was the goal. Findings revealed that goals for students' scores on both subtests of the SAT and the English subtest of the ACT were

already met when the goal was established. The scores on the natural science, social science, and mathematics subtests of the ACT would not be met by 1989. The report's additional findings were that "44 percent of Tennessee's 1986 high school class attended college" and that "47 percent of Tennessee freshmen students in Board of Regents institutions were enrolled in remedial or development courses" in 1987 (p. 40).

The goal for a 15 percent increase in the number of students achieving mastery on the Basic Skills First reading and mathematics tests in kindergarten through eighth grade would be met partially. The report issued a caveat about any improvement in these scores:

Representatives of the department have expressed concern about the possibility that improvements in test scores are attributable, at least in part, to "teachers teaching the test." This means that teachers teach specific items they know will be covered on the test instead of general concepts. The effect would be inflated scores, but an inaccurate measure of the students' real learning progress (p. 48).

The Stanford Achievement Test goal to score above the national average would be met in all content and skill areas except "listening skills." Again, findings showed that nine of the 17 subtest scores were already above the national average when the goal was set and that, in fact, "most states show above-average scores on nationally normed achievement tests" (p. 48). Similar results across the nation gave rise to the expression "The Lake Woebegon Effect," taken from Garrison Keillor's mythical town where "all the children are above average."

In meeting the goals of keeping good teachers in teaching and decreasing the number of waivers that allow teachers to teach classes outside

of their certified subject areas, the state failed. "Both the number and the percent of teachers leaving teaching service more than doubled from school year 1985-1987 (p. 51). The number of teachers with waivers had increased less than one percent but had not been eliminated. Clearly CERA was not having the impact on academic achievement that Governor Alexander and the General Assembly had hoped for, especially with Career Ladder being an incentive program for improved performance. Another finding was that

- 7) Funds available for public education vary considerably from school district to school district in Tennessee (p. iv).

The State Board of Education had developed the BEP to narrow the funding disparities across counties in the state by establishing the basic elements for school operation, by shifting more of the funding burden to the state, and by assigning the local funding share according to the local board's ability to raise revenue. Findings of the audit performance concluded that "the board's proposal will not resolve the disparity in dollars expended per pupil among Tennessee's local school systems because it leaves in place the existing tax structure with its disparities in tax rates and its opportunities for additional taxation" (p. 69).

- 8) Student test scores are not strongly linked to how much money is spent per pupil by school systems in Tennessee (p. v).

To the already boiling stew of unmet educational reform goals and the need for tax reform, the audit performance added heat with this finding. The report cited many factors influencing student performance with the greatest factor being socioeconomic, such as the income and education levels of parents (p. v). The State Department of Education and the State Board of Education concurred with the finding.

- 9) School systems with comparable numbers of students had varying administrative costs per student (p. v).

This was an efficiency item, with the expected finding being that administrative costs per student would decrease as the number of students went up. The actual finding was that small school systems did not always have higher administrative costs per pupil (p. 72).

As with the recommendations of the outside consultant, Donald Thomas, the recommendations of the performance audit would find a home in the EIA of 1992.

####

The findings of the performance audit were issued between the Kentucky funding trial and the Tennessee funding trial—the legal maneuvering of which began in July 1988 and was amended, motioned, delayed, portions dismissed, appealed, bifurcated and finally decided in Davidson County Chancery Court in 1991—creating an environment of weighing the odds between increasing taxes to pay for changes to fund the BEP or delaying legislation with the expectation that the court would hand down a decision in favor of the state. The strategy of choice, ultimately, was “take it to the people,” but not before some planning had been done.

Evidence Turns to Plans

During the time that the performance audit was completed and the trial was initiated, the State Department of Education under the direction of Commissioner Smith was engaged in developing the goals that would show taxpayers the progress made with money spent on education. These goals had

been specifically requested by the legislature (*Banner*, 9/25/89). On the heels of the performance audit report, Smith presented his 21st Century Challenge Plan to the legislature's Education Oversight Committee. The introduction to the plan is a message portraying the consensual goal of "preparing tomorrow's society to compete in a global environment." Highlights of the introduction include the following:

Education is increasingly becoming a top priority for more and more Tennesseans. It has become everybody's business. . . The simple reason is that the youngsters in our schools today and tomorrow represent our state's future. It is their education which will determine in large measure whether our state improves its per capita income and enlarges its industrial base. It is their education which will determine whether our communities improve the quality of life of our citizens. It is their education which will determine whether we increase or decrease our public welfare rolls and our prison population. . .

The educational system we have today is not significantly different from that in place in Tennessee at the midpoint of this century. . . we are seeking to prepare today's youth for tomorrow in schools that were designed for yesterday. . . The world in which we live today is quite different. Today America's economic power is being challenged and tested by Japan, Korea, and many other countries. Today industries are demanding an educated work force as a prerequisite for any plant location. . . No longer is attainment of an education simply a personal goal, valued only by a student and his or her family. It has become—out of necessity—a goal of society as a whole. This is true because the young people who pass through our schools today will be our communities' voters of the future. They will become our communities' work force of the future. And if they fail, they will become our communities' welfare recipients and jail inmates of the future. . . (Smith, 1990, p. 1-4)

The statewide goals for educational excellence—the goals that would lead us to an enlightened citizenry and away from the caves of ignorance, violence, and poverty were as follows: By no later than the first day of the 21st century:

- 1) All entering first grade students—rich and poor, black and white, urban and rural, gifted and disabled—shall be prepared to achieve at the first grade level.
- 2) Every child who completes the third grade shall be prepared to read and write and solve mathematical problems effectively at the fourth grade level. Children with identified disabilities shall be expected to progress in accord with their IEPs (Individualized Educational Program) established by their M-Teams.
- 3) Achievement levels of Tennessee students shall exceed the national average and be in the top one-third of the states in the Southeast.
- 4) The statewide high school completion rate shall be at least 80 percent.
- 5) Teaching shall be a profession of choice for a significant portion of the best and brightest graduates of the state's colleges and universities.
- 6) No less than 90 percent of the Tennessee adult population shall be literate.
- 7) All students graduating with an emphasis in vocational education shall possess the competencies required to compete effectively in the job market within their geographical area of choice and/or to succeed in postsecondary technical education.
- 8) All local school districts shall be prepared to demonstrate conclusively improved performance and productivity.
- 9) All students admitted to state universities shall be prepared to begin college-level work.
- 10) A comprehensive approach to funding schools shall be in place and such an approach shall be linked directly to goals and tied to accountability standards for quality and productivity.
- 11) School-based decision making shall be the rule rather than the exception in all school districts of the state.
- 12) The Tennessee school curriculum shall be on the cutting edge of knowledge and fully responsive to the vocational, academic, and special education needs of all students as well as the employment needs of this state's businesses and industries (Smith, 1990, p. 4-44).

At the time these goals were announced to the Education Oversight Committee, some of the objectives to the goals were in the advanced planning stages; some were being pilot tested; and some had been implemented. Assessment and evaluation were objectives in all three stages

of development. Behind goal number two was the new statewide testing program Tennessee Comprehensive Assessment Program (TCAP) begun in 1989-90. It would provide a "systematic" assessment for the purpose of providing "diagnostic data" to identify and correct "student, teacher, or curriculum deficiencies" (Smith, 1990, p. 9-10).

And, then, behind goal number three were the details of the assessment plan. The "systemic" assessment would use nationally normed, standardized instruments emphasizing the diagnostic use of test results. A "value-added" component would be developed to "determine more precisely the degree to which schools are increasing knowledge and comprehension among students they serve" (Smith, 1990, p. 14-15). Training in this process would be provided by the state testing officials. Though the basis for the new value-added assessment would be nationally normed, standardized tests—something policymakers had been using since the 1960s—traditional teaching techniques and strategies for all disciplines would be assessed to determine current and future validity.

Behind goal number twelve was the new role of the teacher or what the new role of the teacher was not going to be—"dispensers of information and testers of student memories" (Smith, 1990, p. 46). Technology was going to replace past circumstances of large class size and inadequate teaching materials that had limited the teachers' roles to "dispensers and testers," Colleges, universities, and a revamped staff development program would provide new knowledge and strategies for effective teaching. All of these objectives became part of the EIA. What the 21st Century Challenge Plan created was an external evaluation system regulated by external agencies to

provide diagnostic information for teachers to use after they had received external training.

Planning for these goals had built a basis for documenting them in the 21st Century Challenge for future improvement. For example, the McWherter administration had begun a primary grades class reduction initiative based on the first year of data from Project STAR (Student/Teacher Academic Ratio) that "explained about 3 percent of the total variation" in kindergarten students' performance was the result of reduced class size (Tennessee State Board of Education minutes, 1/87).

Another example was the development of performance audit teams at the State Department of Education. These teams would visit school systems to "determine firsthand whether state funded programs are being operated in accordance with guidelines and whether prescribed goals and objectives [were] being fulfilled" (Smith, 1990, p. 39). The new statewide testing program (TCAP) began in 1989-90. And, in June 1989, "the McWherter administration successfully pushed State Board of Education adoption of an unprecedented requirement that local boards of education develop five-year plans with quantifiable goals and objectives" (Smith, 1990, p. 31). A number of incremental changes were made in moving the mechanism in place for the next education reform package.

Other newsworthy developments nationally and at the state level during 1989 were that Albert Shanker, president of the American Federation of Teachers, called for "an immediate end to multiple-choice, standardized testing, saying the results are misleading and may even be hindering school reform" for fear of "dragging down" test scores (*Tennessean*, 10/29/89). This heightened concerns of legislators who questioned the use of standardized

tests. TEA said it was willing to back tax reform including an income tax to support reducing class size (*Tennessean*, 11/3/89). And a University of Tennessee poll showed 61 percent of the state's residents were opposed to a state income tax.

Reminiscent of the education campaigns at the beginning of the century through which the need for public education was taken to the people of the state, thereby influencing the legislators, the governor announced a new statewide tour across the state to get input on the 21st Century Challenge Plan. The difference was now it was the governor leading the charge rather than the impassioned schoolmen. "We are going to sell the program. If it's worthy then the people will buy it and agree to fund it" (McWherter, *Banner*, 3/27/90). The 21st Century Challenge Plan would be called by some "the people's document." The issues of the proposed education reform were not that different from those of the legislation in the 1903-1913 education campaigns—education as a child's right, local autonomy and responsibility, and substantial financial aid from the state.

The State Board of Education accepted the goals of the 21st Century Challenge and said the next step would be to merge it with their own Master Plan. The development of the Master Plan was outlined as a duty of the State Board of Education in the reform acts of 1984. The plan would "provide a comprehensive assessment of the state's needs with regard to education reform" (Snodgrass, Baker & Detch, 1994, p.12). Work was begun to merge the Master Plan and the 21st Century Challenge Plan. The missing piece now was accountability and Donald Thomas provided that piece with his "Tennessee Education for the 21st Century" report. The integrated plan was "organized into the Key Result Areas and major building blocks presented in the Plan.

The Plan further lists Strategies, Implementation Schedules, Costs, and Indicators of Progress associated with each of the strategies" (State Board of Education, November 1990, p. 3).

At President Bush's 1990 Education Summit, accountability emerged as the major theme, with the President recommending measures for accountability such as annual report cards and federal, state and local goals (*Banner*, 9/20/90). McWherter was re-elected as governor in November of 1990 and the General Assembly was introduced to the education reform bill (House Bill No. 752 and Senate Bill No. 1235) in January of 1991.

Dr. Poulton made his introductory remarks about the education reform plan to the Senate Education Committee:

We are far from being able to say we are producing competitive graduates. . . but within a matter of months we will be able to present to anyone a business-like Master Plan that's centered on key goals, measurable goals, for the coming decade. . . We've made the commitment to review all the regulations to make sure we have constructive regulations—regulations that guide, direct, but don't encumber local operations. We intend to bring to you a request to make some amendments to the laws because there are laws which don't help local operations, only confuse. These have been accumulated over the past 50 years. What we want to move to is good, clear, constructive regulations which both encourage performance and provide a means for accountability. . . It is more a question of when and how more than it is a question of what needs to be done. Neither the State Board of Education nor the General Assembly can expect the state to move ahead without additional resources. I urge the General Assembly to engage in this issue. . . This process needs to trigger an enthusiasm for a goal-oriented operation at the local level. We need to nurture that. We can't impose this on folks and expect it to work as well as we need it to work. . . The value of this is a means of doing what they want to do (Legislative minutes, Senate Education Committee, 1/31/90).

Plans Turn to the Testimony of Dr. William Sanders

After the reform bills were introduced, hearings were conducted for several months in both the Senate and the House of Representatives. Dr. William Sanders, a statistician from the University of Tennessee-Knoxville, had been working on a "value-added" statistical model for assessment. Dr. Sanders had been consulting with Commissioner Smith, the State Board of Education, the governor, and members of the General Assembly concerning the use of assessment for accountability purposes. Dr. Sanders' model provided the link to achieve a coherent legislative package for education reform.

The value-added assessment model "analyzes the scale scores students make over a period of three to five years on the norm-referenced items on the TCAP" (Sanders & Horn, 1995, p. 2). Instead of using stanines or percentile scores to rank students against a norming group, the scale scores are statistically aggregated to show the influence of teachers, schools, and school systems on the rate of student learning. "Scale scores are designed to increase from year to year as the student learns" (Sanders & Horn, 1995, p. 2).

This model closed the circle in forming an accountability structure that tied together curriculum, textbooks, and tests to student performance; student performance to teacher performance; teacher performance to quantifiable goals; quantifiable goals to clear lines of authority and responsibility from the Commissioner of Education and the State Board of Education to local boards of education, superintendents, and principals; and clear lines of authority and responsibility to incentives and sanctions for accomplishing the goals. It provided the justification the public and, therefore, the legislature needed for not only an equitable funding formula but also increased funding to

education. Commissioner Smith outlined for legislators what happens when there is no accountability:

We want to restructure our schools in ways that will permit them to be run more like successful businesses. . . We are blessed across this state with an outstanding and dedicated core of teachers and administrators. . . The people are not the problem. The real culprit is the structure. . . In far too many communities, schools are being run in a state of near anarchy—no one really in charge, no one truly accountable. But should we be surprised about such a circumstance? That's what happens when policymaking boards at the local level get involved in day-to-day management and personnel matters. That's what happens when governing boards have no control over the selection of the chief executive officer. That's what happens when superintendents have their hands tied in day-to-day management judgments. That's what happens when front-line administrators have no effective voice in governance. . . I have seen. . . the hard cold reality of the chaos that results when the school structure collapses under the weight of disorganization (Legislative minutes, House Education Committee, 2/12/91).

Value-added assessment was the information process that would help align the lines of authority in the new structure for the 21st century schools.

Sanders spent months in Nashville explaining his model to individuals, small groups, and large groups of educators and legislators. His testimony to the legislative education committees included a presentation of the model based on data from the Knox County school system, the Blount County school system, and the Chattanooga City school system. Throughout his testimony, he emphasized the finding that most of the teachers in the state were doing an outstanding job and that his model could show that. The following exchange was typical of his testimony to the committees:

Dr. Sanders: We believe that an individual teacher or an individual school is only responsible for the improvement that

a child makes while that child is in that individual class. . . Teachers and schools. . . cannot be responsible for all the environmental influences that create differences among abilities of children. They are only responsible for the improvement that child makes. . . What we found is there's virtually no relationship between the ability levels of the kids and the gain [in academic achievement] except for the teacher effects. The teacher effects are the single most important factor that we have found determining how much the children gain or how much they learn in the course of the year. Statistically, it's rather easy to fairly adjust for the ability levels of the children to give you a fair assessment of the educational outcome without worrying about what the ability levels of the children are. . . Educational outcome assessment using student achievement data can indeed be done fairly, objectively, and unbiasedly. . . What are the primary findings? Most teachers are doing a good job. The effects of teachers on rates of student learning are highly, highly significant. Teacher effects have a high degree of repeatability over years. Outstanding teachers can be identified over a wide diversity of socioeconomic strata. . . differences among teachers tend to magnify as grade level increases. . . School effects are not related to the racial composition of the school's student body if measured this way. . . [This value-added assessment process] is easier in the elementary schools because things are more focused but when you get to secondary education you've got so much diversity in the curriculum.

Sen. (Tommy) Burks: . . . if we're going to gather all this information and not know on the front end what we're going to do with it, then, we're just gathering more statistical information to go in a file. But that's been our problem and it's still our problem. We know what we need to do, but do we do it?

Sen. (Douglas) Henry: . . . I think we all agree that education is at the heart of our social and economic development. I think we all agree that this is something that money alone is not the total answer. . . We're going to have to take some bold conceptual leaps of faith to do what needs to be done in this state and, indeed, in this country. We're going to have to be open to information that you're gleaning. . .

Sanders: I agree with the critics, it's not fair to compare the raw averages. . . But it is fair, and we're showing that, to compare how much improvement that a child made in that teacher's

room regardless of the ability level of that child. . . I hope we would have two or three years or four years to fine-tune it and not go overboard with it. Don't attach merit pay to it initially or things like that. What we want to do is let the teachers know where they profile. . .

Sen. (Leonard) Dunavant: Maybe it will be a self-winnowing factor to cause those teachers to improve and do better and find ways to do better and if they can't to eliminate themselves.

Rep. (John) Bragg: . . . we're studying education to death, this General Assembly is just going crazy trying to figure out when school ought to start, how you get a superintendent on board, how long a student should work at Hardee's for the week and how late should they work? And we're spending all our time worrying about those things, it appears, without looking at the main objects. We had testimony the other day there were just three things we really needed to worry about in education. . . 1) we get the best teachers we can find and we pay them as professionals, 2) we keep instructional time sacred, 3) we adapt our work to the culture of the community. Does that sound like a pretty good guideline to follow?

Sanders: . . . That sounds like very good guidelines. Dr. McLean and I come before you today speaking not as experts in education administration. We come as two statisticians saying you can measure it. But I think the results of what we're showing here is that the time that teacher spends in that classroom is powerful.

Rep. Bragg: I appreciate what you're doing because for too long we've only worried about what the input in education is without . . . trying to figure out what the output is. And I appreciate your telling us it can be done (Legislative minutes, House Education Committee, 1/23/90).

Dr. Sanders was to have many such exchanges with the legislators through the winter and into the spring.

Testimony Turns to Reactions and Preaching

Throughout the testimony, disagreements persisted concerning the need to appoint all school superintendents, consolidating districts, how to evaluate the performance of special education students, funding appropriations, and others. In effect, the amending process was like a fine-tooth combing of the education reform bill and everyone had a slightly different opinion as to where the hairs should be parted. A sample of exchanges revealed the different reactions participants had to the idea of value-added assessment and the larger idea of accountability.

Though the Tennessee Organization of School Superintendents would not take an official stand on appointed superintendents, Mike Simmons, appointed Johnson City superintendent said:

From a management/accountability standpoint, the superintendent is in a better position, the school board's in a better position, the entire state's in a better position if it is set up as appointed. . . That does not remove politics from education. . . If that's the intent, we can forget it because we'll always have politics in education. . . 98 percent of the superintendents in the country are appointed. (Legislative minutes, House Education Committee, 2/19/91)

Rep. Bragg responded jokingly, "There are about 15,000 school districts in this country. Three hundred and forty-six of them [superintendents] are elected—79 in Tennessee. But all of the elected live in the old confederacy. It may be hard to do any reconstruction" (Legislative minutes, House Education Committee, 2/19/91)—shades of resistance to outside recommendations reminiscent of the reaction of the legislature of 1869–1870 to the "carpetbagger" administration of Governor Brownlow.

Dewey Branstetter, representing the Tennessee School Boards Association, agreed with tough accountability but explained to legislators how hard it was to dismiss an incompetent teacher because the current system of evaluation "is geared toward improving the delivery of instruction and improving a teacher's capabilities." Sen. Henry asked him how the bill could change that problem. Branstetter answered:

The kind of accountability that is required under this act and the kind that the people in Tennessee want to see is, if you elect or have a school board that appoints the superintendent, the school board sets the policy and the superintendent carries out the policy . . . we feel that there should be more accountability on down the line, not just holding the school system and the school board and the superintendent [accountable]. But, perhaps, if you do remove the school board and the superintendent, then maybe, perhaps, you should cancel all the principals' contracts, teachers' contracts and look at holding those further-down-the-line employees more accountable. . . with this value-added testing . . . you can say this third grade class has not improved but the other two have. Now, what's wrong? Is it the student mix? Is it the way the teacher's teaching? That will certainly go a long way and the authority that the superintendent has over the principal will help too because, then, the superintendent can look at the principal and say, "You're not doing your job. You may have tenure as a teacher, but I'm going to have you out as a principal, and you're going back into the classroom. So that will help. That certainly helps in that process. (Legislative minutes, Senate Education Committee, 3/6/91)

Sen. Pete Springer expressed his sentiment about the new accountability process by saying, "It seems like so much in this entire discussion—it seems like we always come back to the point of where we've got to have a mechanism to get rid of people. And we may need a mechanism to motivate people more" (Legislative minutes, Senate Education Committee, 3/6/91).

Rep. Phillip Pinion wondered out loud if all this process was really necessary and if the credibility of the corporate model was a good one for education.

In Dr. Cheshier's statement yesterday, he made it clear we have always ranked down there below in what we spend per student and yet we always consistently rank above the national average and, of course, the conclusion was we're getting pretty good results for the money we spend. . . In accountability, I don't think good teachers mind being accountable, but it looks like to me we've proven in this state that we've got good teachers and if you've got a dud in the bunch, a good principal should be able to weed that teacher out. And I'm not sure we're going to have to go through all this process because I think if we implement this program. . . the principal will have the authority if he'll utilize that authority to get rid of those bad teachers. I feel sort of strange that business is pushing all this, when I wish we could get as much out of some of our automobiles that we buy as we do out of our teachers that we have in the state already.
(Legislative minutes, House Education Committee, 4/8/91)

Commissioner Smith responded that there was a widespread perception that "we're not doing a good job and that we have a lot of bad teachers. . . This process will show differently—that the majority of the teachers in this state are doing an outstanding job." Pinion responded that the media was creating that perception (Legislative minutes, House Education Committee, 4/8/91).

TEA's president, Relzie Payton, reminded legislators of what the "bottom line" of accountability was and was not:

Everyone wants accountability. But let's face it, people want it for different reasons. Some see it as a way to avoid being held accountable themselves—a way to blame poor schools and poor performance on someone else, not me. Others see accountability as a way to punish someone—a teacher, a principal, a board, a superintendent. Still others want accountability in a business sense—show me that my tax dollars have been well spent. I want accountability in the hope that finally, people will realize

that it takes all of us working together to successfully educate a room full of children, or even an individual child. The bottom line is—this will require a collaborative partnership. . . I want to conclude by asking you to keep three things in mind as you examine and shape the accountability sections of this legislation: 1) On behalf of the teachers of Tennessee, give us the tools, the time, and the authority to do the job and THEN hold us accountable, not the other way around; 2) Make students' interests the focal point of all accountability systems; and, 3) Remember, those students are human beings, just like you and me. And just like us, they are shaped by a variety of experiences. They are not cheaply or easily molded into the adults we all hope they will become (Payton, Governor McWherter's Papers, State Library and Archives, 2/12/91).

The quietest the legislators were in the entire education reform debate was after Carol Stinson, an elementary teacher from Robertson County, spoke on accountability. After speaking of all those accountable for student success in school and of the measurement process for accountability—using achievement test scores and value-added indicators, she asked:

How will parents responsibilities be measured? How will a county commissioner's, a legislator's, the governor's responsibilities be measured in terms of value-added student achievement?. . . The business community is interested in hiring graduates who are competent in the basic academic skills yet who also possess the communication skills necessary to work cooperatively with coworkers. We ask how can an achievement test measure a teacher's effectiveness in the development of socialization skills and democratic principles which are vital to the workplace and society?. . . We question the latest idea that schools and school systems can be run as a business. We do not work with raw materials, assembly lines, production quotas, or final products. If our students miss school, we cannot put another student in his or her place. We must be sure that student receives the instruction that they missed. . . We cannot remove, dismiss, or fire a student that does not meet or maintain certain performance standards. . . We feel Tennessee teachers are doing an outstanding job with the resources they

have (Legislative minutes, House Education Committee, 4/8/91).

There were no questions. Chairman Gene Davidson said, "It takes something like this to bring it back home, to bring it to heart, to what we're all trying to do. Thank you." (Pause) "Now I think Commissioner Smith, Dr. Poulton, and Dr. Sanders are going to address accountability."

####

Policymakers began the process of amending the bill and amending the amendments in late April of 1991. There were over 130 amendments to the bill. If adopted, amendment 110 sponsored in the Senate by Chairman Albright (a similar amendment was sponsored in the House by Rep. Purcell and Chairman Davidson) would be the final piece of "the most cutting-edge accountability plan in the country," according to Albright, who dominated the discussion on this amendment. Sanders was on hand for any last-minute technical questions and did answer a few before Chairman Albright began his sermon:

You can't compare, in this test, you can't compare student to student; you can't compare school to school; and you can't compare district to district. . . What we want to know is who's teaching in a classroom. What amount of teaching is going on, and that's what we're measuring . . . [we] will look at each class, each school, each district to see who's teaching, who's positive and who's negative. If you have a consistent negative [gain] for three years you need some help. If it's in a class then the teacher needs some help. It may be a lesson plan, it may be anything, but it needs to be looked at. If it can't be helped, then we have somebody in the wrong profession because they're damaging our precious product, our children. And there needs to be some change made. You can apply this to the schools and that's why the management plan, that is going to be addressed later, has to stay in place. Somebody has to be responsible; somebody has to be

held accountable. And that's why this whole plan is so essential to Tennessee. . .

It will tell you who is teaching right and who's teaching the best and maybe we ought to learn and maybe we ought to learn from what somebody is doing right. . . We need to have teachers involved in what we teach and how we teach it. This is the best way to do it. But how do you tell people otherwise. "I'm doing a good job." "Well, you think you're doing a good job, we don't think it." This will prove you're doing a good job . . . It'll tell us all kind of things just statistically. I'm not an expert and I can't tell you how to teach in the classroom. I don't know what's the best method. . . But I know one thing. I know how to measure. And we ought to be able to measure what is being taught . . .

I could go further on what this could do, how many millions of dollars this could save us on subjective, subjective evaluations . . . This takes out a lot of subjectiveness out of the teaching. This tells us who's teaching. The same thing take it to the next level. It tells us which school—and under this bill, the schools and the education system itself will have a report card. And on that report card it will tell, how good a school's doing. How the principal is going to have a lot to do with running that school, and that's different than we've ever had in this state and he's running it. Because he's going to be held responsible for what happens. If he has a school that's negative and continues to be negative, he has some problems. And I'm sure he's gonna be looking to see where he can get some help to solve those problems. He's got the results of every teacher in his positive or negative teaching and those that have the negative effect could have a negative effect on his school. Carry it forward even into the schools—each school. If you have too many negative schools you are gonna have a negative district. . . .

This is not entirely brand new. South Carolina has been trying to do it. South Carolina got on 20/20. They were teaching to the test and some of you are concerned about that and I'll address that. I want you to teach to the test because that's the curriculum. . . but you are not going to have a copy of the test and you are not going to have any answers to the test. The test is drawn over the curriculum. Teach to it because if you teach to the curriculum that's your job. . . we have to maintain the security of these tests . . . you can lose your teacher's license. There was [some] thought of even going to criminal penalties on that. This system has got to work. . .

This will put Tennessee, not competing with others, it'll put us way ahead of others. And its much cheaper than any other evaluation that we've attempted to do in the state. It takes out the subjectiveness of one individual judging another individual. It puts it on statistical sound information and that's what this amendment does . . . I'm sorry I get so carried away that I want to tell every citizen in the state how good it is. . . ladies and gentlemen, and if you don't have this, I'll tell you where we are heading. We are heading down national testing and we are headed down California Achievement Test and somebody else who sets the goal and we're gonna be judged in our state and you teachers are gonna be judged on what you do academically. Because the public is absolutely on all of our backs. If you don't believe it, talk to 'em. "We want a change." . . . I want to save public education and I want them to show how good they are. This testing method will do it and it's fair (Legislative minutes, Senate Education Committee, 4/17/91).

The amendment was adopted and the rest, as they say, was history.

The bill was approved by both committees and went to the Senate and the House. Because the two houses could not reach an agreement, the bill was sent to the House and Senate Conference Committee. That committee conducted hearings on the bill in the summer and fall of 1991. A special session of the legislature sent revisions of the bill back to the House and Senate Education Committees, which passed the new version. Agreement was reached in the conference committee and the bill was approved by the General Assembly and signed by Gov. McWherter on March 11, 1992.

The Resulting Action by the General Assembly

The system was poised, much like the Kentucky system had been, at the edge of chaos, ready for transformation, to be moved far from equilibrium by judicial and legislative forces of reform. Expectations were there—

Tennessee schools were to be among the best in the nation. An accountability plan second to none in the nation was in place. A testing program no other state would have had been developed. A massive effort by policymakers to sell the plan to the people had been launched. A refocused effort on education reform had occurred through the small school systems' lawsuit and the proposed state income tax. What happened? The beliefs about teaching, learning, and what constitutes improvement—measured as in the past by higher test scores—remained the same.

The resulting document from the interaction of educators, legislators, business and industry leaders, taxpayers, consultants, researchers, and legal officials was only an artifact of the continuation of equilibrium-seeking processes—processes to maintain a top-down management system tied to specific requirements for performance on the standardized test of the TCAP. The document itself reveals the policymakers' basic assumption that the teaching learning process is linear in that the inputs of the process determine, proportionally, the results of the process if the operations of the system are "efficiently" and "effectively" managed. The production-function model of education was alive and well in Tennessee.

If we refer again to Ballantine's (1993) systems model of education, (Chapter 2) nothing was significantly changed by the new legislation to perturb the system to a higher level of complexity. To determine the output of the system—improved student performance—the input would consist of student responses to multiple-choice, standardized test items.

The metrics chosen to measure student learning must be linear scales covering the total range of topics covered in the approved curriculum to minimize ceiling and floor effects. These metrics should have a strong relationship to the core curriculum for the

applicable grade level and subject (*Tennessee Public Acts*, 1992, p. 26).

Nothing new here except for the way it would be measured—"a mixed model methodology"—and reported—scores would compare student performance against their last year's performance to determine academic gain. Both of these were more fair to teachers if the goal was to improve standardized test performance.

To ensure that teachers would not "teach to the test," the legislation required that "all tests used shall be fresh, non-redundant equivalent tests, replaced each year" (p. 26). If the test items are congruent with the textbooks and the textbooks are redundant from year to year, will test items then also be redundant? A review of the selected reading book appendices for grade levels six, seven, and eight in an Anderson County, Tennessee, middle school indicates high levels of redundancy in the use of literary terms, comprehension and thinking skills, vocabulary skills, guidelines for study and research, and guidelines for process writing (See Table 1).

Table 1: Rate of redundancy among skills from grade level to grade level

Skills	% of content repeated from 6th to 7th grade	% of content repeated from 7th to 8th grade
Literary Terms	76	83
Comprehension and Thinking Skills	84	81
Vocabulary Skills	100	100
Study and Research	100	100
Process Writing	100	100

Though the reading selections are different from grade to grade, a comparison of grade-level reading books shows the rate of redundancy among skills and concepts from grade level to grade level.

The bill also supported "local Boards of Education to operate ungraded and/or unstructured classes in the primary grades, kindergarten through three (K-3)" (*Tennessee Public Acts*, 1992, p. 35), yet, "the goal is for all school districts to have mean gain for each measurable academic subject within each grade greater than or equal to the gain of the national norms" (*Tennessee Public Acts*, 1992, p. 24). How would an ungraded or unstructured classroom be measured for goal attainment? The bill supported " 'break-the-mold schools', contingent upon federal implementation of such schools" (p. 47); however, the school could be discontinued by the Commissioner if it didn't meet its approved (by the Commissioner) plan "including pupil performance standards" (p. 47). There are no "break-the-mold" schools in Tennessee at this time (Personal Communication, Ken Nye, 7/25/96). These examples highlight that the possibility for a difference in structure to occur at the school level was constrained by the use of negative feedback. The performance of the system was controlled for predictable outcomes.

This negative feedback loop was then reinforced with incentives for compliance and sanctions for noncompliance with the goal. The incentives were to "reward those schools that exceed the performance standards recommended by the Commissioner and set by the State Board of Education" and were to be given as grants in addition to money the school would receive according to the Tennessee BEP (*Tennessee Public Acts*, 1992, p. 23). The sanctions, on the other hand, would be a process of probation, remediation by the State Department of Education, and, then, based on compliance, either

removal from probationary status or removal of the superintendent and/or some or all of the local board of education members. Appeals were to be made to the chancery court of Davidson County, but the review process was "confined to whether or not the decision [for sanctions] was made in accordance with the procedures authorized by the [EIA]" (*Tennessee Public Acts*, 1992, p. 25).

While the legislation supported the idea of site-based management, it also established "clear lines of authority and responsibility," creating a barrier to the acceptance of a "shared leadership" idea. While it encouraged site-based management, if a school system did not meet the goals, ultimately, the Commissioner "shall order the removal of some or all of the board members and/or superintendent and shall declare a vacancy in the office or offices" (*Tennessee Public Acts*, 1992, p. 25). In Tennessee, about 200 of the approximate 1554 schools are using some variation of site-based management (Personal Communication, Ken Nye, 7/25/96).

Though strongly recommending an elected school board and an appointed superintendent, the legislation in the end gave people a choice and freed the legislators from the political risks of taking away people's right to vote for the superintendent of their school system until September 1, 1996. The law stated:

In those local education agencies where the director of schools or school superintendent is not currently appointed by the local Board of Education, the legislative body of any such county or municipality may, by a two-thirds vote taken within one year of elections subsequent to the 1993 election and within one year of subsequent elections, elect to retain the current method of appointing or electing such director or superintendent for an additional term of office (*Tennessee Public Acts*, 1992, p. 28).

After September 1, 1996, "unless extended by the General Assembly, all superintendents shall be appointed by the local board of education" (*Tennessee Public Acts*, 1992, p. 28). The political issue of appointed versus elected superintendent was not resolved but delayed with this clause of the bill.

The intent of the reform effort had been sold as an effort to make Tennessee more economically competitive by making those running the system more accountable for "a day when all students will arrive at high school with basic skills and leave four years later with the ability to make it in the workplace or in college" (Commissioner Smith, Legislative minutes, House Education Committee, 2-12-91). The legislation required an exit exam for students going to college and for those going to work, but "there [would] be no passing standards for this examination" (*Tennessee Public Acts*, 1992, p. 32). There is also a backing off of the "cutting-edge accountability plan" in the criteria set for performance standards at the elementary education level: a different standard is set if school districts do not have mean rates of gain equal to or greater than the national norms based on the TCAP achievement test. In this case,

Each school district is expected to make statistically significant progress toward that goal. The rate of progress within each grade and academic course, necessary to maintain compliance with this section will be established after two years of consecutive testing with tests adopted for each grade and subject as provided in subsection (g) of this section [linear scales of measure] (*Tennessee Public Acts*, 1992, p. 24).

The blade of "cutting edge accountability" was dulled with subgoals to the standards.

The multi-million dollar Project STAR, commissioned by the legislature, had shown that "students in primary grade classes with a 15-1 pupil-teacher ratio performed significantly better than those in classes with a 25-1 ratio" (Smith, 1990, p. 8). The legislation called for an average class size in K-3 to be 20 and the maximum class size no larger than 25. A disclaimer was tucked into the language of the section of the bill dealing with class size.

In no school building shall the average size of any grade level unit prescribed herein exceed the stated average, *though any individual class within that unit may exceed the average*, provided that no class shall exceed the prescribed maximum size (Tennessee Public Acts, 1992, p. 33).

Dr. Sanders maintained that in using the value-added model of assessment, "In Tennessee, success or lack of success in student academic gains appear to bear no relation to a school's average income or racial composition. Likewise, average gains do not reflect whether a school is urban, rural or suburban" (Snodgrass, Baker, Xu, & Detch, 1995, p. 51). In his study of funding for Tennessee schools, Green (1995) reports continuing disparities in per pupil expenditures across the state. Oak Ridge City School System has a per pupil expenditure of \$5,827 or 104.92 percent of the national average for per pupil expenditure, while Alamo School System in Crockett County spends \$3,006 or 54.12 percent of the national average (p. 86). Three-year average gain scores on the norm-referenced items of the TCAP achievement test as a percentage of national norm gains for Oak Ridge and Alamo School Systems are 111.9 percent and 89.2 percent, respectively (Snodgrass, Baker, Xu, & Detch, 1995, p. 34). Using the variables of per-pupil expenditure and average gain scores, a correlation study including all school

districts across Tennessee shows no significant relationship between these two variables. Sanders' model is able to control for the variables of socioeconomic status or race because he is tracking a student's progress over time and against that same student's previous school year performance (Personal Communication, Dr. William Sanders, July 8, 1996). Dr. Harry Green, the Executive Director and Research Director of the Tennessee Advisory Commission on Intergovernmental Relations (TACIR) cites the work of Larry Hedges in his discussion of school funding.

Larry Hedges at the University of Chicago has done some pioneering work on education production functions. . . [His] analysis . . . suggests strong support for the idea that education inputs—such as per pupil expenditure, teacher experience, and the pupil/teacher ratio—have an important positive effect on student achievement (Green et al, 1995, p. 28).

Another study (Kim, 1991) comparing the top twelve and the bottom twelve states' performance on the National Assessment of Educational Progress Eighth-Grade Mathematics Exam scores (1990) shows

that the percentage of students with parents who did not finish high school appears to make the largest contribution [to student achievement]. Testing frequency, high emphasis on numbers and operations (basic skills), and urban disadvantage all, interestingly, negatively contribute to student achievement. By contrast, student reading, teacher experience, and expenditures per pupil positively contribute to achievement (in Odden & Kim, 1992, p. 285).

If socioeconomic, racial, and gender factors are masked by the current use of value-added assessment in Tennessee, data are eliminated from evaluation that are found in other studies of student achievement to be critical to the improvement of learning on a child-by-child basis. What cannot be told from the value-added data as reported by the comptroller's

review of Tennessee Value-Added Assessment System (TVAAS) (Snodgrass, Baker, Xu, & Detch, 1995) or the Tenth Annual Report (1995) to the governor is what and how teacher, school, or system interventions are interrupting what Green and Kim found to be the negative relationship of educational performance to socioeconomic status, and further, what changes in educational structures and processes need to be amplified to support these interventions. These changes require other observational and evaluation methods.

The current value-added model makes use of the norm-referenced test items of the TCAP in accounting for the "effect" the teacher, school, or school district has on the educational progress of students (Sanders and Horn, 1994). Murphy (1992) reminds us that

it must also be understood that standardized norm-referenced achievement tests are not designed to assess the extent to which students have mastered the school system's curricula. Rather, they are designed to assess the extent to which students master what the test publishers deemed to be important (p. 191).

This system of assessment maintains the idea of a pre-determined or static curriculum for all students controlled by a centralized agency for efficient delivery of a product at minimum cost, unless curricular changes are reflected in the continual updating of the standardized tests being used. Outstanding student performance in meeting the state's goals is rewarded with high test scores and monetary incentives. Failure to comply with expected mean rates of gain equal to or greater than the national norms based upon the TCAP achievement test results in schools and/or systems being subject to probation. (To date, no school or system in Tennessee has been put on probation; therefore, no superintendents or board members been removed from office.)

Is the purpose of assessment in Tennessee to evaluate cost efficiency and accountability for the delivery of minimum standards or to collaboratively and contextually improve educational processes? It is only in a closed system, where the exchange of data is governed by the institution isolated from its environment, that prediction is possible based on the variables of teacher inputs and student outcomes. Can such a closed educational system exist? Is such a system desirable?

The Education Improvement Act: A Repeated Cycle of Policymaking

The tradition of using mandated assessment for compliance with performance standards began with the "compliance-assistance" policy of ESEA and moved to the "compliance-dominant" policy of P. L. 94-142 with the intervention of the court. A similar phenomenon occurred in Tennessee. CERA was designed to be a compliance-assistance policy with evaluation tied to the use of an instructional model and merit pay for outstanding performance. With the intervention of the court, what began as a fine-tuning process for the Career Ladder became a compliance-dominant policy, the EIA. When faced with a potential system overhaul like Kentucky and when achieving some success with the compliance policies of CERA, policymakers decided a more stringent compliance policy was in order. CERA was not specific enough in its goals, expectations, or lines of authority. Commissioner Smith acknowledged this desire for specifics in addressing the House Education Committee:

You have indicated to us that you would like to see more specifics in the legislation, and I can tell you that today, that we have worked with Dr. Sanders and individual members of the General Assembly, and with the State Board of Education, and

members of my staff. There will be an amendment to the existing bill that will provide more specific information and hopefully as a result of our discussions before you today and whatever later time we need to come back that you will find the language acceptable or perhaps give us additional guidance as to what it should say (Legislative minutes, House Education Committee, 4/9/91).

The EIA would become more compliance-dominant because of the increased role of the legislators in developing evaluation, management, and curriculum components of educational policy. On the other hand, the State Board of Education would become more involved with funding education as a result of developing the BEP. The trade-off was more money for more accountability, and more accountability meant a greater involvement by the legislature in designing the structure for the state public school system. The tradition of mandated assessment policy would have the effect, after 30 years, of increased specificity in the design of assessment and decreased exploration of other assessment practices available to educators. The following excerpts from the ESEA, P. L. 94-142, CERA, and the EIA show this progression. In the ESEA, evaluation is mandated:

Effective procedures, including provision for appropriate objective measurements of educational achievement, will be adopted for evaluating at least annually the effectiveness of the programs in meeting the special educational needs of educationally deprived children (*Statutes at Large*, 1965 p. 31)

In P. L. 94-142, specific procedures for the use of evaluation is mandated:

Procedures to assure that testing and evaluation materials and procedures utilized for the purposes of evaluation and placement of handicapped children will be selected and

administered so as not to be racially or culturally discriminatory (*Statutes at Large*, 1975 p. 781).

Provide for procedures for evaluation at least annually of the effectiveness of programs in meeting the educational needs of handicapped children. . . in accordance with such criteria that the Commissioner shall prescribe (*Statutes at Large*, 1975, p. 784).

The Commissioner shall measure and evaluate the impact of the program . . . and the effectiveness of State efforts to assure the free appropriate public education of all handicapped children (*Statutes at Large*, 1975, p. 791).

Provide for the evaluation of programs and projects assisted under this part through—(A) the development of effective methods and procedures for evaluation; (B) the testing and validation of such evaluation methods and procedures; and (C) conducting actual evaluation studies designed to test the effectiveness of such programs and projects (*Statutes at Large*, 1975, p. 792).

These provisions were made in light of advanced developments in diagnostic and instructional procedures for specific needs of students.

CERA provided for specific tests to be used, levels of proficiency, and a schedule for evaluation:

To study the use of student progress or achievement, as measured by standardized testing or other appropriate measures, as an indicator of successful teaching and effective schools, and to review periodically the standards and criteria used for teacher and principal evaluation in view of the findings resulting from such study (*Tennessee Public Acts*, 1984, p. 31).

A 15 percent increase in the number of students mastering each skill in reading and mathematics as measured on the Basic Skills criterion-referenced tests in grades 3, 6, and 8 (*Tennessee Public Acts*, 1984, p. 64).

It is the legislative intent that nationally normed achievement tests be selected for administration at three grade levels in the 1984-85 school year and that within five years, the statewide

average test scores achieved on such tests by Tennessee students be higher than the national average for comparable grade groups in each subject (*Tennessee Public Acts*, 1984, p. 64).

Finally, in the EIA, the specific construction of the evaluation measurement—value-added assessment—is given in statistically technical terminology along with specific research articles that explain its construction:

The metrics chosen to measure student learning must be linear scales covering the total range of topics covered in the approved curriculum to minimize ceiling and floor effects. These metrics should have a strong relationship to the core curriculum for the applicable grade level and subject (*Tennessee Public Acts*, 1992, p. 26).

The process of making assessment practices more specific without a focus on the teaching/learning process is analogous to continually replacing a thermostat with a more sensitive model without evaluating the whole heating/cooling system for productive operation. It may be more accurate in detecting fluctuations in the performance of the system but it does not give any information as to the nature of the fluctuations. The problem with using one type of academic performance measure to achieve multiple objectives, like accountability for both expenditure justification and student performance, is the probability that a macro-measurement instrument such as a standardized achievement test will be used to micro-manage the system. The test is used, then, to monitor the system for variation from the pre-set standard and then to adjust the standard rather than investigating the variation. Achieving a system for measuring results becomes a priority for policymakers—"the heart" of the education reform bill. Policymakers become a basketball team shooting for the goal by looking at the scoreboard.

The trade-off of accountability for increased and equitable funding had the effect of lowering the political risk for policymakers by decreasing the public's distrust with government intervention and the public's perception of a lack of political leadership on the part of the legislature. Table 2 shows how the dimensions of mandated improvement in student academic performance changed during the various phases of policy formulation.

Table 2: Dimensions of mandated improvements in academic performance

	Resource Intensiveness	Targeting	Institutional Capacity	Political Risk
Beginning Conditions	High	High	High	Low
Context	High	Low	Low	Moderate
Intervening Conditions	High	Low	Low	High
Strategies	High	High	High	Moderate
Consequences	Lower than Expected	High	Low	Low

Beginning Conditions

At the end of Alexander's administration, the desire to improve education in Tennessee was publicly and politically sound. Alexander continued to be hailed as the education governor and Tennessee was considered a leader in education reform. The nation as a whole was concerned about economic competition and the importance of education as the competitive edge within the global marketplace. Conditions created by

previous education mandates such as programs for disadvantaged children, programs for children with special needs, the Career Ladder Program, and the Basic Skills First program put additional financial strain on the state and local resources, making resource intensiveness high—more than 40 percent of the state's general budget was spent for education (Vickers, *Tennessee Statistical Abstract*, 1994/95).

Improvement for school systems, schools, and teachers had been designed specifically in previous mandates by using assessment for diagnostic purposes to target areas of student need. The new Basic Skills First criterion-referenced tests were the diagnostic tool aligned with the curriculum frameworks composed of specific goals and objectives by grade and subject levels. The tests would be used in identifying areas of student academic weakness or deficiencies in the teacher's teaching program.

Evaluation for teachers was designed as to an instrument that examined specific parts of a direct instructional model (based on Madeline Hunter's *Master Teaching*, 1982) in which teachers had received specific training to enable them to implement it in their classrooms. Based on number of years of experience, evaluation performance, and the desire to work additional days, teachers could apply for Career Ladder Levels I, II, or III and receive merit pay and additional money to work with special needs students before school, after school, or in the summer. The process was designed to increase the institutional capacity of the educational system by providing training, by training administrators to use the new evaluation instrument, and by using Career Level III teachers to evaluate and assist peers while at the same time relieving some of the administrative burden for evaluation. The new evaluation system was more time consuming and

paperwork intensive than past local tenure evaluation practices. With the passage of CERA, though costs were high, policymakers had increased capacity and targeted areas of need, thus improving their political favor with constituents and position in leadership in education reform for the country.

Context

The dimensions began to change as McWherter took office in 1987. Along with the demand for increased salaries, promised by Governor McWherter, teachers also were pushing for reduced class size, mandatory kindergarten, updated textbooks and equipment. All of these changes would mean increased funding for education and the public was not in the mood for another tax increase. Education study results were coming in and supporting the need for adjustments in the Career Ladder, the Basic Skills First program, and reduced class size, and showing that Tennessee had one of the highest poverty and illiteracy rates in the country. The business community was alarmed by the lack of knowledge and skill of the graduates they were hiring from Tennessee high schools. A performance audit of the Tennessee system of public education revealed ambiguous progress in meeting the goals of CERA.

Targeting improvement through the reform efforts of CERA had proved less successful than Alexander and the General Assembly had expected. First of all, Career Ladder evaluations were not always successful in distinguishing the best teachers for merit pay. Yet, the "best" teachers, those on Career Ladder II and III, could receive extended contract money for special needs students. Secondly, the school systems in wealthier districts that could pay teachers higher salaries were attracting many of the better teachers. Then

these teachers were applying for Career Ladder II and III status and receiving extended contract money. The money, therefore, was not following student need, necessarily, but was following Career Ladder II and III teachers to wealthier districts. In addition, the sporadic testing schedule of the Basic Skills First Program, the lack of a data management system, and the reporting system for deficiencies within schools and school systems, made targeting schools or school systems for improvement difficult.

Deficiencies in the evaluation system for teachers and administrators such as the lack of timely feedback on performance, the lack of guidance for improved performance, and the lack of local input in evaluation procedures diminished the institutional capacity of the system by constraining teachers' perception of improvement to "teaching the test." The administrative burden of everyone from teachers to State Department of Education and State Board of Education personnel was increased with the Career Ladder evaluation system. Teachers were asking for changes in the evaluation and testing programs. Political risk was increasing as policymakers discussed the need for additional money in terms of increased taxes or possibly a state income tax to continue the education reform started under Alexander.

Intervening Conditions

The lawsuit filed in chancery court in Tennessee to contest the constitutionality of the state's education funding formula turned the political risk of spending more money on education to high. A similar case in Kentucky resulted in the whole state's educational system being declared unconstitutional. The decision was extensive in that it assigned the state to rebuild a system that would adequately, efficiently, and equitably ensure that

all students graduated with the competencies the court, itself, declared necessary. Additional intervening conditions added fuel to the heat. A shortfall in state revenue, a regressive tax system, and an outdated funding formula put policymakers in a precarious decision-making position.

Action/Interaction Strategies

Development of a "cutting-edge" accountability plan as well as a new formula to narrow funding disparities avoided the extreme action of the Kentucky court decision while maintaining political integrity. The plan also increased the targeting of funding and evaluation by developing measurable goals for teachers, schools, and systems effects on student performance. Using the value-added assessment system in conjunction with an information management system allowed for an objective, efficient, and effective method to collect, analyze and return data to teachers for improved teaching. The plan increased the institutional capacity of the system by eliminating 3,700 rules and regulations, establishing clear lines of authority, updating the curriculum for relevancy, providing improved, relevant staff development and allowing for innovation. Political risk was lowered.

Though costs were still high, resulting in a half cent increase in sales tax, all earmarked for education, they were lower than expected since the court did not find the entire system unconstitutional. The consequence of the policymakers' massive effort was accountability for increased money, and assessment had been the magic key. No state income tax. No lost voting rights of constituents to elect superintendents. More money to poorer districts. No court order to completely rebuild the system. Outstanding teachers for the most part. And Tennessee was recognized, once again, as

having one of the most far-reaching reforms in the country. Political risk was lower still.

Unintended Outcomes

Test scores have not changed because schools have not changed. Schools need to create a meaningful learning environment.

James Pellegrino, Dean of Vanderbilt's Peabody College

The unintended consequences of the EIA were that targeting and institutional capacity only appeared to increase. The constraints of using a mandate with incentives and sanctions for compliance were discussed earlier in this chapter. Short-term political gain would lead to another court challenge to equalize teachers' salaries across the state. Using value-added gain scores in teacher evaluation was halted until an outside validity study of TVAAS was completed. A mixed review of academic progress brought education reform into question again.

In February of 1995, the Tennessee Supreme Court handed down a decision in favor of equalizing teachers' salaries across the state. "The Court ruled that omitting teachers' salaries from the BEP was a significant defect and had the potential to impair achieving the objectives of the BEP" (Green, Smith, & Hydorn, 1995, p. 9).

Established by the EIA to monitor school systems' performance in accordance with performance standards and to conduct studies the Governor, the General Assembly or the comptroller determined necessary, the Office of Education Accountability prepared a study on the TVAAS. The findings listed concerns including "unexplained variability in national norm gains

across grade levels," the security of the system, the complexity of the system, and the "high stakes" nature of the TCAP achievement test that the system was based on (Snodgrass, Baker, Xu, & Detch, 1995). Sanders provided answers to the questions posed in the study, but the final recommendation was that all "components of the TVAAS be evaluated by qualified experts knowledgeable of statistics, educational measurement, and testing (p. iv). Sanders welcomed an outside evaluation.

The three-person evaluation team was composed of R. Darrell Bock, a distinguished professor in design and analysis of educational assessment and a professor at the University of Chicago; Richard Wolfe, head of computing for the Ontario Institute for Studies in Education; and Thomas H. Fisher, director of student assessment services for the Florida Department of Education. Bock and Wolfe (1996) agreed with the central concept of the assessment system, but made recommendations for the improvement of data collection, the quality of the test scores, and reporting gain scores. They recommended not reporting the school effects "until the year-to-year stability of estimated gains for schools can be improved," but reporting teacher effects (with their recommended improvements) to "responsible authorities" (p. 34).

Fisher (1996) focused on the state testing program (TCAP) in his analysis. He found the basic structure of the testing program to be acceptable, but recommended verifying the accuracy of the process and improving the management and oversight of the process. Because of needed improvements in data collection, he recommended that the teacher effects not be reported.

All three questioned the use of national norm scores as the comparative standard for student performance progress. All commended Tennessee for its "bold steps" in working toward educational improvement,

but Fisher stated, "There is no single way to foster educational accountability. The search for acceptable and understandable methodology should be seen as an open-ended quest" (p. ii).

In the eleventh annual report (1996) by the State Board of Education to the Governor and General Assembly on student, teacher, and school improvement, results were reported over several years for performance indicators as provided for in the EIA. The attendance rate decreased from 93.8 percent in 1989-90 to 93.4 percent in 1994-95. The drop-out rate decreased from 23 percent in 1990-1991 to 17.9 percent in 1994-95. Tennessee's average 1991 ACT scores (20.1) improved by two tenths of a point (20.3) in 1995, but the national average also improved by two-tenths (from 20.6 to 20.8). Though the number of students taking Advanced Placement courses increased to 7,226 and the number of Honors Diplomas increased to 8,748, this was still a small percentage of the state's graduating students. And though 464,569 students were being served throughout the school year with extended contract money, 22 percent of the programs offered by local school systems were developmental and remedial in design. Teaching permits were up, from 161 in 1986-87 to 360 in 1994-95, and waivers were down, from 469 to 449. Table 3 shows academic progress based on Sanders (1995) graphical summaries of educational findings.

Table 3: Percentage of Tennessee schools achieving national norm gain or better

Subject	1993	1994	1995
Math	33.9%	33.9%	42.7%
Reading	55.6%	43.4%	42.6%
Language	50.4%	33.2%	67.2%
Science	38.4%	49.6%	39.5%
Social Studies	19.9%	46.5%	73.5%

From these indicators, it is not obvious what or how change had occurred in the teaching/learning process of Tennessee schools as a result of increased accountability. It may be too soon for changes to be detected in the data from these performance indicators. Nelson Andrews, chairman of the State Board of Education while the EIA was being formulated, stated his view of the pace of change in education while speaking to the Senate Education Committee:

One of the things that you talked about that's gonna cost a little bit of money is to take those things that we do right and transport them into other systems. I've said this many times. Successful things in education are the least contagious thing I have ever seen in our society. We don't infect good ideas in education at all even when it may be next door. So what we've got to have I think is a clearing house of ideas. . . In my judgment, ten years from now, if we're smart, a schoolroom's gonna look a lot different. I've said this many times. If you brought your great, great, great, great, great, grandparents back to earth today they would recognize the church and the school and probably not anything else (Legislative minutes, 2/7/90).

Conclusions, implications, and recommendations based on these findings are discussed in the next chapter.

CHAPTER 5

SUMMARY, FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

The hope of authentic reform rests in empowerment that generates action at the site of responsibility by encouraging people to assume responsibility for themselves. When [self-organization] replaces bureaucratic control, people are free to work at the edge of their competence and, consequently, to develop in ways that expand the limits of their expertise. Astuto, 1994

Summary

Currently, policymakers' use of legislated assessment practices places an emphasis on accountability for improved test scores rather than for adjusted practice. How outcomes of learning and teaching can be measured for accountability, objectivity and transportability (Bellon, 1982) is often the criterion by which teaching practice is rejected or accepted. The limiting effects of such criteria are devastating if the purpose of schools is the development of life-long learners, responsive to and creative within the environment they find themselves. If the range of content, the ways of knowing and "teaching as thinking" (Shulman, 1987, p. 16) become constrained within schools that are detached from their environment, knowledge is replicated, or at best, simulated rather than constructed as a basis for new knowledge construction.

Value-added assessment based on standardized test scores is the mandated assessment program used to measure content mastery and skill development in Tennessee. The standardized test is probably the most widely used assessment tool in determining what and how students should know. As Gardner (1991) puts it, "the test is the ultimate scholastic invention, a decontextualized measure to be employed in a setting that is itself decontextualized" (p. 132). Current "objective" assessment practices have

possibly only served the educational purposes of teaching skills or specific content (Shulman, 1987) and not that of evaluating teaching and subsequent performance for the educational purpose of adjusting parameters for self-organized learning.

The Problem

Current educational policymakers' use of mandates with incentives and sanctions for compliance and noncompliance with performance goals at or better than the national norm gains on standardized tests is constraining the evolution of the teaching/learning process. Innovative or positive changes are damped by the negative feedback created in the current assessment system, that is, to monitor the system for variation and align it with pre-set standards based on the mastery of content and skill development. To evolve to higher levels of complexity, the schooling system must develop both positive and negative feedback loops. This allows for recursion of information within the system. Information is fed back into the system to monitor basic processes as well as to provide new insights for adjusted practice in meeting the needs of all learners.

Current assessment is holding teachers accountable for improved student performance and limiting them to the use of a linear model to measure the worth or merit of a nonlinear process—the teaching-learning process. If legislators and other educational policymakers continue to view the teacher or school effect as an isolatable, independent variable constituting a linear relationship with student performance on decontextualized measures of achievement, current assessment practice will constrain the evolutionary nature of teaching and learning by limiting the process to “teaching the test.”

This study was designed to understand how legislators select policy tools, and the use of assessment in the design of these tools, by focusing on legislators' understanding of the teaching-learning process and the impact their decisions concerning assessment have on the evolution of this process.

The Purpose

The general purpose of this study was to understand the impact of legislated assessment practices on student academic performance in Tennessee. Specifically, the study was to examine the Education Improvement Act (EIA) of 1992 to determine the strengths and weaknesses of its design in terms of related theory, "best" educational practices, and its potential to advance policy objectives. As a result of the study, recommendations were made to improve legislative decision-making as it affects educational practice. This study offered answers to the following questions:

- How does the Tennessee State Legislature select a policy tool when making educational reform decisions? What role does assessment play in the design of the policy tool?
- What are the intended and unintended outcomes of legislated assessment practices in Tennessee?
- How can assessment as a component of the policy tool be developed to inform educational decision-making from the classroom to the legislative level?

Procedures

The conditions of this study met the primary requirements for using a case study research design. Documents and records were analyzed that preserve the historical context and decision-making processes of educational

policymakers in Tennessee. Specifically, the documents and records related to the development, implementation, and evaluation of the EIA of 1992 and the use of assessment as part of that mandate's design were reviewed. In an effort to stay as close as possible to the historical context of the enactment of the EIA, documents and records were collected from each major policy agency and then analyzed using a process coding within an open coding technique. To corroborate the evidence presented in the analysis of the data, the researcher used a data triangulation strategy, comparing concepts and categories defined in data from the legislative meetings across multiple-agency settings.

Findings

Major findings of this study were as follows:

1. Policymakers chose to mandate the next wave of educational reform in Tennessee using assessment to ensure accountability for results—improved scores on national achievement tests.
2. Judicial intervention created a shift in the importance and type of influences on past legislative decision-making patterns.
3. The intended outcomes of the EIA were to improve education in the state and decrease political risk. The unintended outcome was to constrain educational practice to a more traditional rather than innovative practice.
4. The basic assumptions of educational policymakers supported a mechanistic understanding of schooling systems.
5. The current assessment program strengthened the negative feedback loop in the schooling system; and
6. The EIA did not explicitly provide for a positive feedback loop.

A brief discussion of each of these findings is presented below:

1. *Policymakers chose to mandate the next wave of educational reform in Tennessee using assessment to ensure accountability for results, that is, improved scores on national achievement tests.*

With increasing involvement of Tennessee legislators in educational policymaking not only as the primary funding agency but also as the shaper of management, curriculum, and assessment practices in schools, the use of an education mandate was a matter of routine rather than a considered selection from among other policy tools. A long tradition of using mandates for educational reform undergirded the choice. When confronted with being overinvolved with legislating educational programs, policymakers dismissed the suggestion.

The role of assessment in the design of the EIA was primarily to ensure accountability for expenditures and, secondly, to ensure accountability for improved student performance. The value-added assessment system is the most specific accountability measure the General Assembly has passed to date. It was the first mandated attempt to hold teachers, schools, and systems accountable for specific improvement results on standardized tests. Incentives and sanctions were tied to performance measured by this system in comparison to national gain averages on the norm-referenced test items of the Tennessee Comprehensive Assessment Program (TCAP).

Accountability for improved student performance was measured as progress made toward performance goals and standards using the following indicators: promotion rate, dropout rate, attendance rate, and value-added gain scores. From changes in the indicators, it was difficult to determine how and to what extent teachers had adjusted their practice to improve learning.

Of all the indicators, value-added gain scores were the only designated "learning" indicators. Though they were focal points in the education improvement debate during the policy formulation phase, there is no reporting of performance indicators for learning competencies; no reporting of noted improvement or lack of improvement in graduates hired by business and industry; no reporting of the number of graduates entering remedial or developmental classes at the college level in the State Board of Education's annual report to the governor and the General Assembly.

2. Judicial intervention created a shift in the importance and type of influences on past legislative decision-making patterns.

When faced with the Tennessee Small School Systems lawsuit, the "effects a law will have," "listening to expert information," and "providing for orderly and manageable change" (Keese, 1990, p. 75) increased in importance for Tennessee legislators requiring different information and different interactions with policy actors. Legislators commissioned studies of the career ladder, extended contracts, and assessment. Dr. Donald Thomas, a consultant from Utah, helped in redesigning the funding formula for education.

With accountability emerging as the theme of the 1992 education reform, the business community and the constituents were the most important influence on policymakers, rather than having a medium influence on decision-making as found by Keese (1990) in formulating previous education legislation. The governor, college and university representatives, and the national education reform movement also gained importance as influences on policymakers during the formulation phase.

Keese found these to be among the least influential groups and individuals in previous legislative decision making.

While court intervention changed the degree of influence from outside sources and the degree of legal concern during the formulation of the EIA, legislators continued to rely on "traditional values, personal philosophy, experience or upbringing, and personal goals for education when approving or disapproving educational legislation" (Keese, 1996, p. 106).

3. The intended outcomes of the EIA were to improve education in the state and decrease political risk.

The stated intended outcome of the EIA's assessment program was to improve student performance while preparing students to become part of a workforce that would be economically competitive through increased funding and increased accountability. The unstated intended outcome of the EIA was to reduce political risk for the policymakers. The unintended outcome of the legislated assessment program was to strengthen the negative feedback loop in the educational system by constraining the performance to a uniform standard. Based on the documentary evidence of Chapter 4, the unintended outcome was a consequence of the basic assumptions policymakers held about teaching, learning, and improved performance.

The basic assumptions underlying the use of a mandate with incentives and sanctions based on performance indicators to formulate the EIA were

- The purpose of schooling (K-12) is to prepare students to be competitive members of the 21st Century workforce.

- Teachers and administrators in Tennessee schools have the capacity to improve student academic performance.
- Though affected by many variables, student academic performance is ultimately the responsibility of the classroom teachers.
- Evidence of student academic progress is increased knowledge and skill development and, therefore, a linear process—gains in progress should be proportional to the number of years in school.
- Standardized test scores are an appropriate measure of student academic performance when used as the basis for value-added assessment.
- Measurable academic performance equals learning.
- With adequate resources, clear goals, efficient management, and an external control mechanism, student academic performance improves.

4. *The basic assumptions of educational policymakers supported a mechanistic understanding of schooling systems.*

Their assumptions maintain policymakers' understanding of schools' and school system's organization as being similar to the organization of a machine. The documentary evidence reveals all three of Kronenberg's (1995) characteristics for identifying this metaphoric understanding of social organizations. First, social organizations such as schooling systems are viewed as "social objects and can be evaluated in terms of standards we apply to machines: efficiency and effectiveness in their uses of resources and other inputs" (p. 254). Throughout the education reform debate of 1992, educators, legislators, the business community, and concerned citizens were all preoccupied with the effectiveness and efficiency of schools producing a quality product for increased spending.

The second characteristic of the "machine" metaphor is that social organizations "are under control in political terms to popular will and can therefore be held accountable" (Kronenberg, p. 254-255). The 21st Century Challenge Plan, sometimes referred to as the "people's document," was fashioned by input from individuals throughout the education community and from communities across the state. The commissioner of education stated that he could point to specific parts of the plan and say who suggested the idea. In testimony before the House and Senate Education committees "accountability" and "what the people want" became popular themes for continued education reform.

"Organizations are predictable, unless they break down" is the third characteristic of the "machine" metaphor (Kronenberg, p. 255). The effort to enact education reform was to bring order, in the form of accountability, to the "hard, cold reality of chaos" in which the "school structure [was collapsing] under the weight of disorganization" (Smith, Legislative minutes, House Education Committee, 2/12/91).

The problem in constructing policy with a mechanistic understanding of schooling systems was that policymakers continued to rely on top-down management strategies, to emphasize predictable outcomes, and to de-emphasize the possibility of self-organizing systems at the local level. High stakes testing was the natural choice for assessment in order to maintain the efficiency and effectiveness of the school "machine."

5. The current assessment program strengthened the negative feedback loop in the schooling system.

Though policymakers were aware of the need for innovation (i.e. break-the-mold schools, unstructured and ungraded classrooms, site-based decision making, and family resource centers), their decision to establish a strong negative feedback loop slowed the pace of change in restructuring Tennessee schools. Because the educational system was regarded as a machine, concerns for local teaching practices effecting student improvement were minimal. These concerns were diminished by the policymaking focus on equal learning opportunity and equitable educational funding. Equitable funding throughout the state was seen as a means for reducing variability in student performance. Value-added assessment was seen as the means for monitoring that variability. Finally student output would be measurable in relation to teacher, school, and system input. Value-added assessment was based on the norm-referenced items of a national achievement test. The EIA reinforced the status quo of the educational system in Tennessee by basing student academic performance standards on a pre-set score for achievement—national gain scores on a standardized test.

6. *The EIA did not explicitly provide for a positive feedback loop.*

The EIA did not provide for the iteration of positive feedback concerning teaching practice into the system. No assessment is explicitly provided for in the EIA to amplify the nature of change brought about by innovative teaching practices at the local level. Patterns in student performance on discrete items may be detected in TCAP achievement test scores and in value-added data, but the interpretation of the patterns must come from other data sources. By mandating specific assessment practices, policymakers assumed that teachers and administrators understood the

construction of the assessment, how to look for patterns, and what to do when patterns were found. Though the Value-Added Research and Assessment Center has made a concerted effort to explain and to assist educators with data interpretation, continued use of a single learning indicator—standardized testing—to assess student performance underscored the historic concern for teachers “teaching to the test” rather than adapting practice for increased learning. The lack of information in the eleventh annual educational performance report concerning increased student learning underscored the need for more and varied assessments.

Conclusions

This research indicates that with assessment of student performance (as outlined in the EIA) being developed outside the system, the crucial element of reflection in the teaching-learning spiral (See Figure 2) has been diminished. As one element of the teaching-learning spiral diminishes, the cohesion among all the elements necessary for teaching and learning to occur at higher levels of complexity are diminished. What restructures the teaching-learning process is performance, feedback or reflection on the performance, and adjusted practice based on new insights from reflection on performance as it relates to teaching practice. Current assessment constrains practice by emphasizing content mastery and skill development over all other levels of learning.

As it stands, the EIA separates assessment from the teaching/learning process. Assessment is summative rather than formative. It is developed and scored by an agency outside the individual classrooms, creating a delay in feedback to the process. It measures learning at the level of learning

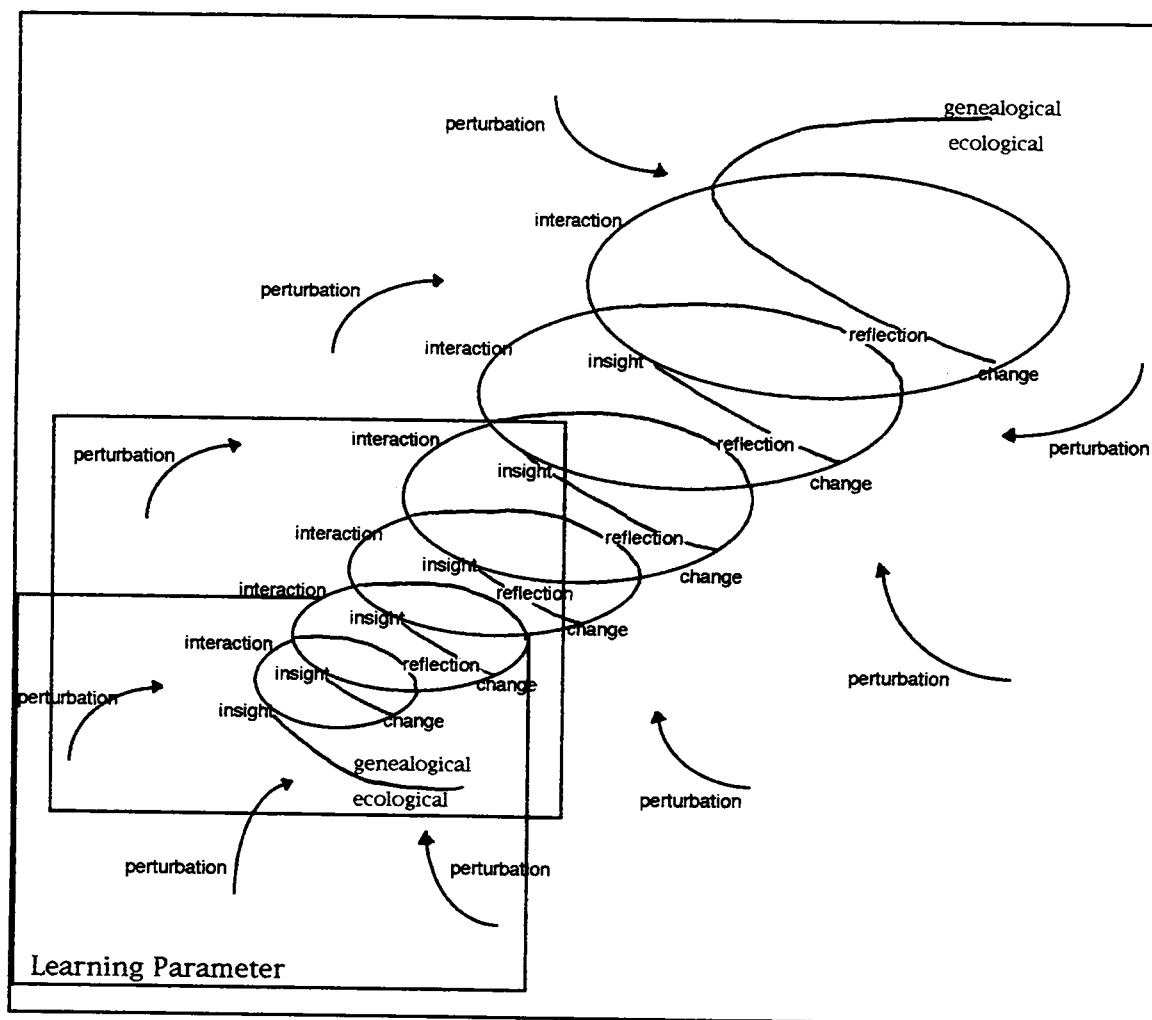


Figure 2: The Evolution of the Teaching-Learning Process

McWhinney (1992) calls "first order change" requiring a high degree of memorization for knowledge acquisition and skill development of concepts and a minimal exploration of concepts by the learner. This separation occurs because the emphasis of assessment in the EIA is increased scores rather than increased learning or learning at a higher level. One does not naturally negate the other unless the measuring process highlights one over the other.

The EIA outlines other performance indicators —"drop out rates, numbers of waivers, local financial contribution to education, local wealth factors related to tax paying ability, attendance rates and others adopted by the State Board of Education" (*Tennessee Public Acts*, 1992, p. 27). While it is necessary for students to be in school and be provided an equal opportunity to perform, these are not learning indicators that will reflect both what students know *and* are able to do—valuable information for teachers to use in adjusting their teaching methods. Performance indicators that reflect knowing as doing and doing as knowing (Maturana & Varela, 1987) are writing samples, projects, experiments, demonstrations, public performances, and/or portfolios—recognized by teachers, parents, and other professionals as achievements of worth, feasibility, and potential impact on the student's learning environment and/or community. None of these performance indicators is explicitly provided for in the EIA.

The assumptions of the policymakers formulating the EIA are similar to those discussed by Astuto et al. (1994) as those that control change in education by forming barriers to restructuring. Evolution of the educational system is a complex, mutually adaptive response to changes in the political and social environment of which the system is a part. Changes within the micro-scales of assessments and/or the macro-scales of policies create

variations in the system's adaptive and creative responses across contexts and time. What is assessed to be valuable will have a bearing on the path of change education will take in Tennessee. The EIA does little to enhance the capacity of teachers or change the system to develop new strategies for adjusted practice at the school level. The EIA reinforced and specified the lines of authority in a top-down configuration. The EIA also reinforced the use of assessments that are summative and decontextualized and that reveal a global pattern of content mastery and skill development.

Value-added assessment is one tool in a reflective professional educator's practice. Any educator with more than one tool is better equipped to construct meaningful learning environments for the richly diverse group of students within any classroom. The policy tools selected and the use of assessment within the design of that selection have the potential to impact what happens to the teaching/learning process. Whether policymakers repeat the history of mandates with an emphasis on high-stakes testing, or take the risk of using different policy tools and expanding the number and type of assessments used in schools by teachers will have a direct impact on the pace of change, the quality of change, and the sustainability of improvement in Tennessee schools for Tennessee students.

Recommendations

Recommendations for future research and policy suggestions indicated by this study are

1. This research should be continued over the full implementation of the funding changes (BEP) of the EIA.

2. In-depth studies of school districts achieving performance goals should be conducted and reported statewide to inform all teachers of innovative practices used to improve student performance.
3. Multiple assessments should be used to evaluate student learning to measure what students know and what students can do with what they know.
4. Professional development should be provided to every teacher in the state in the interpretation and diagnostic use of assessment.
5. As statewide assessment is changed, professional development in the area of assessment should be a budgeted component of the BEP.
6. Policy tool selection for educational improvement should build teacher and administrative capacity to adjust teaching practice for increased learning in relation to changing conditions in student environments. Assessments allowing for capacity-building and systemic change should be formative, contextual, and highly related to the learning targeted for improvement.

REFERENCES

REFERENCES

- Alexander, K. and Alexander, M. D. (1992). *American public school law*, 3rd Edition. St. Paul, MN: West Publishing Company.
- Alexander, K. and Salmon, R. (1995). *Public school finance*. Needham Heights, MA: Allyn & Bacon.
- Astuto, T., et. al. (1994). *Roots of reform: Challenging the assumptions that control change in education*. Bloomington, IN: Phi Delta Kappa Educational Foundation.
- Ballantine, J. (1993). *The sociology of education: A systemic analysis*. Englewood Cliffs, N. J.: Prentice-Hall, Inc.
- Bateson, G. (1972). *Steps to an ecology of mind*. New York: Ballantine Books.
- Bellon, J. and Handler, J. (1982). *Curriculum development and evaluation: A design for improvement*. Dubuque, Iowa: Kendall/Hunt Publishing Company.
- Bellon, J. (1988, Jan.). Career ladder project report. Submitted to the Commissioner of Education, Charles E. Smith.
- Benseler, F., Hejl, P., and Köck, W. (Eds.). (1980). *Autopoiesis, communication, and society: The theory of autopoietic systems in the social sciences*. Frankfurt/New York: Campus Verlag.
- Berliner, D. and Biddle, B. (1995). *The manufactured crisis: Myths, fraud, and the attack on America's public schools*. Reading, MA: Addison-Wesley Publishing Company.
- Berman, M. (1981). *The reenchantment of the world*. Ithaca and London: Cornell University Press.
- Blumer, H. (1969). *Symbolic interactionism: Perspective and method*. Englewood Cliffs, N. J.: Prentice-Hall.
- Bock, R. and Wolfe, R. (1996, Jan. 23). Audit and review of the Tennessee value-added assessment system (TVAAS): Preliminary report. Nashville, TN: Tennessee Office of Education Accountability, Comptroller of the Treasury.

- Bogdan, R. and Biklen, S. (1982). *Qualitative research for education*, Boston: Allyn & Bacon.
- Briggs, J. (1992). *Fractals: The patterns of chaos*. New York: Simon and Schuster.
- Bryk, A. and Raudenbush, S. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, 101 (1), 147-158.
- Caine, R. N. and Caine, G. (1994). *Making connections: Teaching and the human brain*. Menlo Park, California: Innovative Learning Publications.
- Checkland, P. (1995). Systems theory and management thinking. In K. Ellis, A. Gregory, B. Mears-Young, and G. Ragsdell (Eds.), *Critical issues in systems theory and practice*, (pp. 131-135). New York and London: Plenum Press.
- Cohen, D., McLaughlin, M., and Talbert, J. (Eds.). (1993). *Teaching for understanding: Challenges for policy and practice*. San Francisco: Jossey-Bass Publishers.
- Cohen, J. and Stewart, I. (1994). *The collapse of chaos: Discovering simplicity in a complex World*. New York: Viking Penguin.
- Cziko, G. (1989). Unpredictability and indeterminism in human behavior: arguments and implications for educational research. *Educational Researcher*, 18 (3), 17-25.
- Darling-Hammond, L. (1993). Reframing the school reform agenda. *Phi Delta Kappan*, 74 (June), 754-761.
- Darnell, R. (1995-96). *Tennessee blue book: Bicentennial edition (1796-1996)*. Nashville, TN: State Government Publication.
- Davis, D. (1990, Mar. 27). McWherter cabinet to visit 140 school districts. *Nashville Banner*, p. B2.
- De Lange, J. (1995). Assessment: No change without problems. In T. Romberg (Ed.), *Reform in school mathematics and authentic assessment*, (pp. 87-172). Albany: State University of New York Press.

- Deal, D. and Kennedy, A. (1982). *Corporate cultures : The rites and rituals of corporate life*. Reading, MA: Addison-Wesley Publishing Company.
- Dellar, G. (1992). Connections between macro and micro implementation of educational policy: A study of school restructuring in Western Australia. Paper presented at the AERA conference, San Francisco. (ED 347650).
- Denzin, N. (1978). *The research act: A theoretical introduction to sociological methods*. New York: McGraw-Hill.
- DePryck, K. (1993). *Knowledge, evolution, and paradox: The ontology of language*. Albany: State University of New York Press.
- Dey, I. (1993). *Qualitative data analysis: A user-friendly guide for social scientists*. London and New York: Routledge.
- Eidenberg, E. and R. Morey. (1969). *An act of congress: The legislative process and the making of education policy*. New York: W. W. Norton & Company, Inc.
- Elmore, R. and McLaughlin, M. (1982). Strategic choice in federal education policy: The compliance-assistance trade-off. In A. Lieberman and M. McLaughlin (Eds.), *Policy Making in Education*, (pp. 159-194). Chicago: The University of Chicago Press.
- Fisher, T. (1996, January). A review and analysis of the Tennessee value-added assessment system. Nashville, TN: Tennessee Office of Education Accountability, Comptroller of the Treasury.
- Fuhrman, S. (Ed.). (1993). *Designing coherent education policy: Improving the system*. San Francisco: Jossey-Bass Publishers.
- Gardner, H. (1991). *The unschooled mind*. New York: Basic Books, a division of HarperCollins Publishers.
- Gifford, B. and O'Connor, M. (Eds.). (1992). *Changing assessments: Alternative views of aptitude, achievement and instruction*. Boston: Kluwer Academic Publishers.
- Glaser, B. and Strauss, A. (1967). *The discovery of grounded theory*. Chicago: Aldine.

- Glassner, B. and Moreno, J. (Eds.). (1989). *The qualitative-quantitative distinction in the social sciences*. Dordrecht: Kluwer Academic Publishers.
- Gleick, J. (1987). *Chaos: The making of a new science*. New York: Viking Penguin.
- Glesne, C. and Peshkin, A. (1992). *Becoming qualitative researchers: An introduction*. White Plains, NY: Longman Publishing Group.
- Goldstein, J. (1994). *The unshackled organization: Facing the challenge of unpredictability through spontaneous reorganization*. Portland, OR: Productivity Press, Inc.
- Goodwin, B. (1994). *How the leopard changed its spots: The evolution of complexity*. New York: Charles Scribner's Sons.
- Green, H., C. Smith, and S. Hydorn. (1995, June). *Funding Tennessee schools: From reform to restructuring*. Tennessee Advisory Commission on Intergovernmental Relations Report.
- Goodwin, B. (1994). *How the leopard changed its spots: The evolution of complexity*. New York: Charles Scribner's Sons.
- Griffin, G. (1991). Interactive staff development: Using what we know. In A. Lieberman and L. Miller (Eds.), *Staff development for education in the '90s: New demands, new realities, new perspectives*,
- Hall, M. L. W. (1995). Merging the world views of systems science and human values: The organizational perspective. In K. Ellis, A. Gregory, B. Mears-Young, and G. Ragsdell (Eds.), *Critical issues in systems theory and practice*, (pp. 131-135). New York and London: Plenum Press.
- Hanson, G. (1994). Critical issues in the assessment of value added in education. In J. Stark and A. Thomas (Eds.), *Assessment and program evaluation*, (pp. 585-595). Needham Heights, MA: Simon and Schuster Custom Publishing.
- Hayes, P. et al. (1992). *Winners all: A call for inclusive schools*. Alexandria, VA: The National Association of State Boards of Education. The report of the NASBE study group on special education.

- Hodder, I. (1994). The interpretation of documents and material culture. In N. Denzin and Y. Lincoln (Eds.), *Handbook of qualitative research.*, (pp. 393-402). Thousand Oaks, CA: Sage Publications, Inc.
- Holt, A. (1938). *The struggle for a state system of public schools in Tennessee, 1903-1936*. New York: Teachers College, Columbia University.
- Horn, J. (1995). *The emergence of meaning in an interdisciplinary learning environment: A qualitative study of the James R. Stokely Institute for Liberal Arts Education*. Unpublished doctoral dissertation. University of Tennessee, Knoxville.
- Horn, J. and Wilburn, D. (1994). *Complexity: Implications for schooling*. Unpublished paper delivered at the Chaos Society Conference, Johns Hopkins University, Baltimore, MD.
- Humphrey, T. (1994, Jan. 21). McWherter, Alexander found lobbyists powerful opponents. Knoxville News-Sentinel, p. A1.
- Hunter, M. (1982). *Mastery Teaching*. El Segundo, CA: TIP Publications.
- Jaeger, R. M. (1993). Certification of student competence. In R. Linn (Ed.), *Educational Measurement*, 3rd Edition, (pp. 485-514). Phoenix, AZ: The Oryx Press.
- Jaynes, J. (1976). *The origin of consciousness in the breakdown of the bicameral mind*. Boston: Houghton Mifflin Company.
- Joint Committee on Standards for Educational Evaluation. (1981, 1994). *Standards for evaluations of educational programs, projects, materials*. New York: McGraw-Hill.
- Kauffman, S. (1995). *At home in the universe: The search for laws of self-organization and complexity*. New York and Oxford: Oxford University Press.
- Keese, N. (1990). *Educational decision making in the Tennessee state legislature*. Unpublished doctoral dissertation. The University of Tennessee, Knoxville.
- Klausnitzner, D. (1992, March 22). 'Learning community' concept expected to be school winner. *Nashville Tennessean*, p. B8.
- Kliebard, H. (1992). *Forging the American curriculum*. New York: Routledge.

- Köck, W. (1980). Autopoiesis and communication. In F. Benseler, P. Hejl, and W. Köck (Eds.), *Autopoiesis, communication, and society: The theory of autopoietic systems in the social sciences*, (pp. 87–112). Frankfurt/New York: Campus Verlag.
- Kronenberg, P. (1995). Chaos and re-thinking the public policy process. In A. Albert (Ed.), *Chaos and society*, (pp. 15–33). Amsterdam: IOS Press.
- Kugler, P. and R Shaw. (1990). Symmetry and symmetry breaking in thermodynamic and epistemic engines: A coupling of the first and second laws. In H. Haken and M. Stadler (Eds.), *Synergetics of cognition: Springer series in synergetics*, vol. 45, (pp. 296–331). Berlin: Springer-Verlag.
- Lajoie, S. (1995). A framework for authentic assessment in mathematics. In T. Romberg (Ed.), *Reform in school mathematics and authentic assessment*, (pp. 19–37). Albany: State University of New York Press.
- Legislative minutes, Joint Education Oversight Committee, December 17, 1987. Nashville, TN: Tennessee State Library and Archives. (Commissioner Charles Smith, Dr. Brent Poulton, Senator Leonard Dunavant, Senator Robert Rochelle, Representative Leslie Winningham).
- Legislative minutes, House Education Oversight Committee, July 12, 1989. Nashville, TN: Tennessee State Library and Archives. (Dr. Brent Poulton, Representative Paul Starnes, Representative Shelby Rhinehart).
- Legislative minutes, House Education Committee, January 23, 1990. Nashville, TN: Tennessee State Library and Archives. (Dr. William Sanders, Representative Tommy Burks, Senator Leonard Dunavant, Senator Douglas Henry, Representative John Bragg).
- Legislative minutes, Senate Education Committee, January 31, 1990. Nashville, TN: Tennessee State Library and Archives. (Dr. Brent Poulton).
- Legislative minutes, Joint Education Oversight Committee, October 30, 1990. Nashville, TN: Tennessee State Library and Archives. (Representative David Copeland).

- Legislative minutes, Senate Education Committee, February 7, 1990.
Nashville, TN: Tennessee State Library and Archives. (Senator Milton Hamilton, SBE Chairman Nelson Andrews).
- Legislative minutes, House Education Committee, February 12, 1991.
Nashville, TN: Tennessee State Library and Archives. (Commissioner Charles E. Smith).
- Legislative minutes, House Education Committee, February 19, 1991.
Nashville, TN: Tennessee State Library and Archives.
(Superintendent Mike Simmons, Representative John Bragg).
- Legislative minutes, Senate Education Committee, March 6, 1991. Nashville, TN. Tennessee State Library and Archives. (Senator Ben Atchley, John Parish, Dan Frierson, Senator Ray Albright, Dewey Branstetter, Senator Douglas Henry, Senator Pete Springer).
- Legislative minutes, House Education Committee, April 8, 1991. Nashville, TN: Tennessee State Library and Archives. (Tennesseans for Limited Taxation Association, Cavit Cheshier, Representative Phillip Pinion, Commissioner Charles E. Smith, Carol Stinson, Chairman Gene Davidson).
- Legislative minutes, House Education Committee, April 9, 1991. Nashville, TN: Tennessee State Library and Archives. (Commissioner Charles E. Smith).
- Legislative minutes, Senate Education Committee, April 17, 1991. Nashville, TN: Tennessee State Library and Archives. (Senator Ray Albright).
- Lewin, R. (1992). *Complexity: Life at the edge of chaos*. New York: Macmillan Publishing Company.
- Linn, R. (1993). Current perspectives and future directions. In R. Linn (Ed.), *Educational measurement*, 3rd Edition, (pp. 1-11). Phoenix, AZ: The Oryx Press.
- Lofland, J. (1976). *Doing social life: The qualitative study of human interaction in natural settings*. New York: Wiley.
- Lowham, J. (1995). Evolution of intentions: From state policy development to teacher implementation. In D. Carter and M. O'Neill (Eds.), *Case studies in educational change: An international perspective*. (pp. 95-114). London and Washington, D.C.: The Falmer Press.

- Madaus, G. (1988). The influence of testing on the curriculum. In L. N. Tanner (Ed.), *Critical issues in curriculum, Eighty-seventh yearbook of the National Society for the Study of Education*, pp. 83-121. Chicago, IL: The University of Chicago Press.
- Madaus, G., Stufflebeam, D. and Scriven, M.. (1994). Program evaluation: A historical overview. In J. Stark and A. Thomas (Eds.), *Assessment and program evaluation*, (pp. 585-595). Needham Heights, MA: Simon and Schuster Custom Publishing.
- Mainzer, K. (1994). *Thinking in complexity: The complex dynamics of matter, mind, and mankind*. Berlin: Springer-Verlag.
- Marsh, D. and Sevilla, J. (1992). Financing middle school reform: Linking costs and education goals. In A. Odden (Ed.), *Rethinking school finance: An agenda for the 1990s*, (pp. 97-127). San Francisco: Jossey-Bass Publishers.
- Marshall, C. and Rossman, G. (1989). *Designing qualitative research*. Newbury Park, CA: Sage Publications.
- Master Plan For Tennessee Schools: Preparing for the 21st Century. (1990). State Board of Education Document.
- Master Plan For Tennessee Schools: Preparing for the 21st Century. (1995). State Board of Education Document.
- Maturana, H. and F. Varela. (1980). *Autopoiesis and cognition: the realization of the living*. Dordrecht: Reidel.
- (1987). *The tree of knowledge: The biological roots of human understanding*. Boston: New Science Library.
- Mayshark, J. (May 5, 1996). Teacher value scores not to be published. *The Knoxville News-Sentinel*. pp. B1, B7.
- McDonnell, L. (1994). *Policymakers' views of student assessment*. Santa Monica, CA: The RAND Corporation.
- McDonnell, L. and Elmore, R. (1987). *Alternative policy instruments*. Santa Monica, CA: The RAND Corporation.

- McLaughlin, M. (1975). *Evaluation and reform: The elementary and secondary education act of 1965, title I*. Cambridge, MA: Ballinger Publishing Company.
- McWherter, N. (1989, Sept. 25). Press conference release. Collection of Governor McWherter's Papers. Nashville, TN: Tennessee State Library and Archives.
- McWhinney, W. (1992). *Paths of change*. Newbury Park: Sage Publications.
- Michaels, M. (1995). Seven fundamentals of complexity for social science research. In A. Albert (Ed.), *Chaos and society*, (pp. 15-33). Amsterdam: IOS Press.
- Mingers, J. (1995). *Self-producing systems: Implications and applications of autopoiesis*. New York and London: Plenum Press.
- Murphy, J. and J. Schiller. (1992). *Transforming America's schools: An administrators' call to action*. La Salle, IL: Open Court Publishing Company.
- Nashville Banner*. (1987, Jan. 24). Reform not linked to career ladder: TEA survey, p. A7.
- Nashville Banner*. (1987, Jan. 29). Career ladder a travesty to all, p. A14.
- Nashville Banner*. (1987, Mar. 23). Jury out on career ladder, p. A1.
- Nashville Banner*. (1987, Mar. 24). Tennessee teachers under microscope, p. A1.
- Nashville Banner*. (1987, April 24). Youth issues panel sees sunbelt clouds, p. C1.
- Nashville Banner*. (1988, Nov. 18). Full new funding for state schools may take time, p. B1.
- Nashville Banner*. (1989, Mar. 8). TEA aims again at career ladder, p. A8.
- Nashville Banner*. (1989, Sept. 25). Assembly to consider funding \$400 million education program, p. B1.
- Nashville Banner*. (1989, Nov. 3). TEA plan similar to governor's goals, p. A6.

- Nashville Banner*. (1990, Mar. 27). McWherter cabinet to visit 140 school districts, p. B2.
- Nashville Banner*. (1990, Sept. 20). School reformers basing state plan on accountability, p. A1.
- Nashville Tennessean*. (1988, July 13). An unfair school system and an unfair tax system, p. A8.
- Nashville Tennessean*. (1988, Nov. 15). Tighter budget control to avoid more tax hinted, pp. B1, B2.
- Nashville Tennessean*. (1988, Nov. 27). No guarantee for poor areas, p. A1.
- Nashville Tennessean*. (1989, Oct. 29). Standardized test ban asked by AFT chief, p. A1.
- Nashville Tennessean*. (1989, Nov. 3). TEA to tell teachers' view of school needs, p. A1.
- Nashville Tennessean*. (1990, Jan. 27). 21st century schools education aim of McWherter, p. A1.
- Nashville Tennessean*. (1990, Nov. 18). Lawmakers may stay for school reform, pp. B1.
- O'day, J. and Smith, M. (1993). Systemic reform and educational opportunity. In S. Fuhrman (Ed.), *Designing coherent education policy: Improving the system*, (pp. 250-312). San Francisco: Jossey-Bass Publishers.
- Odden, A. (Ed.). (1991). *Education policy implementation*. Albany, NY: State University of New York Press.
- Odden, A. (1992). School finance and education reform: An overview. In A. Odden (Ed.), *Rethinking school finance: An agenda for the 1990s*, (pp. 1-40). San Francisco: Jossey-Bass Publishers.
- Odden, A. and L. Kim. (1992). Reducing disparities across the states: A new federal role in school finance. In A. Odden (Ed.), *Rethinking school finance: An agenda for the 1990s*, (pp. 260-297). San Francisco: Jossey-Bass Publishers.

- Office of Planning, Budget, and Evaluation. (1989). *Sixth annual state education performance chart*. Washington, D. C.: U. S. Department of Education.
- Patton, M. Q. (1987). *How to use qualitative methods in evaluation*. Newbury Park, CA: Sage Publications.
- Patton, M. Q. (1990). *Qualitative evaluation and research method*. Newbury Park, CA: Sage Publications.
- Payton, R. (1991, Feb. 12). Prepared testimony for House Education Committee. Collection of Governor McWherter's Papers. Nashville, TN: Tennessee State Library and Archives.
- Priesmeyer, H.R. (1992). *Organizations and chaos: Defining the methods of nonlinear management*. Westport, Connecticut: Quorum Books.
- Reinharz, S. (1992). *Feminist methods in social research*. New York: Oxford University Press, Inc.
- Rifkin, J. (1989). *Entropy: Into the greenhouse world*. New York: Bantam Books.
- Rist, R. (1994). Influencing the policy process with qualitative research. In N. Denzin and Y. Lincoln (Eds.), *Handbook of qualitative research*, (pp. 545-557). Thousand Oaks, CA: Sage Publications, Inc.
- Romberg, T. and Wilson, L. (1995). Issues related to the development of an authentic assessment system for school mathematics. In T. Romberg, (Ed.), *Reform in school mathematics and authentic assessment*, (pp. 1-18). Albany: State University of New York Press.
- Rules, Regulations, and Minimum Standards for the Governance of Public Schools in the State of Tennessee*. (May 1988 and July 1992 revised). Book 4, Goals of Education and Curriculum Frameworks. Nashville, TN: State Board of Education.
- Rules, Regulations, and Minimum Standards for the Governance of Public Schools in the State of Tennessee*. (1994). Nashville, TN: State Board of Education.
- Sanders, W. and S. Horn. (1994). The Tennessee value-added assessment system (TVAAS): mixed model methodology in educational assessment. *Journal of Personnel Evaluation in Education*, 8 (1), 299-311.

- Sanders, W. and S. Horn. *Educational assessment reassessed: the usefulness of standardized and alternative measures of student achievement as indicators for the assessment of educational outcomes*. Unpublished manuscript, 1994.
- Sanders, W. (1995). Graphical summary of educational findings from the Tennessee value-added assessment system (TVAAS). Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center.
- Sanders, W., and Horn, S. (1995). An overview of the: Tennessee value-added assessment system (TVAAS). A pamphlet printed by authors.
- Schmoker, M. (1996). *Results: The key to continuous school improvement*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Senge, P. (1990). *The fifth discipline: The art & practice of the learning organization*. New York: Doubleday.
- Shulman, L. S. (1987). Knowledge in teaching; Foundations of the new reform." *Harvard Educational Review*, 5.
- Smith, C. (1987, Oct. 15). Text of Prepared Remarks to the Legislative Oversight Committee. Collection of the Governor McWherter's Papers. Nashville, TN: Tennessee State Library and Archives.
- Smith, C. (1988, Jan. 26). Tennessee state department of education news release. Collection of Governor McWherter's Papers. Nashville, TN: Tennessee State Library and Archives.
- Smith, C. (1988, Jan. 27). Tennessee state department of education news release. Collection of Governor McWherter's Papers. Nashville, TN: Tennessee State Library and Archives.
- Smith, C. (1988, June 7). Public's education awareness renewed by state reforms. *Nashville Banner*, p. A9.
- Smith, C. (1989). The qualitative significance of quantitative representation. In B. Glassner and J. Moreno (Eds.), *The qualitative-quantitative distinction in the social sciences*, (pp. 29-42). Dordrecht, The Netherlands: Kluwer Academic Publishers, Inc.

- Smith, C. (1990). Goals and objectives of the 21st century challenge plan. Nashville TN: State Department of Tennessee. (Submitted to and approved by the Legislative Oversight Committee on Education February, 1990. Submitted and approved by the State Board of Education March, 1990.)
- Smith, W. (1952). *The history of the Tennessee education association*. Unpublished doctoral dissertation. George Peabody College.
- Snodgrass, W. (1990, February). Performance Audit. Nashville, TN: Comptroller of the Treasury, Office of Education Accountability Report.
- Snodgrass, W., Baker, A. and Detch, E. (1994, March). The role of the state board of education. Nashville, TN: Comptroller of the Treasury, Office of Education Accountability Report.
- Snodgrass, W., Baker, A., Xu, D., and Detch, E. (1995). The measure of education: A review of the Tennessee value added assessment system. Nashville, TN: Comptroller of the Treasury, Office of Education Accountability Report.
- Stacey, R. (1992). *Managing the unknowable: Strategic boundaries between order and chaos in organizations*. San Francisco: Jossey-Bass.
- Stake, R. (1978). The case study method in social inquiry. *Educational Researcher*, 7 (2), 5-8.
- Sterling, A. (1992). *Human creativity as a function of chaotic dynamics: Implications for research and practice in education*. Unpublished doctoral dissertation. University of Oregon.
- Strauss, A. and Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage Publications.
- Student, teacher and school performance*. (1995). Tenth Annual Report submitted to the Governor and the General Assembly of the State of Tennessee. Nashville, TN: Tennessee Department of Education.
- Student, teacher and school performance*. (1996). Eleventh Annual Report submitted to the Governor and the General Assembly of the State of Tennessee. Nashville, TN: Tennessee Department of Education.

- Taylor, R. (1994). Qualitative Research. In M. Singletary (Ed.), *Mass communication research: Contemporary methods and applications*, (pp. 265-279). New York: Longman Publishing Group.
- TEA News. (1986, July). T-PACE endorsement of McWherter, p. 4.
- TEA News. (1988, April). Bellon report, p. 4-5.
- Tennessee Public Acts. (1984). The Comprehensive Education Reform Act, pp. 20-65
- Tennessee Public Acts. (1992). The Educational Improvement Act, pp. 19-49.
- Thomas, D. (1990, November). Tennessee education for the 21st century. Report to Governor Ned McWherter. Nashville, TN: State Board of Education.
- Thro, W. (1993). *The significance of the Tennessee school finance decision*. St. Paul, MN: West Publishing Company.
- Tuchman, G. (1994). Historical social science: Methodologies, methods, and meanings. In N. Denzin and Y. Lincoln (Eds.), *Handbook of qualitative research.*, (pp. 306-323). Thousand Oaks, CA: Sage Publications, Inc.
- U. S. Statutes at Large. (1965). Vol. Elementary and Secondary Education Act, Washington, D.C. U.S. Government Printing Office. pp. 27-58.
- U. S. Statutes at Large. (1975). Vol. 89. Education for all Handicapped Children Act, Washington, D.C., U.S. Government Printing Office, pp. 773-796.
- Varela, F. (1979). *Principles of biological autonomy*. New York: Elsevier North Holland, Inc.
- Varela, F. (1992). Whence perceptual meaning? A cartography of current ideas. In F. Varela and J. Dupuy (Eds.), *Understanding origins: Contemporary views on the origin of life, mind, and society*, pp. 235-263. The Netherlands: Kluwer Academic Publishers.
- Vickers, B. (Ed.). (1994-1995). *Tennessee Statistical Abstract*. Center for Business and Economic Research, College of Business Administration: The University of Tennessee, Knoxville.
- Villa, R., Thousand, J., Stainback, W., Stainback, S. (Eds.). (1992). *Restructuring for caring and effective education: An administrative*

guide to creating heterogeneous schools. Baltimore: Paul H. Brookes Publishing Co.

- von Glasersfeld, E. (1980). The concept of equilibration in a constructivist theory of knowledge. In F. Benseler, P. Hejl, and W. Köck (Eds.), *Autopoiesis, communication, and society: The theory of autopoietic systems in the social sciences*, (pp. 75–85) Frankfurt and New York: Campus Verlag.
- Warf, J. H. (no date). *Tennessee state department of education, 1900 to 1968*. Tennessee Department of Education.
- Waldrop, M. (1992). *Complexity: The emerging science at the edge of order and chaos*. New York: Simon and Schuster.
- Wheatley, M. (1994). *Leadership and the new science: Learning about organization from an orderly universe*. San Francisco: Berrett-Koehler.
- Weeks, K. (1988, March 25). A study of the Tennessee Foundation Program. Nashville, TN: Tennessee State Board of Education.
- Wiggins, G. (1993). *Assessing student performance: Exploring the purpose and limits of testing*. San Francisco: Jossey-Bass Publishers
- Williams, P., Floyd, C., and Wissner, S. (1990, Sept. 24). Series education reform: A Tennessee state of mind. *Nashville Tennessean*, p. A1.
- Wilson, K. (1990, Feb. 28). Schools need monitoring, funds: adviser. *Nashville Banner*, p. B6.
- Wissner, S. (1990, Sept. 29). Alexander's better schools plan graduates to remedial. *Nashville Tennessean*, pp. A1, A8.
- Wolcott, H. (1990). On seeking—and rejecting—validity in qualitative research. In E. Eisner and A. Peshkin (Eds.), *Qualitative inquiry in education: The continuing debate*, (pp. 121–152). New York: Teachers College Press.
- Worthen, B. (1995). Some observations about the institutionalization of evaluation. *Evaluation Practice*, 16(1), 29-36.
- Wright, C. (1987, Jan. 26). Merit pay program change sought, TEA poll shows. *Nashville Banner*, p. C2.

Yin, R. (1994). *Case study research: Design and methods*. Thousand Oaks, CA: SAGE Publications.

VITA

Susan Denise Wilburn was born in Knoxville, Tennessee on June 23, 1955. She attended the public schools in Kingsport, Tennessee and graduated from Dobyns-Bennett High School in June 1973. She received a Bachelor of Science degree in speech pathology from East Tennessee State University in 1977. She was employed as a speech pathologist and coordinator of special education programs in the Greeneville (Tennessee) City Schools from 1977 to 1991. She earned a Master of Arts in Teaching from East Tennessee State University in 1991. She has worked in the Anderson County (Tennessee) School System as a gifted/resource teacher and coordinator of special education programs. She entered the doctoral program at the University of Tennessee in 1993 in Curriculum and Instruction.

She is currently employed by the Anderson County School System as Director of Staff Development.