The use of performance indicator systems in public higher education

John Milan Davis
William & Mary - School of Education

Recommended Citation
https://scholarworks.wm.edu/etd/1550154048

Follow this and additional works at: https://scholarworks.wm.edu/etd
Part of the Higher Education Commons

This Dissertation is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Dissertations, Theses, and Masters Projects by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.
THE USE OF PERFORMANCE INDICATOR SYSTEMS
IN PUBLIC HIGHER EDUCATION

A Dissertation
Presented to
The Faculty of the School of Education
The College of William and Mary in Virginia

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

by
John Milan Davis

April 2003
# Table of Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter I</td>
<td>Introduction</td>
<td>2</td>
</tr>
<tr>
<td>Chapter II</td>
<td>Review of the Literature</td>
<td>11</td>
</tr>
<tr>
<td>Chapter III</td>
<td>Methodology</td>
<td>51</td>
</tr>
<tr>
<td>Chapter IV</td>
<td>Findings</td>
<td>60</td>
</tr>
<tr>
<td>Chapter V</td>
<td>Summary, Discussion, and Conclusions</td>
<td>100</td>
</tr>
<tr>
<td>Appendix A</td>
<td>Sample Cover Letter and Survey</td>
<td>114</td>
</tr>
<tr>
<td>Bibliography</td>
<td>116</td>
<td></td>
</tr>
</tbody>
</table>
List of Tables

Table 1  Summary of PINS Research Findings 38
Table 2  Data Principles of MGT Associates 43
Table 3  Effects of PINS According to SHEFO 47
Table 4  Summary of PINS Evidence from the World Wide Web 63
Table 5  Status of Reporting Performance 71
Table 6  Primary Use of Performance Data 73
Table 7  Level of Buy-In Among States 78
Table 8  Institutional Uniqueness in Performance Indicators 81
Table 9  Understandability of Performance Indicators 84
Table 10  Nature of Indicators 86
Table 11  Use of Performance Indicators for Benchmarking 88
List of Figures

Figure 1 17
Figure 2 22
Figure 3 51
Figure 4 54
THE USE OF PERFORMANCE INDICATOR SYSTEMS IN PUBLIC HIGHER EDUCATION

ABSTRACT

The use of Performance Indicator Systems (PINS) grew significantly during the 1990s. State Higher Education Finance Officers (SHEFO) were surveyed to determine the degree to which PINS affected state appropriations and how well measures used within PINS conformed to standards set forth in the literature.

Findings suggested that majority of states are engaged in various forms of PINS. However, these forms do not readily conform to the findings within extant literature. PINS have very little impact on state appropriations and that Wildavsky's (1984) theory of budget incrementalism might serve as a better explanation for changes in budget appropriations. Additionally, measures used within PINS varied considerably in terms of quality, utility, and comparability.

JOHN MILAN DAVIS

SCHOOL OF EDUCATION

THE COLLEGE OF WILLIAM AND MARY IN VIRGINIA

vi
This dissertation is dedicated to Heidi for her love, devotion, patience, and support. I also dedicate this work to Grant, who provides me with endless inspiration and joy. Finally, I dedicate this work to the fellows of the Tuesday Night Group, who have helped me in ways few would understand.
THE USE OF PERFORMANCE INDICATOR SYSTEMS
IN PUBLIC HIGHER EDUCATION
Public perception toward higher education has changed dramatically over the past fifteen years. For public colleges and universities much of this change has been made manifest through the imposition of policies intended to improve institutional or systemic accountability. In addition to the general public, state legislators, governing boards, coordinating boards, and governors have increasingly cast a more critical eye toward higher education (Berdahl and McConnell, 1999). This is particularly true with respect to the issue of institutional effectiveness of colleges and universities. A very palpable public perception emerged (supported by evidence both quantifiable and anecdotal) suggesting that the quality of education within America’s colleges and universities was diminishing. Such perceptions grew out of a series of formal and informal, widely-read, published reports including those by the National Endowment of the Humanities in To Reclaim a Legacy: A Report on the Humanities in Higher Education (1984); the Association of American Colleges’ Integrity In the College Curriculum: A
Report to the Academic Community (1985); and the National Governor's Association's *Time for Results: The Governor's 1991 Report on Education* (1991). These reports suggested something needed to be done to prevent the erosion of quality within higher education. As a result, the push for assessment and accountability gained significant momentum. Additionally, colleges and universities were placed in increasingly competitive environments with the growth of proprietary education.

As early as 1984 a movement was afoot (particularly within the state of Tennessee) to develop policies designed to assess the effectiveness of institutions. By the late 1980s most states were actively involved in the development of policies intended to assess the effectiveness of their public institutions. Such activity was revolutionary for higher education. Never before had higher education been subjected to such external scrutiny. This phenomenon and its attendant effects would continue to grow into the twenty-first century.

The ways in which accountability is ensured has taken on a much more formal posture over the last decade. In the early days (mid to late 1980s) responsibility for assessment was largely voluntary and conducted at the institutional level (Banta, Lund, Black, and Oblander, 1996). Many
institutions developed methods to assess various aspects of the undergraduate experience including classroom knowledge, communication skills, specific vocational skills, and co-curricular experiences. However, a number of issues complicated this effort. First, institutions relied largely upon internally derived methods of assessment (Ewell, 1993). Such data were often so contextually bound and thick that they were difficult to understand and interpret to external audiences. Second, each institution (even within the same system) was developing its own methods and definitions of institutional effectiveness that data from similar type of institutions could not be compared. In an effort to improve institutional effectiveness policy makers began to develop ways in which institutions could be judged. However such efforts failed to mollify those most critical of higher education (Gaither and Neal, 1993). As a result, policymakers created more formal systems to ensure accountability. Such systems often coupled performance on measures related to effectiveness with resource allocation.

Importance of the Study

This study is important for a number of reasons. First, the literature suggests that determining performance and accountability is a very active policy initiative on the part of states (Burke, 1997; 2001). Considerable public
resources are being expended to collect, analyze, interpret, and report data that comprise performance indicators. State policies that require institutions to report on measures of institutional effectiveness require a significant amount of time by faculty, staff, and administrators. Often the types of data that are increasingly being required are not readily available. A fundamental question begins to emerge as to whether such efforts are worth the cost.

Second, while much is known regarding who is requiring data and what types of data are being required there is virtually no knowledge with regard to the quality of data generated. Furthermore, there is no extant knowledge suggesting these systems do or do not make a difference where they are used. This is especially true in states that have adopted such systems as a tool for performance funding.

Finally, there is a movement within the higher education accreditation industry toward requiring institutions to have systems in place to demonstrate institutional effectiveness. The Southern Association of Colleges and Schools (SACS) newly adopted comprehensive standards require that institutions identify outcomes and demonstrate program improvement. Failure to comply with these standards can result in an institution’s accreditation being compromised.
The Purpose of this Study

Although previous research has been useful in enhancing the understanding of Performance Indicators Systems (PINS) there are significant knowledge gaps in this area. Virtually nothing is known regarding the impact of such systems (Burke, 2001). Furthermore, there have been no studies to examine the issue of quality. That is, are the specific indicators being gathered and reported actually reflective of the constructs and values being investigated at the level of the campus?

This study was both exploratory and explanatory in nature. Its aim was to enhance the understanding of current policy and implementation as well as understanding how key players involved in this policy view the quality, comparability, and utility of the information being used. The purpose of this study was to explore the effectiveness of the PINS in terms of the indicators' quality, utility and comparability from the perspective of State Higher Education Financial Officers (SHEFO). This was accomplished via a multi-method approach. Additionally, this study attempted to determine whether performance within PINS was, in fact, linked to appropriations. In other words, do states link performance on established measures to rewards or sanctions as a matter of policy or practice? Finally, results from
this study were viewed in light of Wildavsky's (1984) theory of budget incrementalism. The study began with the process of verifying the findings of the Rockefeller Institute of Government surveys (2001). This was accomplished by reviewing each state's higher education web site for evidence of PINS. Next, a survey (see Appendix A) was sent to all SHEFO to gain an understanding of how they perceive the quality, comparability, and utility of the data that comprise their respective PINS. Additionally, the questions probed for the characteristics of indicators advanced by MGT of America Inc. (2001), henceforth referred to in this paper as MGT Associates. Finally, follow-up telephone interviews were conducted with individuals (each of whom dealt with these PINS) from a different perspective. The content of the interviews were then analyzed in order to develop a clearer picture of how (or if) PINS were being used.

Limitations and Delimitations

A primary limitation of this study was that it focused primarily on the perspective of the SHEFO. There are other players involved including those at the level of the institution as well as those who craft policy—legislators, others within the executive departments, etc. In an attempt to mitigate this weakness follow up interviews were conducted with various players within one state. An
additional potential weakness was the fact that some SHEFO may not be very familiar with the particular PINS system in place. SHEFO that have been hired or assumed their position following the development of PINS were not as familiar with the intricacies of their particular system. Another significant limitation of this study was the fact that it focused on four-year institutions to the exclusion of community colleges. In a number of states PINS have been developed exclusively for community colleges to the exclusion of four-year institutions. To assist in making the analysis more consistent, focus was placed on one stratum of higher education, public four-year colleges and universities.

Definitions

Performance Indicator Systems is a term designated to represent formal efforts on the part of state policymakers to systematically collect, analyze, and report data that convey the overall effectiveness of institutions of higher education. Such systems may take different forms and serve simply as a reporting function to state government while others provide performance data that can affect the amount of public money allocated to a particular institution (Burke, 2001). To borrow the terms used by the Rockefeller Institute, in the case of Performance Funding (PF), performance is
tightly coupled with budget allocations. In other words, how well an institution performs along a specified set of criteria will determine the amount of public funding available to that institution. The more common use of PINS is found through the use of Performance Budgeting (PB). Within this scheme, performance is more loosely coupled with resource allocation. In other words, performance may be taken into account (along with other factors) as a means in determining institutional budget appropriations.

To serve as a benchmark for judging the overall usefulness of data this study employed the concepts developed formally by the National Postsecondary Education Cooperative (NPEC). The focus of the NPEC has been to increase the overall effectiveness of educational data by emphasizing the concepts of quality, utility, and comparability. Quality speaks to the degree to which data actually represent the construct under examination. In essence, the issue is one of validity. For example, data such as graduation rates purportedly speak to the degree to which institutions are efficient. But are such data a valid measure of efficiency? Utility refers to the notion that the data that are generated are used in some meaningful manner for policy change. The question is, are the data that are being generated really used in a manner that is productive?
Or, are data being collected for the sake of data collection. Finally, the issue of comparability refers to the idea that data within the system should be able to be judged similarly. This is largely a function of scale and measurement. In other words, are institutions using data that allow for meaningful comparison?

While no study can completely provide for a thorough understanding of the impact that PINS have, this study significantly advanced the knowledge base that currently exists. A considerable amount of research has been conducted detailing the types of data being collected at the state level through the implementation of PINS (Burke and Minassians, 2001; Burke, Rosen, and Minassians, 2000; Burke and Modarresi, 1999; Christal, 1998). It is important now to advance the knowledge further by more closely examining the types of data collected and the impact that such data collection is having upon America's public colleges and universities. SHEFO are in a unique position to address these questions because they are primarily responsible for the collection of performance data at the state level. They serve as "point persons" between the individual institutions and the policymaking bodies within state government. As such, their perspective is invaluable and remains insufficiently tapped within the literature.
CHAPTER II
REVIEW OF THE LITERATURE

The use of performance funding and performance budgeting strategies begs a number of real questions. How prevalent are performance funding and performance budgeting systems within the states? What types of indicators are being used as data generators for performance funding and budgeting systems?

The areas surveyed by this review include the emergence of the accountability movement within American public higher education, the efforts on the part of policymakers to initiate a variety of strategies aimed towards insuring such accountability, and the current state of knowledge regarding performance funding and performance budgeting policies.

The Emergence of the Accountability Movement
During the late 1970s and through the 1980s higher education received an increasing amount of external scrutiny concerning the quality of undergraduate education. A number of national reports and studies published during the 1980s were critical of what was taking place within America’s
colleges and universities (Nettles, Coles, and Sharp, 1997). These included works by the National Endowment of the Humanities in *To Reclaim a Legacy: A Report on the Humanities in Higher Education* (1984); the Association of American Colleges' *Integrity In the College Curriculum: A Report to the Academic Community* (1985); and the National Governor's Association's *Time for Results: The Governor’s 1991 Report on Education* (1991). At the time, the economies of Asia and Europe began to gain leverage against the economic dominance of the United States. A general feeling existed among the American public that undergraduate education was not developing the skills needed to sustain a competitive workforce and the American economic position worldwide. Such feelings also penetrated the K-12 sector of public education. As a way to combat this perceived decline, Ewell (1985) suggested that state policymakers develop effective policies consistent with the general goals of higher education and that institutions be evaluated on the degree to which they meet these goals.

Higher education responded to these public concerns in very significant ways. During the 1980s the assessment movement within the academy emerged in response. This movement was part of a larger national trend that emphasized organizational quality in both the public and private
sectors (Bogue and Saunders, 1992). Due to external pressure (National Governor's Association Center for Policy Research and Analysis, 1986) and calls for reform from within (Roaden et al., 1987), those who crafted higher education policy embarked on creating processes intended to assess both the amount and quality of student learning. By the end of the decade, state policymakers began to develop mandates and formal polices that called for public institutions to actively engage in assessment. Policymakers viewed assessment, in its initial stages, as a tool largely geared toward institutional improvement. That is, the primary audience for assessment results would be faculty and administrators within higher education institutions (Palomba and Banta, 1999).

The Emergence of Performance Indicator Systems

By the early 1990s however, amid a national economic downturn and a diminishing resource base in state appropriations, new questions were asked of higher education. Given that higher education tended to consume a significant portion of state tax dollars and other expensive and competing programs such as corrections and health care came to the fore, assessment results were coming to be viewed as a way of evaluating higher education's claims to public support (Bogue, Creech, and Folger, 1993). Pressure
from a skeptical public about the use of its resources forced higher education into a position of demonstrating its worth relative to the worth of other expenditures of public funds; that is, it was required to demonstrate a return on the public’s investment (Neal, 1995). By 1990 forty states had enacted policies that promoted the gathering of assessment data (Ewell, Finney, and Lenth, 1990).

Despite the rise and institutionalization of assessment both on campus and within state systems, external pressures for accountability and improved quality did not abate. Although, many faculty members resisted ardently (Palomba & Banta, 1999), administrators began to reluctantly accept the notion of assessment. Many leaders believed that by agreeing to more closely examine student learning, pressure from governors, legislators and coordinating boards would diminish. Assessment was viewed as a Faustian bargain—giving up some measure of autonomy out of fear that a failure to do so would lead to even more intrusiveness on the part of external stakeholders (Ewell, 1994).

A number of forces converged to propel the evolution of quality achievement and maintenance forward to a point where less focus was placed on institutional improvement and more emphasis was placed on the values of external accountability and efficiency. Fife (1995) identified three factors that
have led to a continued demand for evidence that higher education is worthy of the resources which it is afforded. First consumers of higher education are less likely to accept the quality of higher education on faith than they once were. Evidence concerning the demonstrated quality of higher education in general (or the experience at a specific institution in particular) is increasingly being demanded. Second, the perception that higher education is an important component in gaining and maintaining meaningful employment continues to increase. Thus, higher education is viewed as an integral factor in spurring economic development. Finally, higher education consumes a larger portion of disposable and discretionary income at the level of the public agency and the individual. In summary, perhaps, the stakes concerning higher education are higher than ever before.

Ewell and Jones (1994) identified four additional forces. First, higher education, as an enterprise, had become increasingly complex with significant proliferation in both the number and types of students attending college as well as the diversity of institutions themselves. Second, in the early 1990s, public institutions experienced a diminishing resource base attributable to state funding cutbacks. Due to economic downturn, higher education found itself competing more aggressively with other state services
to preserve at least its portion of shrinking state budgets. Colleges and universities often lost when pitted against corrections, health care and K-12 education.

A third factor that helped to precipitate this newfound emphasis on quality was the change of perception regarding how citizens viewed their colleges. Higher education had occupied a lofty place in the American psyche for generations. Indeed, education was viewed as a key mechanism in facilitating social mobility. Deference was given to those within the higher education system (faculty and administrators) with regard to the how institutions should be run. However, by the 1990s, no government entity including higher education was above being questioned or second-guessed. A shift occurred in which higher education went from being perceived primarily as a public resource to being perceived as a public investment a major tenet of the accountability movement (Gaither, Nedwek, and Neal, 1994). Such a seemingly slight shift had enormous implications for higher education as a whole. For now, some type or amount of return was expected of this investment. Figure 1 provides an overview visualization of this evolution.
Finally, concerns regarding the quality of undergraduate education continued to remain at the fore of the public agenda. The rigor of the curriculum was questioned and anecdotal evidence provided by critics suggested that college students were learning less than earlier generations of students (Sykes, 1988; Bloom, 1987). While the assessment movement served to stave off or at least stall some of this criticism, legislators, governors, and other external stakeholders became increasingly
skeptical with the notion that institutions of higher education could or would adequately police themselves.

Accreditation agencies have also been stepping up efforts to ensure that colleges and universities more closely examine issues of assessment and institutional effectiveness. At its annual meeting in December, 2001, SACS approved new accreditation criteria that specifically require all colleges and universities to develop quality enhancement plans. Embedded within these plans are efforts to assess student learning and demonstrate improvement. In previous years accreditation had been largely a perfunctory exercise in which institutions showed compliance with a series of "must" statements. This shift toward institutional effectiveness suggests that colleges and universities develop PINS in addition to other reasons cited.

As a result of these confluent factors, new mechanisms designed to indicate the degree to which institutions of higher education were performing their perceived obligations emerged. These mechanisms were designed in an attempt to provide various stakeholders with quick and understandable information. The last decade has seen the development and growth of PINS across the United States (Ewell, 1994; Burke, 2000; Nedwek, Gaither, and Neal, 1995; Layzell, 1999). The intent of such indicators is to present in a public fashion,
pertinent information regarding how well colleges and universities are meeting their goals as well as the expectations of a variety of stakeholders.

Measurement in Context

Despite the effort of those in the assessment arena over the past fifteen years recent research suggests that public institutions of higher education are not able to demonstrate that learning outcomes are taking place. A study by the National Center for Public Policy and Higher Education (NCPPHE) (2000) employing a variety of national experts, graded states’ efforts in higher education along a number of specific areas. While states differed significantly on most items, all states were given grades of “Incomplete” in the measurement of student learning. The authors of the report noted that measures have yet to be developed that allow for any meaningful comparison of student learning between states. Thus, while much attention has been paid to the subject of assessment, in practice, it could be argued that very little progress has been made regarding the ability of states to determine what students learn or how much they learn during their college experience.
Definition and Development of Performance Indicators

Performance indicators most often reduce complex information to simple numerical form. As a result, they have been defined as "policy relevant statistics" (Ewell and Jones, 1994), "an authoritative measure—usually in quantitative form" (Cave, Hanney, and Kogan, 1991, p. 24), as well as "ratios, percentages or other quantitative values" (Taylor, Meyerson, and Massy, 1993 p. x). However, Ewell and Jones (1994) provide for a more holistic and operational definition of performance indicators. In their definition a performance indicator is "a concrete piece of information about a condition or result of public action that is regularly produced, publicly reported, and systematically used for planning, monitoring, or resource allocation at the state or system level" (p. 7). Despite slight variations in definition, a primary characteristic of a performance indicator is its emphasis on efficiency or parsimony (Layzell, 1999; Burke, 1998; Ewell and Jones, 1994). In other words, an important goal of performance indicators is to provide the most information to a variety of stakeholders in the most efficient manner possible. This has led to quantitatively oriented indictors being preferred over those more qualitatively oriented. Burke (1997) defines

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
performance funding indicators as those that tie directly to or affect budget allocations.

By the late 1980s and early 1990s, performance indicators were indeed a very real phenomenon affecting higher education policy (Bogue, Creech, and Folger, 1993; Burke, 2001; Ewell, 1994; Layzell, 1999; Nedwek, 1995). During this early stage, performance indicators often developed quickly from information already being reported in some form at the state level. Ewell (1994) suggested policy was being formed through “legislation by FAX”. This meant that states were exchanging information on performance indicators often as a preemptive strike to avoid the prospect of mandated indicators from the state legislature or governor.

Conceptualization of Performance Indicators

As noted by Nedwek (1995) performance indicators may be presented or framed in a variety of ways to fit particular conceptual models. Astin’s (1985) resource/reputation model (Inputs-Environment-Outputs) provides one way of examining performance indicators. Briefly, inputs refer to the qualities (talents, skills, and characteristics) that students bring with them to higher education. Environment refers to the experiences the student has during college, while outputs are those talents/skills/characteristics the
college is develops above and beyond the inputs. He suggests that this particular model (Figure 2) served as the traditional manner in which institutions are judged to be effective.

Figure 2

![Diagram](image)

Astin suggests that most colleges rely upon a resource and reputation model as a way of demonstrating institutional effectiveness and accountability. In this model, input measures rely almost completely on an institution's resources (human, physical, and capital) and reputation (perception of others inside or outside of the field) as well as the characteristics of the student body, faculty, etc. Thus, little emphasis is placed upon the more important notions of value-added or talent development. Figure 2 illustrates Astin's I-E-O model.
Richardson (1994) suggests that performance indicators can be categorized in terms of the type of quality they purport to represent. He identifies the following five types of quality:

**Transcendent Quality**—Performance indicators of this are largely input or environment-based (in Astin’s terminology) and include such things as students’ entering SAT scores, the size of an institution’s library, or the percentage of faculty members who possess terminal degrees.

**Cost/Benefit Quality**—Performance indicators of this type focus on desire to demonstrate efficient use of campus resources. Examples of such indicators might include space/classroom utilization ratios or instructional or non-instructional staff costs.

**Process-Based Quality**—Overlapping considerably with Astin’s notion of environment, these performance indicators reflect the experience found within the institution itself. Examples of such performance indicators might include the availability of academic programs for students or the percentage of undergraduate students who participate in sponsored research programs.

**Product-Based Quality**—Product-based quality represents those things that are “added” to the product (primarily the student population) in order to improve it. Ideally,
product-based quality controls for inputs and is therefore similar to Astin's notion of talent development or value-added education. In terms of performance indicators an example might be the reported results of student assessment efforts that measure value-added in various ways.

**User-Based Quality**—This category of quality is intended to reflect a client's satisfaction with their educational experience. Examples of performance indicators in this area might include student/alumni satisfaction surveys or employer surveys.

Similar to Richardson (1994), Ewell and Jones (1994) divide performance indicators into four separate categories depending on the particular value at play. In their scheme, performance indicators fit into an **Input-Processes-Output** model (practically indistinguishable from Astin's I-E-O model) if the focus centers around Astin's concept of value-added or talent development. This type of performance indicator seeks to answer the question - In what ways is higher education improving or adding to the skills and abilities that students bring with them to the college experience? Ewell and Jones (1994) also suggest that performance indicators can be typed according to their emphasis on **Resource Efficiency and Effectiveness, State Need and Return on Investment, or "Customer" Need and Return**
on Investment. Within the resource efficiency and effectiveness emphasis are embedded values that place great importance on the degree to which an institution maximizes its production capacity. This category bears striking resemblance to Richardson's conceptualizations of process-based quality and cost-benefit quality. Emphasis is placed on optimizing the resources available to the institution. State need and return on investment represents a "macro" approach to performance. Embedded within this construct is the idea that policymakers view higher education as a significant public investment from which a sufficient return is expected and that it would be useful for performance indicators to resemble something akin to a stock portfolio summary. For example, performance indicators related to an identified state need for a type of workforce would focus on the degree to which that need has been met by higher education. Finally, this same concept is applied at the individual level in terms of customer need and return on investment. In this case, performance indicators would attempt to measure the degree to which a customer's, i.e. student's, need is met.

Burke (1998), borrowing from Richardson (1994), Astin (1985), and Ewell and Jones (1994) suggests an Input-Processes-Output/Outcome model. While inputs and processes
are consistent with prior conceptualizations by Astin, output and outcome are purposefully differentiated. Burke's distinction is that output represents the quantity of a product produced (graduates, credit hours, etc.), while outcome represents the quality or impact of programs on users (more consistent with Astin's constructs of value-added and talent development).

Additionally, Burke (1997), employing previous writings by Marshall, Mitchell, and Wirt (1989) and Richardson (1994) proposes a conceptualization of performance indicators that emphasizes policy-oriented, global and societal values. That is, performance indicators can be categorized by the policy value they advocate. These values include the concepts of choice, efficiency, equity, and quality. The concept of choice emphasizes institutional uniqueness. A performance indicator that reflects this concept would measure the unique nature or educational experience of a particular institution. Performance indicators oriented towards efficiency would focus on weighing the amount of resources used against the results achieved, a sort of cost/benefit analysis. Equity is a value represented in performance indicators that examines issues of access and diversity within an institution or system. (e.g. percentage of full-time faculty who are minorities). Finally, performance
indicators based on quality would examine institutional performance in a way consistent with Astin's notion of value-added or talent development.

Finally, Burke (1997) suggests that performance indicators can be classified according to other criteria. Most prominent are classifications that denote whether performance indicators were developed internally (within the institution) or externally and whether they were mandated (by legislation, executive order, etc.) or entered into voluntarily. Additionally, performance indicators may be classified with regard to their relationship to funding. It is important to note that Burke distinguishes between Performance Funding and Performance Budgeting. Performance funding is a scheme whereby some portion of an institution's funding is contingent upon the degree to which an institution performs on its set of indicators. The relationship between indicators and funding is tight and linear. Performance budgeting, however, is a looser concept whereby performance may be taken into account as a factor in funding. Performance budgeting is the strategy more commonly found within the states.

While initially PINS were developed as a method to demonstrate institutional effectiveness and accountability, the last several years have witnessed an increasing
occurrence of PINS being connected to the funding or budgeting process in some manner. Recent research (Ewell, 1994; Nedwek, 1995; Burke, 1997) suggests that pressure for accountability within the environment has caused policymakers to increasingly tie performance indicator results to higher education funding. The most notable case of this development is in South Carolina where an increasing proportion of an institution’s budget is contingent on performance along 37 separate indicators.

Use of Performance Indicators

A number of studies have chronicled the implementation and use of performance indicators in a variety of settings. Much of the early literature regarding performance indicators was more international in scope (Gaither, Nedwek, and Neal, 1994). Works by Kells (1993), Cave, Hanney, and Kogan (1987), and Cuenin (1987) helped to crystallize the issues surrounding the development and use of performance indicators in Europe. As described earlier, the genesis of performance indicators resulted from a general concern over the quality of higher education. The mid 1980s saw these concerns play themselves out within the policy arena. A notable quality of these early systems was their emphasis on institutional distinctiveness (Ewell, 1994). Strenuous efforts were made to prevent inter-institutional comparisons.
Such policies were consistent with the beliefs and values of many involved in the assessment movement (Ewell, 1994). The purpose of such initiatives was to spur quality improvement within public colleges and universities. Until the late 1980s and early 1990s only Tennessee’s performance funding initiative formally linked performance with budget allocations. Over the last decade a number of efforts have been made to determine the degree to which states are implementing PINS. A National Association of College and University Business Officers (NACUBO) study sought to identify the functional areas in which performance indicators could be used. Another effort by Peterson’s and the Association of Governing Boards of Universities and Colleges (AGB) developed a set of strategic indicators. However, both of these efforts are limited in the fact that they focus on traditional or “reputation-oriented” notions of quality. That is, much like the earlier stages of assessment and accreditation, emphasis was placed on reporting inputs and outputs. None of the indicators forwarded by these efforts addressed the effects or impact of the college experience.

Only recently has any systematic effort to examine the impact of performance indicators taken place. Surveys by the Rockefeller Institute of Government at SUNY-Albany (Burke
and Mondarresi, 1999; Burke, Rosen, Minassians, and Lessard 2000; Burke and Serban, 1997, 1998) have chronicled the use of performance indicators through the latter part of the 1990s. According to a survey conducted by the State Higher Education Executive Officers (SHEEO) approximately three-fourths of states either report or use performance indicators in some manner (Layzell, 1999; SHEEO, 1997). Nonetheless, studies conducted thus far have examined or compared performance indicators across a limited number of states. The Education Commission of the States (ECS) undertook one of the most prominent early studies of performance indicators during the early 1990s through support by the Fund for the Improvement of Postsecondary Education (FIPSE). The ECS study provided case studies of ten states that had developed PINS of some type. States studied included: Colorado, Florida, Illinois, Kentucky, New York (SUNY), South Carolina, Tennessee, Texas, Virginia, and Wisconsin. Each case was examined to provide an overview of the development of performance indicators among many of its public institutions as well as to identify and categorize performance indicators used.

In his review and summary chapter Richardson (1994) attempted to compare performance indicators according to the Input/Output/Outcomes model and attempted to categorize them
according to his quality definition model. In general, he found that while input measures were collected states surveyed placed greater importance on output indicators such as retention/graduation rates, student credit hours by discipline, the amount of sponsored research funds generated, etc. Performance indicators identified as outcome measures were less frequent. That is, performance indicators overall seemed to focus on volume (production) rather than quality. Interestingly, the only outcome measure used by a majority of the states examined in the study was the passing rate of graduates on various professional licensure exams. Examining the measures according to Richardson’s typology of quality shows that states implemented indicators that were primarily focused on product-based quality. Examples of such performance indicators include: enrollment, progression, retention, and graduation by race, gender, etc.; pass-rates on professional licensing exams; and external or sponsored research funds.

In a series of studies, Burke (1997, 1999) examined the use of performance funding indicators and found results dissimilar to those found in the ECS study. Little commonality among states was found. Of eleven states examined in his 1999 study only eight out of a possible sixty-seven indicators were used by four or more states.
These indicators included: retention and graduation rates; two to four-year transfers; faculty workload; institutional choice; licensure test scores; transfer graduation rates; and workforce and training development. Burke’s findings suggest that performance-funding indicators are more diffuse than might be expected.

Policy Values and Models of Excellence

The studies by Burke also examined the primary policy value (efficiency, equity, quality, and choice) embedded within by performance funding used by the eight states. Overall, findings suggested that often, indicators tend to emphasize efficiency and quality. The findings however varied significantly by state with both Missouri and Tennessee placing more emphasis on quality. In a separate study by incorporating the models of Astin (1985) Richardson (1994), and Ewell (1994), Burke (1998) developed a hybrid archetype based on three models: Resource/Reputation Model (faculty-oriented); Strategic Investment/Cost-Benefit Model (state-oriented); and Client/Customer-Centered Model (student-oriented). Burke (1997) found the strategic investment/cost-benefit model paired with the client/customer-centered model to be most reflective of the indicators implemented in the states examined.
While this research has proven useful in promoting our understanding of the characteristics and usage of performance indicators there are serious limitations to the studies cited. As increasing numbers of states have implemented or plan to implement performance funding/budgeting programs, attention has tended to focus on a small number of states that implemented such programs early on. Additionally, recent research has focused almost exclusively on the use of performance funding indicators and not performance indicators in general. In fact, no comprehensive study of the use of performance indicators at the national level has taken place since the SHEEO survey of 1997.

In their seminal work on performance indicators Sizer, Spee, and Bormans (1992) identified five general uses for performance indicators. These include:

1. Monitoring—to promote ongoing assessment of a program, institution or system
2. Evaluation—to measure the attainment of goals and objectives
3. Dialogue—to build a foundation for communicating with others about abstract concepts and goals
4. Rationalization—to promote a rational and coherent policymaking process
5. Resource allocation—to provide a rational basis for the allocation of resources

A review of the literature suggests that there are performance indicators that represent each of these uses.
In fact, various performance indicators may fall into more than one category. Unfortunately, there has been no research that has sought to categorize and summarize performance indicators from various states according to these uses.

Accountability vs. Institutional Improvement

One of the great tensions emerging from the use of performance indicators is whether they are primarily geared for external accountability or internal, institutional improvement. Richardson (1994) suggests that these categories are not necessarily mutually exclusive. In his analysis of performance funding indicators, Burke (1997) found that the vast majority (nearly two-thirds) of them reflected external concerns. This was most clearly demonstrated in states where indicators were mandated and prescribed.

Issues Concerning Performance Indicators

Ever since performance indicators appeared on the policy radar their limitations and weaknesses have been highlighted. Nedwek (1995) has articulated many of these concerns in describing how policy action can be impeded due to the shortcomings of PINS. A primary shortcoming is what Nedwek calls our "primitive understanding of process" (p. 76). By this he suggests that our understanding of what
occurs in the "black box" or process (analogous to Astin's environment framework) is not well understood. In other words, how inputs interact with particular environments to produce particular outputs or outcomes remains largely a mystery. Therefore, performance indicators tend to reflect those things that are easier to measure—namely inputs and outputs. Second, he suggests that there exists a lack of consensus regarding the proper use of performance indicators. While some might advocate the use of performance indicators as a method of ensuring accountability, others view their purpose as evaluating and improving institutional effectiveness. The same sort of discord might also be found by examining performance indicators across purposes suggested by Sizer, Spee, and Bormans (1992) above.

A bias toward quantitative information was the third weakness cited by Nedwek (1995) as well as by Gaither, Nedwek, and Neal (1994). This is largely driven by the desire for parsimony and efficiency in examining a wide range of indicators (Layzell, 1999). Yet, while performance indicators do tend to be concise, the result is a failure to establish a causal link between inputs, environment, and outcome. Essentially, they speak very little to the how and why of outcomes. Therefore, indicators are limited by our lack of data concerning the relationships between and among
actors and environments in the educational process. Fourth, it is suggested that the policy process itself serves as an impediment to effective performance indicator development and usage. Performance indicators suggest what is important to society or policymakers at a particular point in time. Nedwek (1995) suggests however that the environment in which policy is crafted can be quite volatile. What may be important to one particular player at any given time may be unimportant to another player or the same player at a different time. Fifth, a number of authors (Nedwek, 1995; Gaither, Nedwek, and Neal, 1994; Ewell, 1994) suggest that performance indicators, in general, have not undergone rigorous examination with regard to their validity and reliability. This assertion buttresses a final weakness -- that policy has been formed and enacted prior to adequate conceptualization on the part of policymakers. Perhaps, to a certain degree, there is a bandwagon effect taking place in the policy arena consistent with Ewell's (1994) notion of "legislation by FAX".

Impact and Understanding

The task of synthesizing the findings from descriptive studies of performance indicators is complicated by the fact that researchers use different models (and different definitions within models) to classify indicators (See
Figure 3). As noted previously there is considerable overlap in most of the models used. Clearly, performance indicators and performance funding are increasing in popularity (Burke, 2000) as a policy mechanism. However, drawing any further inferences can become quite difficult. As mentioned, the experience with performance indicators in a number of states has been closely examined (ECS, 1994; Burke 1997; Freeman, 1997; Gaither, 1997; Cunningham, 1997; Stein, 1997; Sanders, Layzell, and Boatright, 1997). What emerges from these studies can be categorized as largely descriptive in the sense that they explain how PINS were developed contextually and what performance indicators were implemented in actual policy. However, there has been no study to date that has systematically analyzed the impact of PINS on various organizations and structures within higher education (funding levels, student learning, teaching, curriculum, etc). Much of this may be attributable to the relative "newness" of these programs. Nonetheless, our knowledge and understanding of this important phenomenon is severely limited. Table 1 provides for a summary of the major contributions made by various authors with regard to PINS.
The degree to which the use of performance indicators has affected higher education remains unknown. Burke and his colleagues at the Rockefeller Institute of Government have addressed this issue in a very general sense through a survey of performance budgeting and funding. One item asked

<table>
<thead>
<tr>
<th>Author</th>
<th>Model</th>
<th>Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Astin (1985)</td>
<td>Input-Environment-Output</td>
<td>System</td>
</tr>
<tr>
<td>Richardson (1994)</td>
<td>Quality-Based</td>
<td>Quality</td>
</tr>
<tr>
<td>Ewell &amp; Jones (1994)</td>
<td>Input-Processes-Output</td>
<td>System</td>
</tr>
<tr>
<td>Ewell &amp; Jones (1994)</td>
<td>Marketplace</td>
<td>Entrepreneurial</td>
</tr>
<tr>
<td>Burke (1997)</td>
<td>Policy Values</td>
<td>Values</td>
</tr>
<tr>
<td>Burke (1998)</td>
<td>Input-Processes-Output/Outcomes</td>
<td>System</td>
</tr>
<tr>
<td>Burke (1998)</td>
<td>Mixed Design</td>
<td>Mixed</td>
</tr>
</tbody>
</table>
campus leaders (N=1,910) among five states the degree to which they agree or disagree with the following statements.

Performance funding has...

Improved performance at institutions
Increased accountability to the state
Increased Responsiveness to state needs
Increased state funding for higher education

Source: Layzell, 2001

The respondents to the survey were equally split among the "agree", "neutral", and "disagree" categories for all items except accountability. Here, there was consensus that performance funding had lead to increased accountability to the state. It is important to keep in mind however, that this survey item inquired about performance funding and not performance indicators per se.

The experience with performance indicators over the last decade suggests that no perfect system has been developed. Nonetheless, the suggestions above indicate that states are seeking to hone existing performance indicators and develop others that will adequately inform various stakeholders. It is also helpful to keep in mind that the concepts of accountability and institutional effectiveness are evolving and not static. Clearly, the research agenda in this area needs to be forwarded and expanded to more closely examine the effects of these policies. In essence we know
what is being asked of the states but we are unaware of the effects such policies are having on the actual appropriations of public institutions. In addition to the number of terms and models that can add confusion to the understanding of PINS there is nature of politics and policy making at the state level. As Ewell (1994) has articulated, states have been in somewhat of a frenzy to implement measures of accountability within higher education. The policy environment could perhaps be described as mercurial and volatile. Essentially, policy has moved so quickly in this arena that extant literature has failed to keep up with what is taking place. Previous studies have provided for "policy snapshots" over the last decade. Revisiting this issue to determine "where states are" in this process would prove useful to policymakers and researchers within higher education.

Focus of Study

The purpose of this study will be to provide a complete, thorough overview with regard to the use of PINS within the fifty states. While research by Burke and Minassians (2001) and Burke, Rosen, Minassians, and Lessard (2000) and the previous annual Rockefeller studies on performance funding and performance budgeting dealt with the use of PINS, they largely failed to inform as to what type
of indicators are being implemented by the states. While this has been done in a limited manner (Ruppert, 1994; Rockefeller Surveys; 1997-2000; Burke, 1997) no data has been collected which specifically addresses the issue of data quality within PINS. In addition, evidence gathered from the World Wide Web suggests that recent research may not accurately or fully reflect the efforts by states to enact Performance Funding/Performance Budgeting with concomitant performance indicators. While previous studies (Burke, 1997; Burke and Serban, 1997; Gaither, Nedwek, and Neal, 1994; Richardson, 1994) have categorized performance indicators using a variety of conceptual schemes there is a dearth of research regarding the attitude and perception of policy makers toward performance indicators adopted within the states. Serban (1997) gathered the opinions of a variety of stakeholders in a survey that probed for attitudes toward performance funding and performance budgeting. Using the conceptual framework of Burke (1997), Serban categorized stakeholder preferences regarding performance funding/performance budgeting according to their primary emphasis (quality, efficiency, equity, and choice) with quality and efficiency being the dominant concerns.

The mission of the National Postsecondary Education Cooperative (NPEC) which was established in 1994 is "to
identify and communicate on-going and emerging issues
germene to postsecondary education and to promote the
quality, comparability, and utility of postsecondary data
and information that support policy development,
implementation, and evaluation." (www.nces.ed.gov/npec/)
This provides a simple and useful way of examining data
generated by PINS. In other words, it would be very helpful
to determine the policymakers' views regarding the degree to
which the data (generated by their own system's performance
indicators) conform to standards of quality, comparability
and utility. Similarly, a paper by MGT Associates (2001)
developed a list of characteristics or guiding principles
for the development of performance indicators. Such a
conceptualization provides a useful way of assessing the
quality, comparability and utility of these indicators.
<table>
<thead>
<tr>
<th>Guiding Principle*</th>
<th>Definition*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Linkage to Mission, Strategic Plan, and Policy Goals</td>
<td>The performance indicators should have internal and external credibility among all institutional stakeholders.</td>
</tr>
<tr>
<td>2. Stakeholder Involvement and Consensus</td>
<td>The performance indicators should incorporate and reinforce institutional missions and strategic plans, as well as broad policy goals.</td>
</tr>
<tr>
<td>3. Simplicity</td>
<td>The performance indicators should be developed through negotiation and consensus among key stakeholders.</td>
</tr>
<tr>
<td>4. Reliant on Valid, Consistent and Existing Information</td>
<td>The performance indicators should be based on data that are valid and consistent and that can be verified by third parties when necessary. The indicators should also be based on established data sources where possible in order to maximize credibility and minimize additional workload.</td>
</tr>
<tr>
<td>5. Recognizes Range of Error in Measurement</td>
<td>The performance indicators should be established with wide recognition that there are certain unavoidable ranges of error in any performance measurement activity.</td>
</tr>
<tr>
<td>6. Adaptable to Special Situations</td>
<td>The system of performance indicators should accommodate special institutional circumstances where possible.</td>
</tr>
<tr>
<td>7. Minimizes Number of Indicators</td>
<td>The performance indicators chosen should be kept to the smallest number possible in order to minimize conflicting interactions among the indicators and to maximize the importance of each indicator.</td>
</tr>
<tr>
<td>8. Reflects Industry “Standards” and “Best Practices”</td>
<td>The performance indicator system should reflect “industry” norms and standards where possible in order to allow for benchmarking and peer comparisons.</td>
</tr>
<tr>
<td>9. Incorporates Input, Process, Output, and Outcome Measures</td>
<td>The performance indicator system develop should have a balance of measures related to institutional inputs, processes, outputs, and outcomes.</td>
</tr>
<tr>
<td>10. Incorporates Qualitative and Quantitative Measures</td>
<td>The performance indicator system developed should incorporate both quantitative and qualitative measures in order to present the most complete picture of institutional performance possible.</td>
</tr>
</tbody>
</table>

*Source: MGT Associates (2001)
Data collected by this study would serve to fill a gap in our knowledge about the use of PInS in state systems. It would also serve to inform stakeholders on the degree to which senior state-level policymakers agree or disagree that the data generated by their systems is useful.

This study is based on the conceptual framework advanced by Burke (1997), Burke and Serban (1997; 1998), Burke and Moderessi (1999), and Burke, Rosen, Minassians, and Lessard (2000). This framework recognizes that various states are engaged in a variety of performance funding or performance budgeting schemes. At its core, it recognizes that decisions on funding rest upon performance of some type. However, the research thus far leaves a number of knowledge gaps in our understanding of performance indicators.

According to the most recent study (Burke and Minassians 2001), twenty-seven states were engaged in performance budgeting while nineteen states were engaged in performance funding—the two groups not necessarily being mutually exclusive. Yet two performance-funding states and eleven performance budgeting states indicated that PInS were not utilized. How then, do these states determine performance? The fact that they are engaged in schemes that, to varying degrees, appropriate funds based on performance
yet do not utilize performance indicators seems inconsistent and illogical. The basic question can be asked—upon what are these states basing performance? Finally, the question remains—do these performance indicators conform to the basic principles forwarded by NPEC and more operationally defined by MGT Associates (2001).

While a considerable amount of literature exists detailing the nature of performance indicators, there is a dearth of information regarding the practice of implementing PINS. The literature assumes that the existence of PINS indicates that such policies are, in fact, used to drive policy and budget making decisions at the state and institutional level. Is this, in fact, the case? If so, are institutions willing partners in this exercise? From the perspective of those who should know, are PINS having any effect whatsoever? Burke and Minassians (2001) addressed this issue in a very superficial manner. In the fifth annual survey by the Rockefeller Institute SHEFO were asked to what extent did PF/PB/PR have on state funding or overall performance of campuses. Perhaps, not surprisingly, most SHEFO indicated that such schemes had a moderate effect on funding or performance. Many states also indicated that it was too early to make any sort of definitive judgment on the
issue. Table 3 provides summary findings taken directly from the latest Rockefeller Institute survey (2001).
<table>
<thead>
<tr>
<th>Effect of Performance Budgeting on Funding</th>
<th>Effect</th>
<th>Percentage</th>
<th>Number of States</th>
</tr>
</thead>
<tbody>
<tr>
<td>Great Extent</td>
<td>0%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Considerable Extent</td>
<td>11%</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Moderate Extent</td>
<td>37%</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Minimal Extent</td>
<td>26%</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>No Extent</td>
<td>11%</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Cannot Judge</td>
<td>15%</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Extent of Performance Budgeting on Institutional Performance</th>
<th>Effect</th>
<th>Percentage</th>
<th>Number of States</th>
</tr>
</thead>
<tbody>
<tr>
<td>Great Extent</td>
<td>5%</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Considerable Extent</td>
<td>16%</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Moderate Extent</td>
<td>16%</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Minimal Extent</td>
<td>16%</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>No Extent</td>
<td>5%</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Cannot Judge</td>
<td>42%</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Extent of Performance Budgeting that Improved Institutional Performance</th>
<th>Effect</th>
<th>Percentage</th>
<th>Number of States</th>
</tr>
</thead>
<tbody>
<tr>
<td>Great Extent</td>
<td>3.7%</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Considerable Extent</td>
<td>7.5%</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Moderate Extent</td>
<td>33.3%</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Minimal Extent</td>
<td>18.5%</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>No Extent</td>
<td>15%</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Cannot Judge</td>
<td>22%</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Extent of Performance Reporting that Improved Institutional Performance</th>
<th>Effect</th>
<th>Percentage</th>
<th>Number of States</th>
</tr>
</thead>
<tbody>
<tr>
<td>Great Extent</td>
<td>0%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Considerable Extent</td>
<td>13%</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Moderate Extent</td>
<td>36%</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Minimal Extent</td>
<td>15%</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>No Extent</td>
<td>8%</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Cannot Judge</td>
<td>28%</td>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>
While such information is useful forcing SHEFO to place their states within discrete categories limits the understanding or the meaning behind these figures. Gaining a deeper understanding of this information provides the real intent behind this study.

The basic principles of quality, comparability and utility advocated by NPEC served as a frame of reference for these basic questions. Some of the basic issues with each of these principles included:

- **Quality**—Do the measures conform to the "best practices" found within the literature? Are these measures considered to be of high quality by those who report them?

- **Comparability**—Do the measures allow for useful comparison between and among similar institutions? Or, are the measures contextually bound?

- **Utility**—Do the measures have any significant or practical policy impact? Is the data used in any meaningful way?

The intent of this study was to fill a gap in the knowledge concerning the use of PINS. Such information will prove valuable to those involved in the development, implementation, or refinement of PINS. Essentially, information gleaned from this study serves the very practical purpose of informing policymakers about the degree to which PINS collect data that are of high quality and the
degree to which such data have any practical bearing on the appropriations/budgeting/funding process. In other words, are the data actually used in any real way to inform the process of appropriating public funds to individual institutions?

Perhaps the most dominant theory over the last forty years concerning the politics of budgeting has been that of Wildavsky (1964). Wildavsky’s theory of budget incrementalism has had a tremendous influence in understanding how budgeting within public agencies operates (Parker, 1997). Essentially, Wildavsky suggests that budgeting is a political process composed of small negotiations or “tactics” that result in relatively small, incremental changes in budgets appropriations. Advocates for particular parts of a budget asked for “a little more than last year” and often received it. Those who were responsible for developing and passing a budget (members of a public legislature) started out with the existing (previous year’s) budget and added a small amount to it. Wildavsky and his colleagues conducted a number of groundbreaking studies that seemed to confirm this hypothesis (Meyers, 1999). However, by the late 1980s incrementalism was viewed as an incomplete theory because, it was believed, Widalvsky had operated under assumptions in the 1960s that were no longer valid.
Nonetheless, incrementalism (and its variants) still have popular appeal both inside and outside of the academy. For this reason it was used in this study to serve as a frame of reference upon which to evaluate PINS. In their most pure form PINS would appropriate (or recommend appropriation) based on institutional performance. That is, money would be allocated based on data received from institutions. The question is whether budget decisions are based on such data or revert to a type of incrementalism suggested by Wildavsky.

Summary

PINS and its sub-variants have been at the fore of higher education public policy for nearly fifteen years. A considerable body of literature exists that addresses the purposes of these systems. Moreover, various scholars have developed their own typologies for such systems. While much is known about the types of data being collected, which states are collecting such data, and the intended policy outcomes of such systems, very little is known with regard to whether funding is affected in any manner by performance.

The intent of this study was to begin to fill that knowledge gap by collecting data from those who would be in a likely position to know. Additionally, this study intended to determine the degree to which components of PINS adhere to sound educational practice.
CHAPTER III

METHODOLOGY

What is known about PINS remains largely descriptive. That is, studies have simply focused on how many states have implemented PINS, the relationship between PINS and public funding, as well as attempts to group and categorize indicators by type, value constructs, focus, result, etc. This study was multiphase in its execution and utilized several approaches to inquiry. Figure 3 demonstrates the path of this study.

Figure 3

Stage One
Document Analysis
Investigated use of PINS via policy documents on WWW

Stage Two
Nominal Data
Check for conflicts or inconsistencies
Based Upon Web Findings and extant research
Identified states that are utilizing PINS in various forms

Stage Three
Sent survey instrument to SHEFO

Data from surveys were interpreted/analyzed/ with follow-up interviews.

Interviews are transcribed and data were analyzed, etc.
The initial phase of this study involved a confirmatory analysis concerning the use of PINS within the fifty states. Although Burke et al. (2000) provided information regarding which states were engaged in PF/PB techniques along with states that had developed PINS, the pace of policy development is such that revisiting this issue was worthwhile. Lincoln and Guba (1985) identify records as written communication that have an official purpose. All states have made at least some of their public records and policy proposals readily available via the World Wide Web. Glesne (1999) suggests that the analysis and review of documents serve to corroborate other forms of data. In this case, a review of documents served to either strengthen or weaken the findings published by the Rockefeller Institute. A list of keywords was used to identify documents that pertain to performance indicators. As of yet, there is no uniformly accepted terminology for what Burke (1997) identifies as performance indicators or performance indicator systems. Therefore, each state agency may employ its own unique terminology for such systems. Once these records had been accessed they were analyzed for meaning.

Additionally, the content of these documents served as an initial foray into the use of the deeper mode of
inquiry. It seemed quite possible (and perhaps likely) that policies within states changed recently. Inconsistencies between what was found on the World Wide Web and the most recent research findings were noted identified. During the next phase of this study a survey instrument was sent to all SHEEO. A list of members, with concomitant contact information, was provided by a staff person within the administrative office of SHEEO. The purpose of this survey was to enhance understanding with regard to whether PINS possess some of the basic characteristics outlined in the literature and suggested by MGT Associates (2001). Most important was whether data derived from such indicators conform to NPEC standards of quality, comparability, and utility. Additionally, the study sought to determine whether the implementation of PINS bore any practical implications to "normal" operations.

The survey instrument was comprised of both open and close-ended questions. Initially, surveys were sent electronically via e-mail with follow-up e-mails sent after two weeks had passed. It was anticipated that the response rate for this survey would be quite high. As noted in the previous chapter, very little is known about the effects that such policies are having upon a variety of stakeholders. The primary lens through which this study
would be examined is the perspective of the SHEFO who is ultimately responsible for collecting and disseminating PINS data. Such individuals are in a unique position to judge the efforts of the respective PINS utilized in their state.

The findings from the surveys were recorded, stored and subsequently analyzed for meaning. During the next phase of the study individual telephone interviews were conducted with representatives of various entities from one state that utilized PINS.

Figure 4
Methodological Approach

The primary purpose of this study was to gain an understanding of the specific components that comprise various PINS and the effect that the implementation of such policies has had from the perspective of SHEFO. The three NPEC principles as well as many of the MGT Associates principals served as the framework for these questions. The following list contains the survey questions and its concomitant NPEC and/or MGT Associates principle.

1. In what way(s) are the public four-year institutions within your state required to report their success in meeting specific performance criteria?

2. In what way(s) is this performance information used?

3. In your opinion, to what degree do these criteria determine in any direct or indirect way the amount of funding appropriated to individual institutions?

4. In your opinion, to what degree do these measures have “buy-in” among institutional stakeholders? [NPEC-Quality, Utility; MGT Principle #2]

5. In your opinion, to what degree do these measures attempt to reinforce each individual institution’s distinctive mission and strategic plan(s)? [NPEC-Comparability, Utility; MGT Principles #2, #6]

6. In your opinion, to what degree are these measures understandable to the public? [NPEC-Quality, Utility; MGT Principle #3]

7. Overall, to what degree are the measures valid, i.e., they measure what they intend to measure? [NPEC-Quality; MGT Principles #4, #5]

8. Do these measures rely on quantitative data (numerical data from various tests, instrument, or
formulas) or qualitative data (verbal and written prose)? [NPEC-Quality; MGT Principle #10]

9. In your opinion how well does this set of measures, taken as a whole, assess institutional performance? [NPEC-Quality; MGT Principle #9]

10. In your opinion, to what degree do these measures allow for benchmarking and peer comparisons? [NPEC-Comparability; MGT Principle #8]

These questions were pilot tested for understandability and validity. They were submitted to a former official of one state's central coordinating board who was quite instrumental in developing a set of new performance measures for that state. The suggestions made were incorporated and the questions adapted accordingly. While the intent of this study was not to develop a formal theory per se, it was intended to provide an understanding of whether PINS were comprised of indicators that were meaningful and useful.

As survey results were received (all electronically), data for each question were analyzed for content and meaning. Strauss and Corbin (1990) define this as "the process of breaking down, examining, comparing, conceptualizing and categorizing data" (p.61). The purpose of this step was to develop conceptual categories for the data. One way of accomplishing this is to take an entire document (in this case, survey results) and ask the
question—What appears to be going on here? What makes data from this document different from others? As Creswell (1998) suggests, a constant comparative approach is used in an attempt to “saturate” the conceptual categories. That is, continue looking for information until no new categories can be formed. In other words, responses to each survey question were analyzed for their overall thematic content. Then as data continued to be analyzed, different groups or categories of responses within each question began to emerge. Consequently, responses fell into distinct conceptual categories based on the overall theme of their content. This same process followed the interviews conducted with higher education officials in the mini case-study state. The ultimate purpose of this exercise was to reduce the data to a reasonably manageable set of themes or categories. At this point the data were analyzed to determine the degree to which they “fit” within Wildavsky’s theory of budget incrementalism.

Credibility and Validity of the Study

A strength of this study was its intent to capture an entire population of actors (SHEFO from all fifty states) rather than a portion of the population. Unlike other modes of inquiry, the approach used contained elements that helped to ensure credibility within the study (Creswell,
Thus, conceptual categories were "grounded" in the data. Furthermore, propositions generated were related back to the extant literature to check for consistency and fit. This study was also strengthened by the fact that multiple sources of evidence (document analysis, survey results, and individual interviews) were employed. Triangulation of these phenomena lent credibility and rigor to the study (Gall, Borg, and Gall, 1996). Finally, a strength of this study was the inductive nature of the data collection. Instead of requiring respondents to categorize and label their states' PINS this study purposively abandoned the use of jargon so as not to influence any of the responses.

Weaknesses

There were a number of areas of potential weakness in this study. This study did not attempt to track appropriations backwards through the decision-making process. Thus, some of the empirical evidence that could have demonstrated how performance affects appropriations is missing. No legislators were interviewed to determine how they themselves made appropriations decisions.

Additionally, the data collection used in this study meant that one person (the survey respondent) spoke for the entire state in matters related to PINS. This was a shortcoming of previous research (Rockefeller Institute) as
well. While some of the respondents were the SHEFO themselves others were staff members to whom the survey instrument had been forwarded via e-mail. In other words, how representative and knowledgeable was the survey respondent concerning the status of PINS within a particular state? While this is viewed as a shortcoming it proved to be the only efficient manner in which to try to collect such data for fifty states. The rationale was that, of all the individuals operating with the context of public higher education at the state level, the SHEFO or a staff person within their office would be the most equipped to address the questions on the survey instrument.

Ethical Considerations

Approval of the Human Subjects Review Committee of the College of William and Mary was obtained prior to data collection. The purpose of the study was presented to all survey respondents and interviewees.
CHAPTER IV
FINDINGS

The drive for increased accountability has spurred higher education policymakers to develop systems in which performance of colleges and universities can be measured. To date, this effort has continued to gain momentum within the public policy arena particularly at the state level. PINS are a manifestation of this effort to hold public higher education accountable to a variety of stakeholders. Surveys conducted by the Rockefeller Institute indicate that states are continuing to develop systems by which colleges and universities must demonstrate performance (Burke and Minassians, 2001). The purpose of this study was to take an inventory on the state of these PINS and thus confirm (or not) previous research findings. However, perhaps, more importantly this study intended to determine the degree to which PINS were affecting public appropriations to colleges and universities. A research approach was used that purposefully abandoned some of the assumptions of previous research.
This study started with the initial question, "What is the status of PINS within the fifty states?" In order to develop a more thorough understanding regarding the effectiveness of such programs it was important to determine, as much as possible, which states appeared to be engaged in the use of PINS. As previously mentioned, the Rockefeller Institute has been publishing an annual status report of performance funding, performance budgeting and performance reporting for a number of years. The most recent report (Burke & Minassians, 2001) indicated that 27 states were engaged in performance budgeting while 19 states were engaged in performance funding. It is also important to remember that the two groups are not necessarily mutually exclusive. The findings by this most recent survey served as a backdrop to an important issue. Although SHEFO (Burke & Minassians' surveyed population) indicated whether or not their states were engaged in various PINS, was there evidence available to support these claims? Therefore, the initial phase of this study was a review of the status of PINS conducted via a search on the World Wide Web. However, based upon the review of documents via the World Wide Web it was extremely difficult to determine what specific policy mechanism was in place, i.e. performance funding, performance budgeting, or performance reporting. Therefore,
the search focused on locating evidence of any of these mechanisms. The website of the State Higher Education Executive Officers' website (www.sheeo.org) provided active links to each state's central higher education agency. Each agency's website was reviewed regardless of the Rockefeller Institute's findings. Findings for each of the fifty states are in the table below. The most recent Rockefeller Institute report by Burke and Minassians (2001) was used as a comparator to the findings.
### Table 4

Summary of PINS Evidence From the World Wide Web

<table>
<thead>
<tr>
<th>State</th>
<th>Burke &amp; Minassians PF, PB, PR</th>
<th>Evidence from WWW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>PB</td>
<td>A review of Alabama Commission on Higher Education documents found no evidence to support existence of PINS.</td>
</tr>
<tr>
<td>Alaska</td>
<td>PR</td>
<td>A review of the State of Alaska FY2003 Governor’s Operating Budget (<a href="http://www.alaska.edu/swbudget/FY03Redbook/redfinal.pdf">http://www.alaska.edu/swbudget/FY03Redbook/redfinal.pdf</a>) confirmed that data (performance measures) are collected and reported in a systematic manner.</td>
</tr>
<tr>
<td>Arizona</td>
<td>PR</td>
<td>In the Arizona Board of Regents 2000 Report Card (<a href="http://www.abor.asu.edu/1the_regents/reports_factbook/rptcrd/2000reportcard.PDF">http://www.abor.asu.edu/1the_regents/reports_factbook/rptcrd/2000reportcard.PDF</a>) included reference to performance measures. However, determining what those measures include was more difficult.</td>
</tr>
<tr>
<td>Arkansas</td>
<td>PF</td>
<td>According to the Arkansas Department of Higher Education’s Strategic Plan of the 2003-2005 biennium (<a href="http://www.arkansashighered.com/pdfs/StratPlan.pdf">http://www.arkansashighered.com/pdfs/StratPlan.pdf</a>) institutions are required to submit goals related to institutional mission. However, it would appear that these goals are rather general in nature and performance indicators per se.</td>
</tr>
<tr>
<td>California</td>
<td>PB, PR</td>
<td>The public document entitled “Performance Indicators of California Higher Education, 2000: The Seventh Annual Report to California’s Governor, Legislature, and Citizens in Response to Assembly Bill 1808” by the California Postsecondary Education Commission contained the latest results of California’s PINS. Thus, it would appear that CA is actively engaged in collecting performance data.</td>
</tr>
<tr>
<td>Colorado</td>
<td>PF, PB</td>
<td>Various Colorado Commission on Higher Education publications including the Performance Funding Process for FY 2001-02 (<a href="http://www.state.co.us/cche/qi/performanceprocess.html">http://www.state.co.us/cche/qi/performanceprocess.html</a>) support previous findings that CO is engaged in PF.</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th>State</th>
<th>Evidence from WWW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connecticut</td>
<td>A report by the Connecticut Department of Higher Education entitled “Higher Education Counts: Accountability Measures for the New Millennium” (<a href="http://www.ctdhe.org/info/pdfs/Accountability2002.pdf">http://www.ctdhe.org/info/pdfs/Accountability2002.pdf</a>) detailed the use of a PINS within CT. However, it was difficult to determine whether such a PINS had any effect upon the funding or budgeting processes.</td>
</tr>
<tr>
<td>Delaware</td>
<td>A review of the Delaware Higher Education Commission website confirms that Delaware is currently not engaged in the use of PINS.</td>
</tr>
<tr>
<td>Florida</td>
<td>The Florida Board of Education’s “State University System Accountability Report” (<a href="http://www.fldcu.org/planning/accounreport/2001_Final_Accountability_Report.pdf">http://www.fldcu.org/planning/accounreport/2001_Final_Accountability_Report.pdf</a>) (2001) details Florida’s use of performance measures. However, it is not clear the degree to which these measures have any relationship with the budgeting or funding processes.</td>
</tr>
<tr>
<td>Hawaii</td>
<td>The University of Hawaii’s “Benchmarks/Performance Indicators Report, 2000 Update” (<a href="http://www.hawaii.edu/ovppp/assessment/benchmarks/benchmark00.pdf">http://www.hawaii.edu/ovppp/assessment/benchmarks/benchmark00.pdf</a>) details the UH system’s PINS. However, it is not clear the degree to which these performance indicators affect funding or budgeting.</td>
</tr>
<tr>
<td>Idaho</td>
<td>No evidence could be found relating to the reporting of performance measures within Idaho’s Board of Education website.</td>
</tr>
<tr>
<td>State</td>
<td>PF, PB, PR</td>
</tr>
<tr>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td>Illinois</td>
<td>PB, PR</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Indiana</td>
<td>None</td>
</tr>
<tr>
<td>Iowa</td>
<td>PB</td>
</tr>
<tr>
<td>Kansas</td>
<td>PF, PB, PR</td>
</tr>
<tr>
<td>Kentucky</td>
<td>PR</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Louisiana</td>
<td>PF, PB, PR</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Maine</td>
<td>PB, PR</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Maryland</td>
<td>PB, PR</td>
</tr>
<tr>
<td>State</td>
<td>Evidence from WWW</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>A review of the Massachusetts Board of Higher Education website (<a href="http://www.mass.edu/p_p/home.asp?id=4">http://www.mass.edu/p_p/home.asp?id=4</a>) provided evidence that some type of PINS is under development.</td>
</tr>
<tr>
<td>Michigan</td>
<td>A review of the state of Michigan’s government website (<a href="http://www.michigan.gov">http://www.michigan.gov</a>) provided no evidence that Michigan is engaged in the use of PINS.</td>
</tr>
<tr>
<td>Minnesota</td>
<td>A review of the Minnesota State Colleges and Universities website (<a href="http://www.mnscu.edu">http://www.mnscu.edu</a>) provided no evidence that Minnesota is engaged in the use of PINS.</td>
</tr>
<tr>
<td>Mississippi</td>
<td>Within the Mississippi Board of Trustees Institutions of Higher Learning website a document Plan of Excellence: IHL Strategic Plan (<a href="http://192.103.84.28/strategic_plan.asp">http://192.103.84.28/strategic_plan.asp</a>) details the measures which comprise the PINS.</td>
</tr>
<tr>
<td>Montana</td>
<td>A review of the Montana University System (<a href="http://www.montana.edu/mus/">http://www.montana.edu/mus/</a>) website found no evidence of PINS development.</td>
</tr>
<tr>
<td>Nebraska</td>
<td>A review of the Coordinating Commission for Postsecondary Education (<a href="http://www.ccpe.state.ne.us/PublicDoc/CCPE/default.asp">http://www.ccpe.state.ne.us/PublicDoc/CCPE/default.asp</a>) revealed no evidence of PINS.</td>
</tr>
<tr>
<td>Nevada</td>
<td>A review of the University and Community College System website (<a href="http://www.nevada.edu">http://www.nevada.edu</a>) revealed no evidence of PINS.</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>A review of the University System of New Hampshire website (<a href="http://www.usnh.unh.edu">http://www.usnh.unh.edu</a>) revealed no evidence of PINS.</td>
</tr>
<tr>
<td>New Jersey</td>
<td>The New Jersey Commission on Higher Education’s Sixth Annual Accountability report is one of several documents (<a href="http://www.state.nj.us/highereducation/ar06.pdf">http://www.state.nj.us/highereducation/ar06.pdf</a>) that detail the implementation of PINS.</td>
</tr>
<tr>
<td>State</td>
<td>Evidence from WWW</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>New Mexico</td>
<td>The New Mexico Commission on Higher Education report, &quot;Aiming for Excellence&quot; (<a href="http://www.nmche.org">http://www.nmche.org</a>) details the state’s PINS.</td>
</tr>
<tr>
<td>New York</td>
<td>A review of the New York State Department of Education, Office of Higher Education website (<a href="http://www.highered.nysed.gov/Quality_Assurance/home.html">http://www.highered.nysed.gov/Quality_Assurance/home.html</a>) reveals the state’s effort at developing PINS.</td>
</tr>
<tr>
<td>North Carolina</td>
<td>The document “Accountability Overview and Report on Campus Visits” (<a href="http://www.northcarolina.edu/docs/assessment/AccOvuRptCampVis00-01.pdf">http://www.northcarolina.edu/docs/assessment/AccOvuRptCampVis00-01.pdf</a>) published by the UNC Board of Governors indicate that the state is engaged in the use of PINS.</td>
</tr>
<tr>
<td>North Dakota</td>
<td>The North Dakota University System’s “Accountability Measures Report” outlines the state’s effort in developing PINS.</td>
</tr>
<tr>
<td>Ohio</td>
<td>The Ohio Board of Regents’ College and University Performance Report: Student Outcomes, Experiences and Campus Measures (<a href="http://www.regents.state.oh.us/perfrpt/">http://www.regents.state.oh.us/perfrpt/</a>) provide evidence that the state is engaged in the use of PINS.</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>A review of the Oklahoma State Regents for Higher Education website (<a href="http://www.okhighered.org/">http://www.okhighered.org/</a>) revealed efforts at data collection but lack of systematic approach characteristic of PINS.</td>
</tr>
<tr>
<td>Oregon</td>
<td>The Oregon University System’s Performance 2000-2001: System and Institution Summaries and Report Cards (<a href="http://www.ous.edu/aca/performance-ex-sum-12-01.html">http://www.ous.edu/aca/performance-ex-sum-12-01.html</a>) detail the state’s use of PINS.</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>The Pennsylvania State System of Higher Education’s, “The Accountability Imperative” (<a href="http://www.sshechan.edu/sppi3.htm">http://www.sshechan.edu/sppi3.htm</a>) provides evidence that the state is actively engaged in the use of PINS.</td>
</tr>
<tr>
<td>Rhode Island</td>
<td>A review of the Rhode Island Board of Governors for Higher Education (<a href="http://www.ribghe.org/riohe.htm">http://www.ribghe.org/riohe.htm</a>) revealed no evidence of PINS.</td>
</tr>
<tr>
<td>State</td>
<td>PF, PR</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>South Carolina</td>
<td>PF, PR</td>
</tr>
<tr>
<td>South Dakota</td>
<td>PF, PR</td>
</tr>
<tr>
<td>Tennessee</td>
<td>PF, PR</td>
</tr>
<tr>
<td>Texas</td>
<td>PF, PB, PR</td>
</tr>
<tr>
<td>Utah</td>
<td>PB, PR</td>
</tr>
<tr>
<td>Vermont</td>
<td>None</td>
</tr>
<tr>
<td>Virginia</td>
<td>PB, PR</td>
</tr>
</tbody>
</table>
An initial glance at the data suggests some inconsistencies between data gathered by the Rockefeller Institute and data available via the World Wide Web. For example, based on information gathered on the web, there was no evidence to suggest that Wyoming was engaged in any type of PINS scheme. Nonetheless, Burke & Minassians (2001) found that Wyoming was engaged in performance reporting. In fact, evidence gathered from the World Wide Web failed to support the claim that various forms of PINS were used in the following states: Alabama, Idaho, Kansas, Maine, Michigan, Minnesota, Nebraska, Nevada, and Wisconsin.

To gain a deeper understanding regarding the use of PINS within the states, a survey (see Appendix A) was mailed to all fifty SHEFO. This survey had several purposes. First,
it would serve to provide more evidence regarding the status of PINS within the states. More importantly, however, it would probe more deeply to determine if such PINS (in particular PF and PB) were having any real effect on appropriations. Finally, it would provide evidence regarding the degree to which PINS in practice support the principles forwarded by NPEC and MGT Associates.

Survey Data Collection

Surveys were sent to all SHEFOs in May of 2002. In some cases surveys were completed by the SHEFO themselves while, in most cases, the survey was forwarded to and completed by a staff person within the state's central higher education coordinating agency. Surveys were sent to listed SHEFO in all fifty states. Responses were provided by thirty states for a return rate of 60%. The responses to the questions were analyzed for content and meaning. What follows is a question-by-question summary of the findings.

Question 1. In what ways are the public four-year institutions required to report their success in meeting specific performance criteria?

The purpose of this question was to determine if performance criteria were reported in some systematic manner among the four-year public colleges and universities. The
question purposefully did not include the terms performance funding or performance budgeting. The intent was to get a pure sense of what was asked of institutions without SHEFO having to commit to a predefined scheme.

Responses to this question were placed in one of three categories. As illustrated in Table 5, the first category was one in which the states indicated that they were not currently required to report criteria. Three states fell within this category. Next, a number of states indicated that there were plans for developing reporting requirements but that such plans were not yet formal policy. Six states fell within this category. Finally, the majority of states indicated that they were required to report performance at some level as the result of an existing formal policy. Twenty states fell within this category.

Table 5

<table>
<thead>
<tr>
<th>Status of Reporting Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not required (3)</td>
</tr>
<tr>
<td>Delaware</td>
</tr>
<tr>
<td>Indiana</td>
</tr>
<tr>
<td>Nebraska</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Question 2. In what way(s) is this performance information used?

The purpose of this question was to determine if the information that is gathered is used in any meaningful way. Responses to this question were gathered from the twenty-six states that either had or were developing some type of PINS (Table 6). Some states provided multiple responses to these questions. The data were coded based on whether the indicator usage was based on information gathering/reporting (I), evaluation (E), or appropriation of funds (A) either through PF or PB. Examples of I responses included: "used in policy discussions"; "inform constituents"; "generate reports to governor/legislature". Examples of E responses included: "measure against peers"; "evaluate college presidents"; "measure progress toward 2020 goals"; and "evaluate institution in relation to the strategic plan". Finally, examples of appropriation of funds responses included: "determine funding"; "to drive performance funding"; determine presidential raises and contract extension"; and "report back to allocate a pool of money". Some responses included multiple uses of the data. For example, "used by legislators to ask questions in appropriation hearings" contained elements of all three
purposes—providing information that is perhaps used in an
evaluative sense for the purpose of appropriation.
Information—Responses that centered on the theme of
information focused reporting to some external audience.
The most common audiences cited were the legislature,
governor’s office, and general public. Information was the
dominant theme in ten of the comments. Evaluation—Responses
that centered on the theme of evaluation focused on the
evaluation of the institution and, in one instance, the
college president. Evaluation was the dominant theme in
four of the responses. Finally, responses that centered on
the theme of Appropriation focused on information being
used to drive appropriation decisions in most cases by the
legislature. Appropriation was the dominant theme in nine
of the responses.

Table 6

Primary Use of Performance Data

<table>
<thead>
<tr>
<th>Use</th>
<th>Appropriation</th>
<th>Evaluation</th>
<th>Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of States</td>
<td>9</td>
<td>4</td>
<td>10</td>
</tr>
</tbody>
</table>
Question 3. In your opinion, to what degree do these criteria determine in any direct or indirect way the amount of funding appropriated to individual institutions?

The purpose of this question was to find out from the SHEFO whether reporting affected appropriations in any way. It is important to remember that SHEFO were not asked if their state engaged in performance funding or performance budgeting as previous surveys have done. It seems quite possible that a state may have a formalized policy of performance funding or performance budgeting, but, in fact, the criteria may not affect appropriations. Additionally, Wildavsky’s notion of budget incrementalism was not presented as an alternative to explain the degree of variation in appropriations.

For the most part, states that responded indicated the impact of reporting had a marginal effect on the amount of funding appropriated to institutions. Only one state, Tennessee (5.45%), mentioned a percentage of the budget affected by performance reporting of greater than one percent. A number of states mentioned that the impact on appropriations was minimal. Nevada, for example, noted that such policies had “no impact currently” on state appropriations. A respondent from Washington noted that they have “zero impact.” In most cases, enrollment targets
largely drove any appropriation based on performance. The response from Idaho indicated that the impact was "...very limited, only enrollment is used to drive a very small part of the funds." A respondent from Ohio added "some of our funding for campuses is directly tied to performance i.e., enrollment growth." Additionally, a number of states indicated that appropriations were affected when new money was available. Interestingly, those who view incrementalism as the model that best explains legislative behavior suggest that it is more commonly found in situations in which budgets are increasing. However, during flat budgets or deficits the impact PINS was negligible. A prime example is California as the respondent noted, "this state is subject to tremendous deficits and cuts were made. Measures are not used in face of this crisis—the system goes out the window in times of crisis." Similar sentiments were echoed by South Carolina, "[appropriations are affected] very directly when there is new money—not as directly when budgets are flat."

Only a few states indicated that their PB/PF plan operated as conceptually defined regardless of contextual factors. The respondent from Colorado suggested, "new general fund is determined by the legislature, the performance funding system allocates the percentage of the total that will go to each institution's governing board."
Question 4. In your opinion to what degree do these measures have “buy-in” among institutional stakeholders?

The purpose of this question was to determine the degree to which performance measures conform to the NPEC standards of quality and utility. Measures with significant “buy-in” would likely be of higher quality and have greater utility. Additionally, this question addresses the second principle of performance indicator development as described by MGT Associates, stakeholder involvement and consensus.

Responses to this question were somewhat varied. An analysis of the data revealed four distinct responses emerging: significant buy-in, gradual buy-in, buy-in by measure to type of institution, and little to no buy-in. For the most part, respondents indicated that there was a significant level of buy-in mainly due to institutions having the opportunity for input when the measures were developed. For example, the respondent from Idaho noted, “institutions and board developed measures together, so they have bought in.” Similarly, New Jersey’s respondent noted that “institutional representatives were consulted in the process of formulating PF measures and any changes are discussed with the colleges.” Continuing with this trend, Ohio’s respondent suggested “our performance reporting
process has heavy buy-in among institutional stakeholders as they helped craft the process."

Some respondents indicated that buy-in among institutional stakeholders was more gradual. Missouri’s respondent noted that, "Measures developed in consultation with stakeholders increased buy-in over the years...."

According to the respondent from Pennsylvania, "At first universities were skeptical of both performance funding... Now that they have seen them in operation those who did well like them and those who didn’t don’t."

Some states suggested that either a portion of the measures had buy-in or that a particular segment of institutions viewed the measures favorably. The respondent from South Dakota noted that, "The campuses did buy in to some measurements." Similarly, Wisconsin’s respondent offered this response, "There is some variation in the level of buy-in depending on the specific stakeholder and specific measures they address...." The response from Texas echoes these same sentiments, "Virtually all institutions except the flagship universities complain that the standard set of performance measures does not accurately reflect the mission of their institution."

Finally, a number of institutions indicated that there was little to no buy-in among institutional stakeholders
regardless of institutional type. The response from South Carolina illustrates this point, "Institutions do not like this system. They do not like to be compared with each other in this state on a numerical basis". Washington's respondent noted that, "There is little buy-in institutionally though some institutional researchers seem interested in the planning process. By and large institutions comply, usually grudgingly, with legislative requirements." Nevada's respondent added "Very little at this time." In summary the responses can be categorized in Table 7.

Table 7

<table>
<thead>
<tr>
<th>Level of Buy-In</th>
<th>Number of States</th>
</tr>
</thead>
<tbody>
<tr>
<td>Significant buy-in</td>
<td>9</td>
</tr>
<tr>
<td>Gradual buy-in</td>
<td>4</td>
</tr>
<tr>
<td>Contextual buy-in</td>
<td>4</td>
</tr>
<tr>
<td>Little to no buy-in</td>
<td>4</td>
</tr>
</tbody>
</table>

Question 5. In your opinion to what degree do these measures attempt to reinforce each individual institution's distinctive mission and strategic plan?

The purpose of this question was to determine the degree to which performance measures took institutional
uniqueness into account. This purpose is consistent with the NPEC principle of comparability as well as the sixth principle of performance measures development by MGT Associates. That is, the measures should possess some adaptability to context.

Responses to this question varied greatly. Some responses indicated that measures did or were crafted with the intention of reinforcing the distinguishing features of institutions. However, other responses indicated a more "one size fits all" model was the norm in their state.

First some respondents indicated that the measures did serve to reinforce each individual institution's distinctive mission and strategic plan. Respondents indicated that at least some of the measures served this purpose. The respondent from South Carolina indicated, "Very much so—of the fourteen indicators four are institutionally mission-based and directly tied to strategic planning, mission, and goals. North Dakota's respondent suggested, "There is a direct link to campus strategic plans" In the case of Virginia, the respondent offered, "Each institution is represented by a separate report introduced by its mission statement. The reports include a section of institutions specific measures selected by the institutions to highlight its mission, achievement, or strategic direction." The
respondent from Washington indicated, "We ask the institutions to develop institution specific measures in addition to the statewide measures. In some cases measures reflect the priorities of the institution. In other cases they don't."

Another common response to this question centered on the notion that an effort was made to develop measures that would take institutional distinctiveness into account. Colorado's respondent noted that, "each individual institution's role is taken into account". While Iowa's respondent said, "The measures are supposed to be designed round the institutional strategic plans so that progress can be identified and marked." Such responses indicate and understanding of the need to include such information. However, the degree to which such PINS actually do reinforce institutional distinctiveness remains unclear.

Finally, there were states whose respondents indicated that there was very little reinforcement of institutional distinctiveness. Kentucky's respondent said simply, "Not at all." Nevada's respondent added, "They don't. The criteria are the same for all institutions." A respondent from South Dakota added, "Only one measure reinforces institutional mission."
The purpose of this question was to determine the degree to which states took institutional uniqueness into account. In some cases, there were a number of measures that accomplished this goal. In others, there may have been only one that would fall into that category. Still, there were states that indicated that none of the measures took institutional uniqueness into account. Three respondents noted that measures were tailored for different classification levels of institutions versus individual institutions. For purposes of this study, states were categorized as having at least one indicator that took institutional uniqueness into account or not (Table 8).

Table 8

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of States</th>
</tr>
</thead>
<tbody>
<tr>
<td>At least one of the measures took</td>
<td>17</td>
</tr>
<tr>
<td>institutional uniqueness into account</td>
<td></td>
</tr>
<tr>
<td>Measures did not take institutional</td>
<td>4</td>
</tr>
<tr>
<td>uniqueness into account</td>
<td></td>
</tr>
</tbody>
</table>

Question 6. In your opinion, to what degree are the measures understandable to the public?
The purpose of this question was to determine if measures were something that one of their intended audiences (in this case, the general public) could understand. Such a purpose is consistent with the third MGT Associates principle of performance measures development—simplicity. It also related to the NPEC principles of utility. Measures that are recondite in nature may be less likely to be used in any practical policy sense.

Responses from the states to this question varied. Some respondents indicated that the public probably had a very limited understanding of the measures. Others suggested that the measures were very understandable. Most respondents, however, indicated that the understandability of the measures used varied considerably.

A few respondents indicated that the measures weren't understandable. The respondent from South Carolina suggested, "Not very much—the system is numerical and complicated.". The respondent from South Dakota indicated that the measures would be understandable to policy makers but not the general public. At the other extreme were states like Kentucky, whose respondent noted, "[the measures] are very understandable—we concentrated on making the measures easily understood." Similarly, the Ohio respondent noted, "I
believe the measures are understandable to the public. Of course, as author of the report I am somewhat biased.”

Most respondents indicated that some of the measures were understandable while others were not. Virginia’s respondent noted, “The general public seems to find the profile measures (descriptive statistics) particularly interesting and useful, but many of the other performance measures do not have much meaning for the public because of the jargon used...” The measures most often cited as being understandable to the public were: graduation rates, class size, and first-year retention.

Interestingly, two respondents noted that the measures were not or had not been developed for public consumption. North Dakota’s respondent noted that, “To date it has not been compiled for public consumption.” Pennsylvania’s respondent added, “The public does understand the productivity measures but the internal measures have not been shared with them.”

In summary, most states (15) indicated that at least some of the measures were understandable to the public while (4) indicated that they were not or none of the measures were made available to the public (Table 9).

Table 9
Understandability of Performance Indicators

<table>
<thead>
<tr>
<th>Understandability</th>
<th>At Least Some</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of States</td>
<td>15</td>
<td>4</td>
</tr>
</tbody>
</table>

Question 7. In your opinion to what degree are the measures valid, i.e. they measure what they intend to measure?

The purpose of this question was to determine the degree to which measures complied with the NPEC principle of quality and the fourth and fifth principles of performance measurement development by MGT Associates. The fourth principle relates to the reliance upon "valid, consistent, and existing information" while the fifth principle "recognizes range of error in measurement".

The vast majority of respondents indicated that at least a substantial portion of the measures carried some validity. Ohio's respondent suggested, "We believe they are quite valid..." while Nevada's respondent, noted, "They are valid and measure certain important accomplishments." Some respondents indicated that a significant amount of time had been invested in the development of the measures with the implication being attainment of validity. For example, the respondent from South Carolina said, "They are very valid as we have spent seven years assuring that we are measuring
apples to apples". Additionally, a number of respondents indicated that they realized the complexity of this issue. Wisconsin's respondent mentioned that many of the indicators were straightforward while other "are complete and more difficult to measure" and that it is "necessary to use proxy measures that indicate performance in more indirect manner (e.g. time with faculty outside of the classroom as a proxy for faculty mentorship)." The respondent from Colorado noted that validity was "a never ending debate—these measures are as valid as any others."

Question 8. Do these measures rely on quantitative data or qualitative data?

The purpose of this question was to determine if multiple types of data were being used to measure performance. Such a purpose is consistent with NPEC's principle of quality as well as the tenth principle of performance measurement development by MGT Associates. The tenth principle simply states that effective measures "incorporate qualitative and quantitative measures".

Of the states that responded to this question (n=22), nine indicated that all of the performance measures used were quantitative (Table 10). Five states indicated that the measures were predominately quantitative while six indicated
that "both" were used. One respondent indicated that they would "need to look".

Table 10

<table>
<thead>
<tr>
<th>Nature of Indicators</th>
<th>Types of Indicators</th>
<th>Quantitative</th>
<th>Predominately Quantitative</th>
<th>Mix</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of States</td>
<td>12</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

Question 9. In your opinion, how well does this set of measures, taken as a whole assess institutional performance?

The purpose of this question was to determine the degree to which SHEFO believed the measures did, in fact, measure institutional performance. Such a purpose is consistent with NPEC's principle of quality as well as MGT Associates ninth principle. MGT Associates ninth principle suggests that effective performance measures "incorporate input, process, output, and outcome measures". In other words, do the indicators measure the whole range of institutional performance?

Responses to this question varied considerably. Only in a few states (MD, MO, NV, OH) did respondents indicate that institutional performance was measured "fairly well". Some
(ND, SD) respondents indicated that the measures were “minimally” effective in gauging institutional effectiveness. However, most respondents indicated that the measures provided for an incomplete picture of institutional effectiveness. The respondent from Wyoming noted that, “they only tell part of the story but we are continually trying to refine the measures...” Similarly, Washington’s respondent indicated that, “The measures provide a picture of a few important aspects of institutional performance but they certainly do not capture much.” From the viewpoint of the respondents there appeared to emerge some consensus that the set of measures was incomplete.

Question 10. To what degree do these measures allow for benchmarking and peer comparison?

The purpose of this question was to determine if performance indicators allowed for benchmarking and peer comparison. Such a purpose is consistent with NPEC’s principle of comparability and MGT Associates performance measures development eighth principle. This principle advocates standards that, “where possible...allow for benchmarking and peer comparison.”

Of the states whose respondents (n=22) answered this question fourteen indicated that at least some of their measures allowed for benchmarking and peer comparison. The
other eight respondents either indicated that their data were not used in this way or were not developed with this goal in mind. A number of states (NV, TX, VA) indicated that their measures were not intentionally developed with peer comparison in mind but that the data could be used in that manner.

Table 11
Use of Performance Indicators for Benchmarking

<table>
<thead>
<tr>
<th>Provide for Benchmarking?</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of States</td>
<td>14</td>
<td>8</td>
</tr>
</tbody>
</table>

If there is a general theme that has emerged as a result of analyzing the data, it is that there currently exists a strong degree of variation from state to state along a number of the factors cited by NPEC and MGT Associates. While the responses provided significant illumination regarding the use of PINS within the states this study it is necessary and useful to look more closely at a particular case. As a way to increase the validation of this study, one state was examined more closely to determine if other players within the higher education arena would respond similarly to the SHEFO. As a means of
verification it was decided that additional oral interviews would be conducted. A follow-up interview would be conducted with someone within the states' higher education coordinating agency. Additionally, at least one telephone interview with an individual responsible for reporting data at the institutional level and an interview with someone who worked in the legislative staff area would be given. Given the richness of the data provided and the fact that the state had been actively engaged in the use of PINS, Texas was selected as a mini case study.

It was discovered upon a telephone follow-up interview that the SHEFO response for Texas was, in fact, furnished by a senior staff member of the Legislative Budget Board (LBB) of the Texas Legislature. In order to substantiate the findings of the SHEEO survey follow-up interviews were conducted with three professionals who dealt with PF/PB in Texas. A follow-up interview was conducted with the original survey respondent (staff member of the Texas LBB) as well as interviews with a senior member of the institutional research staff at a large public university as well as with two staff persons in the Texas Higher Education Coordinating Board (THECB).
Follow Up Interviews

Within the state of Texas, institutional performance data are collected and reported to the THECB and the LBB. The first follow-up interview was with the original survey respondent, a senior staff member of the Texas LBB. Notes were taken during the interview. The respondent was asked to elaborate upon answers provided in the original SHEFO survey.

The responded was asked about the use of the data. The respondent noted that legislators use the data "as a basis for questioning and evaluating institutions during appropriation hearings." However, interest in such data varied considerably among legislators. It was noted that what often dictates the use of the data depends upon who is occupying key positions with the legislative committees. In the case of Texas, the Chairmen of the Senate Finance Committee and the House Appropriations Committee often dictate the degree to which the data will be used.

When inquired about the utility of the data, the respondent indicated that the indicators could be divided up into "key" and "non-key" categories, the former being comprised of graduation and retention rates and the latter comprised of most everything else. Additionally, legislators are able to view an institution's previous five years of data. Interestingly, the respondent also noted that, to
date, no institution has been penalized or significantly rewarded based upon the data provided to the legislature.

"The rule of thumb is incrementalism," noted the respondent. When the legislature convenes and committee meetings begin, legislators typically ask institutions (represented by their respective president) the following questions. "What have you done over the last fiscal term?" and "What are your problem areas?" Legislators who point to poor institutional performance in the data are usually met with a response by the president that suggests that the indicators do not adequately capture the mission and goals of the institution.

Another interview was conducted with a senior institutional researcher at one of Texas' major research universities. Because this respondent had not answered the original questions found within the SHEFO survey, it was decided that the interview would focus on those questions so as to provide some basis of comparison with the other interviewees. The following are the responses to the SHEFO questions.

1. In what way(s) are the public four-year institutions within your state required to report their success in meeting specific performance criteria?

The respondent noted that the institution is required to file an annual report to the TX HECB that indicates how well the institutions performed along selected indicators.
"If the institution comes in at five percent above or below its target, it has to submit a written explanation" noted the respondent.

2. In what way(s) is this performance information used?

To this question the respondent simply noted that the information was used to determine areas of funding need.

3. In your opinion, to what degree do these criteria determine in any direct or indirect way the amount of funding appropriated to individual institutions?

The respondent noted that, in the case of her institution, there had been no observable effect on appropriations.

4. In your opinion to what degree do these measures have "buy-in" among institutional stakeholders?

The respondent noted that buy-in among institutions was about as much as one could reasonably expect given the circumstances. She noted, "Buy-in as about as good as it can get."

5. In your opinion, to what degree do these measure attempt to reinforce each individual institutions distinctive mission and strategic plan(s)?

"They don’t” replied the respondent. She noted that the legislators don’t fully understand that not all of the
performance measures are particularly applicable to all types of four-year institutions.

6. In your opinion, to what degree are these measures understandable to the public?

The respondent noted that there was a general demand for information (particularly graduation rates) that was understandable to the public at large.

7. In your opinion, to what degree are the measures valid, i.e., they measure what they intend to measure?

The respondent noted that she believed they were valid in that they were very straightforward in nature.

8. Do these measures rely on quantitative date or qualitative data?

"They are all quantitative" she replied.

9. In your opinion how well does this set of measures, taken as a whole, assess institutional performance?

"It’s about as good as we can get," replied the respondent. She noted that there were too many variables (institutional type, student demographics, etc.) for these measures to truly assess institutional performance.

10. In your opinion to what degree to these measures allow for benchmarking and peer comparison?

The responded noted that the reports were used for benchmarking purposes and that the data generated from the
reports were used to make peer comparisons with both in-
state and out-of-state institutions.

Finally, interviews were conducted with two
representatives of the THECB. The first representative
worked in the area of finance within the THECB. He
responded to the first five questions of the survey because
they dealt more directly with finance issues. The following
are his responses.

1. In what way(s) are the public four-year
institutions within your state required to report their
success in meeting specific performance criteria?

“In my opinion we haven’t been doing a lot,” he noted.
He also noted that the THECB had recently asked for
efficiency measures. However, he added, “There has not been
a standardized mechanism for reporting.”

2. In what way(s) is this performance information
used?

“We can extract a lot of information. However, there
is very little funding attached to it—no definitive
programs tied to it. We’re talking small dollars and an
infinitesimal part of the overall budget. It has been
suggested that anywhere from three to five percent of the
budget rest on performance. That has been recommended to
the coordinating board but it has not gone anywhere.”
3. In your opinion, to what degree do these criteria determine in any direct or indirect way the amount of funding appropriated to individual institutions?

“It is very small. We’re in the midst of budget cuts right now. To the extent that they meet targets that’s good.”

4. In your opinion, to what degree do these measures have “buy-in” among institutional stakeholders?

“There is no standardization, each institution has set its own. There has been no attempt to come up with different types of measures.”

5. In your opinion, to what degree do these measures attempt to reinforce each individual institution’s distinctive mission and strategic plan(s)?

“I haven’t seen any impact in that area.”

The individual referred me to another THECB staff member who is more involved in developing the actual reports. Her responses are as follows:

1. In what way(s) are the public four-year institutions within your state required to report their success in meeting specific performance criteria?
"There is a report that goes to the LBB—lots of special request for information. For example, the number of graduates by CIP code."

2. In what way(s) is this performance information used?

"Essentially, these things decide how much money goes to higher education. It is a fair system because it is based on semester credit hours. Formula—this level of the formula at 81% or so and so."

3. In your opinion, to what degree do these criteria determine in any direct or indirect way the amount of funding appropriated to individual institutions?

I think that semester credit hours really drive the system. The other things "true performance" account for a small percentage of funds.

4. In your opinion, to what degree do these measures have "buy-in" among institutional stakeholders?

"Formula funding system has buy-in among institutions."

5. In your opinion, to what degree do these measures attempt to reinforce each individual institution's distinctive mission and strategic plan(s)?

Depends on how many students there getting.

6. In your opinion, to what degree are these measures understandable to the public?

The public doesn't know much about this stuff.
The respondent noted that she did not have a sufficient background and knowledge to address questions seven through ten. However, she noted that performance did seem to have an effect on several specialized programs developed in Texas. These were incentive-based programs that focused on workforce development in regions along the border of Texas and Mexico. Additionally, funds were available for colleges and universities that focused on teacher education programs. The funding for these very specific programs were the only ones that were considered performance based.

What became clear after these interviews was that institutions within Texas were, in fact, required to report performance. Yet, the degree to which performance affected the amount of appropriations institutions received was cloudy at best.

Verification

The findings of this chapter were sent to three noted experts in the field of higher education performance measures. This served to assist in verifying and corroborating the major findings of this study. One expert noted, "Overall, your findings are what I would have expected. I think many folks in higher education view performance indicators as just another reporting exercise with dubious utility." Another expert noted, "I am surprised
at the high response for the validity and reliability of the PINS. I bet you would get a different response from campus representatives." This, of course, points to one of the weaknesses of the study, its almost exclusive reliance on SHEFO data. Finally, after reading the results, one expert keenly noted, "If states are filing and acquiring all these data on higher education performance indicators and major decision makers are not paying any attention or the data have no effect, why are we doing all this?"

Summary

Responses to many of the survey questions were quite varied. Nonetheless, there was little evidence to suggest that appropriations allocated to institutions were affected by performance along a set of criteria. Instead, incrementalism seemed a more plausible explanation for changes in appropriations.

As for the indicators that comprise PINS, there was wide ranging opinion regarding the quality, comparability, and utility of the data. However, there was a general consensus that the quality of the data being gathered, while less than ideal, is as good as can be expected.

In summary, the findings suggest that PINS do not operate in a clean, linear fashion as has been suggested in the literature. While many such systems have been written
into policy and enacted across many of the states there is little evidence to support the notion that higher education is operating more effectively or efficiently than before.
CHAPTER V

CONCLUSIONS AND IMPLICATIONS

The purpose of this study was to develop a more thorough understanding as to the status of PINS within the fifty states and to determine the degree to which PINS work in practice. The most recent studies by Burke & Miassians (2002) indicated that at least forty-five states were engaged in at least one type of PINS scheme, PB, PR, or PF. However, data gathered on the World Wide Web (to links to the states' higher education coordinating agency via www.sheeo.org) showed a number of inconsistencies. First, based on the web findings alone, it was nearly impossible to ascertain in which type of PINS states were engaged. Any evidence of the three types was considered evidence of engagement in PINS. There were a number of states that, according the literature, were engaged in PINS but for whom no evidence of such involvement was found. Such were the cases of: Alabama, Idaho, Kansas, Maine, Michigan, Minnesota, Nevada, and Wisconsin. There are a number of plausible reasons for such discrepancies. The quality of state coordinating board websites varied considerably. Many
websites contained voluminous amounts of higher education reports and data while others were much more sparse. It would seem possible that some state sites failed to present the most current information available. However, this discrepancy was somewhat resolved with the SHEFO survey findings. States for which evidence of PINS involvement was expected but not found would be identified in the content of the SHEFO responses to question one. In fact, SHEFO responses from Idaho, Nevada, and Wisconsin indicated that these states were involved in PINS. SHEFO responses from Alabama, Kansas, and Minnesota indicated that PINS were in the development stage. Responses for the remaining states in this category (Maine and Michigan) were not provided. In the case of Nebraska, Burke and Minassians (2002) indicated Nebraska's involvement in PB. However, no evidence from the state's coordinating board served to corroborate this assertion. Additionally, the Nebraska SHEFO who responded to the survey indicated that the state was not currently engaged in any type of PINS.

Performance Indicators

A key component of this study was the focus on the use of PINS against the backdrop of what is considered to be good practice. The question emerges as to how well the components of PINS conform to the standards set out by NPEC.
and MGT Associates. A survey of the indicators and their use suggests there are both strengths and weaknesses in what is currently being used.

Strengths

Of the ten principles outlined by MGT Associates (2001), two areas emerge as strengths in terms of what is currently in practice. It would appear, based on the evidence, that current PINS do rely on valid, consistent, and existing information (Principle 4). Most of the PINS use data that are already gathered and reported (admissions, enrollment, graduation, matriculation, etc.). There are accepted standards (IPEDS, Common Data Set), which create a level of consistency and confidence in the information. Another strength is Principle 7 (minimizing the number of indicators). In only two cases (South Carolina, Texas) were the number of indicators considered excessive. It would seem that, in their current state, PINS lack the sophistication to go much beyond information that is already being reported.

Weaknesses

Two areas covered in the MGT Associates (2001) principles appear to be weaknesses of current PINS. Principle 9 suggests that PINS incorporate input, process, output, and outcome measures. Based on the evidence it would appear that current PINS lag considerably in terms of
true outcome measures. As defined by Burke (1998), outcomes represent the quality or impact of programs on users as conceptualized by the value-added and talent development constructs of Astin (1985). Current PINS do an insufficient job (and many do not even attempt) in trying to measure true outcomes of education such as student learning, cognitive/affective growth, and abstract constructs such as critical thinking and effective communication skills.

Another weakness of current PINS is their excessive reliance upon quantitative data. With very few exceptions, data are provided only in numerical form. While this has the advantage of efficiency and understandability there is a downside as well. Often, a piece of quantitative data does not convey any real meaning to an audience. It is as if all of the cognitive, affective, and behavioral experiences and outcomes of students can be represented by a number of digits. In many cases it would be helpful to have more meaning i.e. qualitative prose behind the quantitative data.

For most of the MGT Associates (2001) principles, the findings suggest no clear evidence exists to determine whether the criteria have been met by current PINS. Perhaps, current PINS lack the sophistication needed to make judgments about measurement error (Principle 5) or best practices (Principle 8). Additionally, responses were quite
varied with regard to mission linkage (Principle 1), stakeholder involvement (Principle 2), simplicity (Principle 3), and adaptability (Principle 6). There seemed to be no pattern to and a high degree of variability within the responses.

With respect to the NPEC principles of quality, utility, and comparability it would appear that current indicators used in PINS fall short of ideal standards. As mentioned previously, there is insufficient evidence to demonstrate that the indicators being used are of sufficient quality. Most of the indicators are input and output in nature and fail to address the more important outcomes of education. Such a finding is consistent with the latest Report Card of Higher Education (2002) published by the National Center for Public Policy in Higher Education. In terms of utility, the findings appear more complex. On the surface, data are being collected and reported. However, the degree to which these data are actually being used to drive decisions on appropriations remains unclear. What is clear is that many of those responsible for collecting and reporting these data do not exhibit confidence that they are being used in any systematic and logical manner. At this point it would also appear that PINS are insufficiently developed to provide for effective comparability of institutions. It would appear
that most SHEFO recognize the delicate tension between the myriad of types of higher education institutions and the need to provide common data that the public can understand. Based on the response it would appear that tension has yet to be resolved.

Use of PINS

If there is one theme that emerged regarding the use of PINS data, it is that many seem to view the process and exercise in almost fatalistic terms. In other words, the data generated by PINS were not going to make any real difference in appropriations. Legislators would use their own rationale (likely incrementalism) in developing the budgets. In qualitative research what is not said is often as valuable as what is said (Glesne, 1999). It is interesting and, perhaps, not surprising that no respondents (including South Carolina) suggested a tight, linear relationship between performance and appropriations. What seemed to emerge from these surveys was a sense of minimal confidence that the data finds its way to the legislature or plays an active role in legislative decision-making. No respondents indicated that funding had been cut as a direct result of poor performance on specific indicators. The only indicator that seemed to play a role in the appropriation process was enrollment. However, as mentioned earlier,
enrollment is the driving force in formula funding and not PINS per se.

How the information was used also interesting. As noted previously, responses were coded according to the dominant theme conveyed. In this case the three themes that emerged were: information gathering/reporting (I); evaluation (E); and appropriation of funds (A). Of the respondents, using data for the purpose of information gathering/reporting was the most cited response (N=10). Nine responses focused on appropriation of funds while four focused on evaluation. This suggests that PINS are primarily used to convey information and determine appropriation levels. Yet, there was very little indication that appropriation levels were directly affected by performance.

PINS vs. Incrementalism

The evidence from the Texas mini case study clearly suggests that decisions regarding appropriations conform with the theory of budget incrementalism developed by Wildavsky (1964). While PINS were used for appropriation in a number of the states there was no evidence to suggest that data from PINS were used, in any significant way, to determine appropriation amounts. The fact no SHEFO indicated that institutions had been punished for poor performance suggests that appropriation decisions are based on something
else. In fact, many SHEFO suggested that institutional budgets had remained stagnant or slightly declined over the last several years. While there is no evidence to support the notion that appropriations may have been decremental in nature it would seem safe to suggest that PINS played a negligible role in determining the amount of money going to particular institutions. The question of whether appropriations truly reverted to an incremental approach cannot be answered given the data limitations of this study. However, it is interesting to note that a number of respondents suggested an approach in operation, similar to forwarded by Wildavsky.

Themes

A significant theme that developed in examining these qualitative data is that there would appear to be tremendous variation among the states in terms how higher education is coordinated and managed. Rather than PINS having the effect of homogenizing the states' public higher education systems, they seem to highlight and reinforce the unique differences among states. In other words, PINS seem to adapt to the unique culture and circumstance of the higher education environment within a particular state as opposed to making them look more alike.
The findings from this study also suggest that if there were a tighter, more linear relationship between performance and funding, institutions of higher education would have more incentive to perform better. A unique aspect of this study was the phenomenon of "pushing to the test". In other words, what is measured is what becomes important. As long as what are measured are inputs and outputs, those are what will be perceived as most important. This serves to reinforce Astin's (1985) notion of a reputation-based model of excellence versus a talent-development model.

Costs Versus Benefits

It was not within the purview of this study to total the real costs associated with developing, implementing and maintaining a PINS within a state system of higher education. Suffice it to say the number of man-hours involved in developing, collecting, and reporting data from indicators plus the costs associated with developing technological infrastructure to support PINS are considerable. This seems very important in light of a recent economic downturn and retrenchment on the part of higher education. Most states are suffering from staggering deficits and reducing higher education spending is a convenient (and sometimes politically viable) manner in which to balance the budget (Selingo, 2003). It is interesting to note that many of the
current PINS were developed and implemented through the 1990s. That decade saw increased spending on higher education and growth in budget surplus. In a sense, PINS were crafted to help legislators and policymakers make decisions on appropriations in a time of relative prosperity. Are these same systems effective or even utilized in times of economic recession? In the case of Texas, a state which is undergoing its own fiscal crisis, decisions on appropriations (in this case, reductions) did not appear to emanate out of data. As the interviews from Texas suggest, there would appear to be no direct link between performance and appropriations. If this is the case, then PINS may be no more than window dressing developed by politicians to appease disgruntled taxpayers. However, Layzell (2003) points out that there are costs associated with not engaging in PINS. He notes that higher education would have likely failed to develop PINS on its own were it not for external pressures. Had higher education failed to implement PINS (and, in most cases had a voice in developing them) something more drastic would likely have been imposed.

Public pressure to account for the performance of colleges and universities is clear. Accrediting bodies and the majority of states now require institutions to collect data, measure performance, and report to the public. But
110

neither the technical adequacy nor the utility of the data now collected have been well established. States do not uniformly put performance data to use, raising questions about the costs and benefits of collecting and reporting. Institutions and the states that support them clearly have differing expectations about how performance data will be used. For institutions, the focus remains on institutional improvement particularly as it relates to student outcomes. Many of those who represent the state including legislators and taxpayers, while not discounting student outcomes, place a tremendous emphasis on efficiency, accountability and cost savings.

The results of this study suggest a reconsideration of the way performance data are collected and used. At the very least, good data can be used to increase transparency, and transparency is the foundation for accountability. Because "performance" remains difficult to define, and comparisons among institutions difficult to validate, rewards and sanctions may be impossible to connect to performance in any meaningful way. Nonetheless, operating transparently is at least a second-order result of PINS that the public and their elected representatives would understand and appreciate. Perhaps individual institutions could be asked to construct their own PINS in the spirit of greater
transparency, and rewards and sanctions based on how clearly and completely institutions account for their operations.

Remaining Questions and Directions for Future Research

The findings from this study raise several questions and suggest directions for future research. First, the findings of this study suggest that data derived from PINS are used minimally in the appropriations process. Additional studies employing a case-study methodology are necessary to gain a sense of what is taking place within each of the states. Based on the findings from this study in Texas, it would seem that use of PINS are contextual.

Second, empirically based studies which attempt to track appropriations "back" to the decision-making process would prove valuable as well. This could be accomplished via a regression analysis that attempted to ascertain the most powerful determinants of appropriations.

Third, studies that focused more directly at the campus level would prove useful as well. The perspective of longer term administrators at universities in states like Tennessee and South Carolina would illuminate many of the subtleties and nuances associated with PINS.

Finally, studies that focused more exclusively on the legislators (and their staffs) would provide for increased
understanding with regard to the attention span or life-cycle of policies such as PINS.

Summary

This study did not find evidence that PINS were affecting state appropriations within higher education. To the contrary, the evidence suggested that budget incrementalism may provide a better explanation for the primary effect on state appropriations. In terms of the indicators that comprise PINS, there seemed to be varying levels of meeting the NPEC standards of quality, comparability, and utility. In essence, it would seem that the indicators lack sufficient utility because they are not being used for the express purpose determining appropriations. There is some evidence that suggests that engagement in PINS has some value in itself in that it forces institutions of higher education to more closely examine themselves. However, as a policy lever they would appear to have very limited impact.

The case of Texas provided a significant window into actual PINS operation. What came out of that case was a sense that data collected provided some information (particularly to legislators and aides) that was useful but not terribly consequential. The fact that no one interviewed could connect performance (outside of enrollment) with
appropriations suggests that PINS, while appealing in theory, may not be useful in practice.
APPENDIX A

SAMPLE COVER LETTER AND SURVEY

Dear SHEFO:

I am conducting a study about the use of data in performance reporting. While much is known about the development and implementation of performance indicator systems very little is known about how or if performance data collected by the states are actually used. I would greatly appreciate a few moments of your time to get your professional perspective on this issue.

All responses are confidential and anonymity will be preserved throughout this study. Based upon the findings of this questionnaire we may follow-up with a question or two for clarification. The questions are included both within the text of this e-mail as well as in a MS Word attachment. Please respond using whichever format you prefer. Should you have any questions regarding this study please feel free to contact myself or the chairperson of my dissertation committee, Dr. David Leslie, at (757) 221-2349 or at dwlesl@wm.edu.

Your responses may be returned via e-mail to jmdavi@wm.edu or in hardcopy form to my address:

634 River Bend Court #203
Newport News, VA 23602

Thank you for your time and help.

Sincerely,

John M. Davis

enclosure
The purpose of this survey is to gather information regarding the reporting and use of performance data by public four-year colleges and universities. Please offer your own professional judgment based on your state's experience. We may follow up for clarification or further information.

1. In what way(s) are the public four-year institutions within your state required to report their success in meeting specific performance criteria?

2. In what way(s) is this performance information used?

3. In your opinion, to what degree do these criteria determine in any direct or indirect way the amount of funding appropriated to individual institutions?

4. In your opinion, to what degree do these measures have “buy-in“ among institutional stakeholders?

5. In your opinion, to what degree do these measures attempt to reinforce each individual institution's distinctive mission and strategic plan(s)?

6. In your opinion, to what degree are these measures understandable to the public?

7. In your opinion, to what degree are the measures valid, i.e., they measure what they intend to measure?

8. Do these measures rely on quantitative data (numerical data from various tests, instruments, or formulas) or qualitative data (verbal or written prose)?

9. In your opinion, how well does this set of measures, taken as a whole, assess institutional performance?

10. In your opinion, to what degree do these measures allow for benchmarking and peer comparisons?

Thank you for your help. Please feel free to contact me (tel. 757-249-0944 or e-mail jmdavi@wm.edu) if you have anything to add or any questions about the study.
BIBLIOGRAPHY


Ewell, P.T. and Jones, D. (1994). Pointing the way: Indicators as policy tools in higher education. In S. Ruppert (Ed.), *Charting higher education accountability: A


Layzell, D. (1999). Linking performance to funding outcomes at the state level for public institutions of


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Vita

John Milan Davis

Birthdate: May 24, 1970

Birthplace: Tyndall Air Force Base, Florida

Education: 1994-1996 The Florida State University
            Tallahassee, Florida
            Master of Science

            1990-1992 The Florida State University
            Tallahassee, Florida
            Bachelor of Science

            1988-1990 Gulf Coast Community College
            Panama City, Florida
            Associate of Arts