2006

Perceptions of School Performance Measures: A Study of Principals in the United States and Head Teachers in the United Kingdom Using Q Methodology

Rene Velez
University of North Florida

Perceptions of School Performance Measures: A Study of Principals in the United States and Head Teachers in the United Kingdom Using Q Methodology

By

René Velez

A dissertation submitted to the Department of Leadership, Counseling and Instructional Technology
in partial fulfillment of the requirements for the degree of

Doctor of Education
In Educational Leadership

UNIVERSITY OF NORTH FLORIDA
COLLEGE OF EDUCATION AND HUMAN SERVICES

December 2006

The dissertation of René Velez is approved:

Signature Deleted

Nov. 27, 2006

Katherine M. Kasten, Ph.D., Chair

Signature Deleted

Nov. 27, 2006

Joyce T. Jones, Ed.D.

Signature Deleted

Nov. 27, 2006

Russell O. Mays, Ed.D.

Signature Deleted

Nov 27, 2006

Steven K. Paulson, Ph.D.

Accepting for the Department:

Signature Deleted

January 8, 2007

Joyce T. Jones, Ed.D., Chair
Department of Counseling and Educational Leadership

Accepting for the College:

Signature Deleted

10/30/07

Larry G. Daniel, Ph.D., Dean
College of Education & Human Services

Accepting for the University:

Signature Deleted

30 OCTOBER 2007

David E.W. Fenner, Ph.D.
Dean, The Graduate School

ACKNOWLEDGEMENTS

There are numerous friends, family members, and colleagues whose efforts and support made the completion of this study possible. Although I am constrained in my ability to thank each of them here, this does not diminish my gratitude to each and every one of them for their part. The contribution of the head teachers and principals who found time in their extremely busy schedules to participate cannot be overstated. Their desire to contribute to the educational body of knowledge and the professional development of this doctoral candidate makes them, in my humble opinion, outstanding examples of educational leadership. This superior level of educational leadership was also commonplace for my dissertation committee, especially my chair, Dr. Kathe Kasten, as they provided constant encouragement and ensured I realized the full benefit of this experience. Last, but by no means least, I must acknowledge my wife Mayda's patience and efforts towards realizing this goal.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Performance measures have been used throughout the business sector as a means to assess productivity, allocate resources, and increase profitability. More recently, they have been utilized to answer increasing calls for accountability in public education. Legislation has been passed in both the United Kingdom and the United States that implements performance measures as a means to measure student achievement and assess school performance. This study, conducted both in the United States and the United Kingdom, examined the perceptions of 15 primary and 15 elementary school leaders with regard to the transnational issue of school performance measures.

Q methodology was used to examine the opinions and perceptions of these leaders for the purpose of providing insight for stakeholders and identifying future areas of research. The data from the participants revealed patterns of opinion within the head teacher group, the principal group, and the participants as a whole. Common opinions included the balanced use of performance measures, the political nature of school performance measures, the appropriate use of standardized test scores, and the consideration of economic and social factors. This study also demonstrated the use of Q methodology in qualitative educational research by both obtaining and analyzing rich and insightful participant data.

CHAPTER 1

Introduction and Purpose

A host of socio-economic, behavioral, educational, and policy issues face school leaders in the United Kingdom and United States today. The societal impact of these issues is particularly felt in schools, as they closely mirror their environments. Schools in economically deprived areas are challenged in ways schools from affluent areas are not. Student achievement, attendance, well-being, and ability to learn are affected by the economic and social environment. In many ways, particularly those challenging for a leader, schools are a microcosm of the larger environments they are situated in.

Irrespective of these issues and their impact, the school leaders in these two countries are responsible for both the day to day administration of their schools and efforts to improve existing practices and processes. Included in their many responsibilities are mentoring and training staff, providing a conducive learning environment, and addressing various stakeholder requirements. These requirements are as diverse as the stakeholders themselves. These stakeholders include parents of enrolled children, oversight authorities, and school staff, to

name but a few of the most obvious. Stakeholder requirements, needs, and expectations are considered by these leaders in their leadership roles.

Perhaps one of the most important roles these school leaders have is implementing change and addressing its impact. As change agents, they are expected to implement change that originates both internally and externally. External change for schools often results from issues that have been decided in the public and legislative forums. Change may address any of a host of areas, from the manner in which their schools are operated to the manner in which they are assessed. In some of these public issues, school leaders may only represent a very small voice in the cacophony of discussion on what changes will be effected to address the issue. Nevertheless, they are charged with both implementing the change and dealing with its impact.

One of the issues at the forefront of the public and legislative forum in the United Kingdom and the United States has been the quality of the taxpayer-funded education in schools. In fact, few public service issues claim as much time in the public and media spotlight. The prevalence of this issue is due to its strong personal, public, political, and professional aspects. For many parents, whose children are the recipients of such education, there is hardly a more personal issue with perhaps the exception of healthcare. For the larger public audience, the assurance that a large portion of their taxes is being spent using the best fiscal

practices has become increasingly important as the costs of social services, including public education, have skyrocketed in both countries. One need only consider the 41% real increase in the United Kingdom school budget between 1997 and 2005 (Education Ed., 2005) to appreciate this upward trend of these costs.

As the public's interest goes, so follows political focus. Politicians in both countries have placed this issue at the forefront of their political platforms, declaring goals and passing legislation to address this issue. Calls for schools to adopt proven business and organizational improvement processes are common in both countries as politicians seek to incorporate tangible and measurable indicators of performance. The prevailing sentiment regarding the cost of public education is maximizing value for money rather than past views that largely did not monitor these costs very closely.

Legislation has been passed in both countries to standardize requirements and measurements for schools. In the United Kingdom, the National Curriculum and the nationwide testing that are currently in practice resulted from the passing of the Education Reform Act in 1988 (Black, 1994). In the United States, the No Child Left Behind Act of 2001 mandated that publicly funded elementary schools would be the first schools in which federal nationwide performance measures would be utilized (U.S. Dept. of Ed., 2002). In both instances, this

legislation was focused on measuring the performance of schools against a nationwide standard for the first time. From a chronological standpoint, the use of such measures in United Kingdom state schools has been in practice for almost two decades, whereas the elementary schools of the United States began implementation less than five years ago. In the United Kingdom, performance measures are an integral aspect of monitoring performance of publicly funded state primary schools.

Implementing the changes associated with these school performance measures has taken considerable effort by school leaders and their staffs in both countries. School curriculum was modified and training programs developed for school staff. School leaders have been left with the unenviable task of both defending and improving the performance of their schools. The measures that are in use have been developed externally and are meant to be applied across schools in different environments.

Statement of Purpose

The purpose of this descriptive cross-group study was to examine the perceptions and attitudes of a group of school leaders, both in the United Kingdom and the United States, on the use performance measures in their schools. As leaders of their schools, these head teachers and principals share similar leadership responsibilities for implementing performance measures in

their educational organizations. In order to conduct this examination, this study first determined the attitudes and perspectives of 15 principals and 15 head teachers. These individual perceptions were then examined collectively, both within the two groups and as an aggregate, for patterns of opinion or common perceptions.

The methodology that was used in this study to collect data both on individual perceptions and to reveal patterns of opinion is more often employed in the field of psychology than education. Q methodology provides a means to identify any clustering of like-minded perceptions (Brown, 2004) in the context of small groups such as were examined in this study. A secondary purpose of this study was to validate the use of this methodology for educational research as a means of obtaining a much richer insight of attitudes and perspectives than might result from traditional Likert scale surveys.

Research Questions

In keeping with the previously stated purpose, five research questions were developed to provide a framework for this study. The research questions addressed both individual perceptions and group patterns of opinion. A final research question addressed the validity of Q methodology towards this end. The five questions were as follows:

1. What are the attitudes of these principals in the United States and these head teachers in the United Kingdom with respect to the use of performance measures in their schools?
2. Are there any patterns of opinion on the use of performance measures in their elementary schools among these principals in the United States?
3. Are there any patterns of opinion on the use of performance measures in their primary schools among these head teachers in the United Kingdom?
4. If there are patterns of opinion, are there any differences or similarities when comparing the opinions of the principals in the United States with those of the head teachers in the United Kingdom on this issue?
5. Is Q methodology an effective means to determine individual and common attitudes and perceptions regarding an educational issue?

Definition of Terms

The definitions of terms that were used throughout this study are contained in this section. Among these terms will be those associated with the methodology, the issues, and the organizational structure of the schools in the United Kingdom and the United States.

Accountability, best described in the glossary of *Education Week* (n.d.) for the purposes of this study, is the:

> State or district policies related to holding districts, schools, and/or students responsible for performance. School and district accountability systems typically include efforts to assess and rate schools or districts based on student performance and other indicators, to publicly report on school or district performance, and to provide rewards and sanctions for schools or districts based on performance or improvement over time. (p. 1)

Balanced performance measures are a selection of performance measures that include a proportional use of operational and non-operational based measures.

Department for Education and Skills (DfES). DfES is the United Kingdom counterpart to the United States Department of Education.

Elementary schools in the United States typically provide education to students between the ages of 6 and 11, in levels kindergarten through fifth grade.

Head teachers are the organizational leaders of primary schools in Gloucestershire County, in the United Kingdom. The terms head master and head mistress have been replaced by this term in state schools.

Non-operational performance measures are qualitative measures such as customer satisfaction in business and teacher morale in education.

Operational performance measures are quantitative-based measures such as units of production in business and standardized test scores in education.

Performance measures are measurable aspects of an organization that are used to indicate progress towards organizational improvement goals.

Primary schools in the United Kingdom typically provide education to students between the ages of 5 and 11, in the levels of reception through year five.

Principals are the organizational leaders of elementary schools in Duval County, Florida, in the United States.

Public schools in the United States are funded primarily by taxpayer revenue and administrated by state governments with funding assistance from the federal government and oversight by the U.S. Department of Education. These schools have no enrollment fees.

Q methodology is a qualitative research methodology primarily used in the field of psychology to determine the attitudes and opinions of an individual or small group of participants. The instrument used for this methodology consists of a group of statements that cover the range of opinion on an issue and a forced distribution scale on agreement. Participants complete a sort by placing the opinion statements on the scale while considering their own personal beliefs. The statement order in this sort is then correlated among the participants. A

factor analysis is conducted using an "inverse" data matrix in which persons define data columns and then responses define the rows. The resulting "person factors" are usually rotated to achieve simple structure. Statements are then assigned z-scores with respect to the factors, producing a factor arrays that is defined by a number of sorts. This factor array or combined sort reveals patterns of opinion for the researcher's interpretation.

A *sort* is the rank ordering of a sort set on an agreement scale by the participant, who considers their personal perceptions and attitudes with respect to each statement.

A *sort set*, or *Q-set*, is the group of opinion statements that covers a broad range of opinion on a topic.

State schools in the United Kingdom are no enrollment-cost schools that are funded primarily with taxpayer revenue and with oversight by the Department of Education and Skills.

Significance of the Research

This research is significant in two different respects. First, this research provides information to the stakeholders of both public elementary schools in the United States and state primary schools in the United Kingdom. This information, in its simplest form, is feedback from those in the most important leadership position in these schools. In its most complex form, this information

takes the form of recommendations and guidance for these stakeholders. In the second respect, the field of educational research, this research is significant in that it identifies areas for further research on this complex issue. This study use of an infrequently used research methodology for education also validates a significant qualitative research tool for the educational researcher.

Principals and head teachers, as the leaders of elementary and primary schools respectively, are directly affected by the implementation and consequences of performance measurement. In their unique leadership position on the front line, their perceptions and attitudes regarding the efficacy and impact of these measures can provide insight into a national level education issue in both countries. As noted earlier, the use of national performance measures is more mature in primary schools, where head teachers have been impacted by their use longer than their principal counterparts. In the United States, elementary school principals are at the forefront of public schools on this issue given the initial implementation of the No Child Left Behind Act in grades three through eight (U.S. Dept. of Education, 2002a.). The chronological difference provides an interesting dynamic to consider as these two groups provide opinions tempered by different periods of exposure to nationally implemented school performance measures.

The insight gained from this relatively small group of leaders can be useful to the numerous stakeholders involved with this issue. Although these perceptions are unique to these participants and cannot be attributed to any larger group, this does not detract from their value when considered in the appropriate context. These insights can be considered by policy makers in both countries as they continue to refine the use of performance measures for schools. The perceptions of these participants with respect to the efficacy of current performance measures contain obvious recommendations for more effective implementation. For those education professionals aspiring to these leadership positions, these insights provide advance detail on a complex leadership issue they will likely face. For the parents of children in these schools, this research provides another perspective that can be weighed against the torrent of media and government discussion on the issue. These are but a few of the stakeholders that can utilize the insight gained from this study.

This research is significant to the field of educational research in two ways. First, as is common in most research, this study generated potential areas of further study on this educational issue. The perceptions and attitudes of other stakeholder groups associated with the implementation of these performance measures in one or both of the two countries could be explored and examined for similarities or differences. Various stakeholder groups could be similarly

compared to discover broader patterns of opinion on this issue. These broader patterns of opinion could provide similarly valuable insight to the stakeholders in the execution of their duties.

The successful use of Q methodology in this study provided the other manner in which this study was significant for educational researchers. Q methodology accomplishes data analysis by "the sequential application of three sets of statistical procedures: correlation, factor analysis, and the computation of factor scores."(McKeown & Thomas, 1988, p. 46). This uniquely quantitative approach to qualitative research provides rich data as it allows participants to express their views in a far more complex manner than simple surveys. Furthermore, its versatility does not limit its use to individual participants. Groups that use the same sort can be qualitatively examined for common perceptions or attitudes with regard to educational issues. This commonality or lack of commonality, depending on the results of a study, can be used to both reveal and answer a variety of research questions.

The significance of this research is evident both in its use for United Kingdom and United States stakeholders and the international field of educational research. These head teachers and principals provide insight to stakeholders on a current and intensely debated issue regarding taxpayer-funded primary and elementary schools. Educational researchers, regardless of what

country their research is conducted in, can use this study to assist in their research on this issue. The validation of Q methodology in this study also served to provide researchers with an alternative method to do so.

Summary

This chapter began by introducing the reader to the demanding leadership role our primary and elementary school leaders hold and closed by detailing the significance of this research on a transnational and current issue faced by these school leaders. Legislation passed in the United Kingdom and the United States was briefly covered to provide a sense of the high level of government involvement with regard to improving schools. The attitudes and perceptions a group of these leaders have regarding this oversight, and the mechanisms put in place to accomplish it, are the basis for the first four research questions that were detailed. Terms were defined to allow the reader to understand important aspects of performance measures, Q methodology, and the participants of this study. The applicability of Q methodology as a means to collect data for this study that will answer the research questions was also reviewed.

The literature review that follows will provide further information and context, both in the United Kingdom and the United States, on this current leadership issue. The history and implementation of performance measures both in the business and education sectors will be presented to provide further

context. Current legislation and its impact on schools will be discussed to better understand the data obtained from the participants. The strengths of Q methodology will be presented so as to explain why this qualitative tool was so well suited to answer these research questions.

The description of the methodology that follows will provide the detail of how Q methodology was used to obtain the data to answer the research questions. These data, presented in Chapter 4, were synthesized in factor arrays that were interpreted to reveal patterns of opinion. These patterns, both within the two groups and as one aggregate group, provided the recommendations discussed in Chapter 5, the concluding chapter.

CHAPTER 2

Literature Review

This review of the literature encompasses five areas related to the use of performance measures in schools. The first aspect is an overview of their use, both in international business and the U.K. and U.S. public education sectors. Following this overview, the review then examines the current educational policy in the context of performance measurement required by legislation in the two countries. With the historical and current policy context established, the review examines the most prevalent issues faced by both countries as the result of implementing this legislation. Finally, the leadership impact is examined by focusing on how the role of the school leader is affected by the use of performance measures in their schools.

In the business sector, the historical evolution of performance measures from an initial narrow focus, to the manner in which successful businesses use them today, provides insight as to how they can be similarly applied in non-business organizations. Early educational performance measures used by state governments in the United States consisted primarily of standardized tests and

linked funding to school performance (Massy, 2003). In the United Kingdom, similar efforts to enforce accountability were used to identify failing schools so they could be closed (Bell, 1999). The use of education performance measures in post secondary education in both countries has increasingly included a genuine organizational performance focus. In the United States, the business community has actively endorsed the use of performance measures as a means to improve public education accountability (Hoff, 1999).

The legislation enacted in the United Kingdom and the United States both increased the national government's oversight and reflected government attempts to apply this business approach of performance measurement to schools. The liberal use of performance measure targets by the two governments largely focuses on student achievement and seeks to satisfy taxpayer concerns on value for money. Schools are evaluated as passing or failing based on progress towards these publicly reported targets. Application of this business approach, and the validity of its underlying assumptions, has raised significant issues that are being openly debated in the public and legal domain of both countries.

The issues that have emerged following these government efforts to improve school performance revolve around the end to which these performance measures are being used. The focus on accountability, instead of performance improvement, has raised issues regarding the relative importance of some

performance measures and the ultimate goals of this approach. In the business arena, the focus is more on improving performance in a competitive marketplace than holding the organization accountable. The further focus on one operational measure, standardized test scores, has also raised similar issues in both countries regarding the adverse effect on other non-operationally measured areas.

School leaders in the United Kingdom and the United States face the challenge of improving their schools' performances while also meeting public calls for accountability as reflected in legislation. The implementation of these laws has not changed their basic leadership responsibilities in spite of the far-reaching consequences in a number of areas. Although school leaders continue to be responsible for motivating their staffs, that is made more difficult by an environment in which the teacher assessment of student progress is effectively trumped by an externally mandated standardized test. The issue is further exacerbated by the known limitations of such a broadly administered test in assessing individual student achievement. This degree of external influence has significantly increased the complexity and challenges of the school leader's role.

Performance Measures in Business

Performance metrics have enjoyed widespread use throughout the business sector for over four decades as a means to improve business processes, the quality of production, and market share. Performance measurement is a

complex methodology that has evolved since the business sector first began using it to improve productivity. The initial and crude use of performance measures focused on the singular use of one operational metric, units of production. In time, companies came to realize that having to work with a variety of key performance variables meant that such a singular focus was inadequate for accomplishing real improvement (Harbour, 1997).

The concept of measuring performance is built on the principles used by the American "father" of Total Quality Management (TQM), W. Edwards Deming (Anderson, Cuellar, & Rich, 2003). The use of Statistical Quality Measures, although flourishing for a brief period in the 1930s, did not gain widespread acceptance until Deming's success in post-war Japan (Walton, 1986). It was this advocacy of measuring quality that was a significant departure from traditional business practices. During one of his speeches in Japan in the 1950s, Deming summarized the traditional approach: "Manufacturers used to think of manufacturing in three steps: Design it, Make it, and Try to sell it. These steps were thought of as completely independent" (as cited in Scherkenbach, 1991, p. 9). Among a number of shortfalls with this focus, Deming thought the lack of customer interaction or measuring of the customers' satisfaction was a crucial aspect that was missing.

Deming's statistical approach to measuring areas such as quality and customer satisfaction was eventually embraced by business organizations throughout the world and formed the foundation for holistic business performance measures used today. His fourteen points for management implored American industry to adopt a philosophy of constant improvement with less attention to objectives and numerical goals (Deming, 1995, chap. 2). Deming clearly felt that his fourteen points could be applicable outside of private industry, summarizing the common problem in the following way:

> Efforts and methods for improvement of quality and productivity are in most companies and in most government agencies fragmented, with no overall competent guidance, no integrated system for continual improvement. (p. 465)

This fragmentation and lack of integration when measuring performance indicators can easily negate any of the benefits of using such a system. Those involved in these types of nugatory efforts incorrectly perceive that they are using a viable improvement process for their organizations.

Sustained benefit has not been easily achieved by the business sector in attempting to follow Deming's guidance. The improper use of performance measures resulted in many instances of initial success followed by setbacks. "It's a cliché that you get what you measure, but managing to measures by itself

rarely leads to superior value growth. To succeed, a company needs to manage performance rather than just measure it" (Siesfeld & Pape, 2004, p. 52). Aside from providing an inaccurate assessment of the organization, using metrics improperly can negatively influence members of an organization. A business magazine survey found that "more than a third (37) of the respondents registered dissatisfaction with how metrics are used in their companies to monitor purchasing, sourcing and supply management functions and performance of outside suppliers" (Morgan, 2000, p. 26).

The business sector has increased the benefits of performance measurement by adopting a holistic approach that utilizes more than productivity or operational measures. This holistic approach forsakes the traditional business practice of focusing on one performance metric, usually financial, for a multidimensional view that looks at other areas as well (Frost, 2000). This new view is best exemplified by the Balanced Scorecard model developed by Robert S. Kaplan and David P. Norton, who studied the performance metrics at leading organizations (Frost, 2000). The Balanced Scorecard model developed by Kaplan and Norton (1992) was a performance measurement system that considered not only financial measures, but also customer, business process, and learning measures.

The use of performance measures in the business sector has followed a path that included missteps and the refinement of a balanced approach that has proven successful for many companies. Deming's early guidance on the importance of selecting an integrated approach to performance measurement that did not focus solely on productivity has allowed the business sector to take full advantage of this method for organizational improvement. Holistic and balanced measures have been developed that give equal weight to qualitative aspects of improvement. This approach is the culmination of more than 50 years of trial and error since evolving from the traditional business model that was inadequate for large-scale manufacturing in a competitive worldwide marketplace. Compared to the education sector, performance measurement in the business sector is far more mature.

Performance Measures in Education

There are a number of notable aspects regarding the implementation and progress of education performance measures in the United Kingdom and the United States. The three that will be briefly examined here are the historical beginnings of such measures, their use in post secondary education, and the support of the business community in the United States for their use.

The historical beginnings of public education performance measures in the United Kingdom and the United States focused on accountability and were,

in some cases, linked to funding. This was the case for post secondary education in both countries. Many state college and university performance measures in the United States were linked to funding; Tennessee used such measures as early as 1974, and 34 states were using some type of performance funding by 1997 (Massy, 2003, chap. 10). Public demands for funding accountability had increased, and states were using performance funding as a means to ensure accountability for public colleges and universities (Burke & Modarresi, 2000). In the United Kingdom, institutional research funding was similarly allocated on the "basis of measured performance" (Massey, 2003, p. 290). For K-12 education in the United States and primary education in the United Kingdom, performance measures were also used for accountability, but the focus on student achievement as measured by standardized tests was unique to this level of education. The conservative government of the United Kingdom in the late 1980s crafted legislation for England, Wales, and Northern Ireland that established a national curriculum and national curriculum testing for students at three different ages (Olson, 2004). By 2000, almost all of the state governments in the United States were administering their own standardized tests in K-12 education and publishing results (Elmore, 2002). In both countries, the development of educational performance measures at all levels was driven primarily by the perceived need to hold these publicly funded institutions accountable for results.

The growing use of post secondary school performance measures in both countries appears to increasingly recognize the need to assess organizational improvement in additional to monitoring accountability. Advocates encourage the use of a balanced model found in successful businesses rather than one focused on a narrow measure such as student achievement. This balanced use of performance measures had been proposed for educational institutions attempting to reach customer satisfaction and efficiency goals. Massy proposed a scorecard for colleges and universities that measured outputs, market, internal processes, finance, and organizational learning and growth (Massy, 2003). The use of such business-like measures for post secondary education in the United Kingdom is being clearly directed by the government:

> ...there has been strong Government pressure on the higher education funding agencies and on universities to demonstrate the existence of effective quality measures for teaching, learning, and the student experience, and to publish the results of these measures. The significant influence of this concern reflects the dominance of national Government funding of teaching activity in U.K. higher education and a cross-party political determination in a "customer is king" society to ensure good value for money. (Assn. of Research Libraries, 1999, p. 2)

Another such "customer is king" model, focusing on both student achievement and the customer, has been developed from the health care industry and offered

for use in evaluating the quality of nursing education (Anderson, Cuellar, & Rich, 2003). These instances indicate how post secondary educational performance measures are shifting towards a business sector focus of balanced measures and provide insight as to how early public education can do the same.

In the United States, the business sector has been one of the major proponents of performance measures as a means to ensure accountability. They have advocated the use of performance measures through both government and private organizations. The National Institute of Standards and Technology, a part of the U.S. Commerce Department, administers the Baldrige National Quality Program, which promotes performance excellence among U.S. manufacturers, service companies, educational institutions, and healthcare providers (National Institute of Standards and Technology [NIST], 2005). The U.S. Department of Commerce felt these exacting quality standards could boost performance for educational institutions, so education was added to the possible recipient areas in 1999, with awards for research-based and accountable initiatives (Arif & Smiley, 2003). The fact that the Commerce Department, and not the Department of Education, emphasized this linkage first is worthy of consideration.

The focus of the performance measures advocated by the U.S. Commerce Department is noteworthy. Using its extensive experience with effective business practices, they developed the criteria by which an educational institution or

organization would be assessed for their Baldrige Award. These organizational performance areas parallel many of the business award categories and include others unique to education. The areas are as follows:

(1) student learning results

(2) student- and stakeholder-focused results

(3) budgetary, financial, and market results

(4) faculty and staff results

(5) organizational effectiveness results, including key internal operational performance measures

(6) leadership and social responsibility results. (NIST, 2005, p. 7)

An examination of the areas indicates that operational or productivity metrics (student learning results/standardized exam scores) are only one of several areas these business leaders believe a successful educational organization should focus their efforts on. In fact, the authors of the award criteria make this point explicitly: "The use of this composite of measures is intended to ensure that strategies are balanced—that they do not inappropriately trade off among important stakeholders, objectives, or short- and longer-term goals" (NIST, 2005, p. 7). This guidance emphasizes a multi-faceted and balanced approach to organizational improvement. Conversely, it could be inferred from these criteria that educational institutions focusing a disproportional effort on student

achievement, as measured by standardized test scores, could not fully benefit from the use of performance measurement as an organizational improvement tool.

In the United States, the business sector has continued to increase its role as an external stakeholder of government-funded education to this day. In preparation for congressional hearings on the reauthorization of the NCLB Act scheduled for 2007, these stakeholders are organizing to protect the law from significant changes (Hoff, 2006). Hoff further described this trend of support by the business community for performance measures and accountability in schools:

> While corporate America has long supported national education initiatives, many observers say that business leaders are now more prominent and more focused on specific details than ever before. Although business leaders supported efforts to set national education goals in the late 1980s, for example they weren't as involved as they are now in advocating specific policy measures. (p. 2)

This influence of the business community on educational policy will most likely continue to increase as this external stakeholder seeks to instill more business proven improvement processes in public education.

Current Policy in the United Kingdom and the United States

Although the targets for performance measures are set at different government levels in the United Kingdom and the United States, the current policy in both countries similarly focuses on accountability. In addition to this similar focus, there is a heavy dependence on standardized tests to measure student achievement, and by extension, school performance. Current policy in the United Kingdom is derived from the Education Reform Act passed in 1988 which had accountability as one of its' key features (Bell, 1999). In the United States, efforts that began in the mid 1980's by the National Governors Association to introduce performance-based accountability (Elmore, 2002) culminated in the passing in the No Child Left Behind in 2001 (U.S. Dept. of Education, 2002). Both pieces of legislation established national benchmarks for student achievement and mandated the use of standardized tests as the means to measure it (Olson, 2004; U.S. Dept. of Education, 2002). A review of this legislation and how it has shaped current policy in both countries reveals many similarities with respect to intent and the mechanisms that are employed to meet the requirements of the law.

Before reviewing the provisions of the legislation in the United States, it is worth noting that federal legislation to support funding for K-12 education is relatively recent given the age of U. S. public educational systems. The primary source of federal K-12 support began in 1965 with the enactment of the

Elementary and Secondary Education Act (ESEA) (U. S. Dept. of Education, 2005). Although no substantial changes have been made in the law since its inception, this changed during George W. Bush's first term as president. The No Child Left Behind Act (NCLB) of 2001 was a reauthorization of ESEA, and the law's expressed purpose was to close the student achievement gap through accountability (U. S. Dept. of Education, 2005).

The provisions of the NCLB law include the state governments in the goal of improving the nation's public schools. States are required to assess reading and math every year for every child in grades three to eight (American Federation of Teachers [AFT], 2002). States set standards, in consultation with the federal government, to gauge progress towards the NCLB Act's goal of all students reaching a state-defined level of proficiency by 2014 (National Education Association [NEA], n.d.a).

Although states have been given this opportunity to develop their own tests and assessments (AFT, 2002), the federal government has mandated an independent nationwide benchmark as well. The NCLB tasked the National Assessment of Educational Progress (NAEP) to conduct nationwide mandatory tests in reading and math during the 4th and 8th grades (NEA, n.d.c). The NAEP, in its role as an unbiased congressionally mandated project, had been conducting non-mandatory nationwide student testing of various subject areas since 1969 to

provide student learning assessments (NEA, n.d.b). The state governments' results are essentially verified against the NAEP results. Although other aspects of the NCLB Act are important such as improving teacher quality and the school environment, the clear focus is on accountability measured by achievement.

In the United Kingdom, the use of standardized tests to evaluate student achievement against national standards, and subsequently school performance, is far more pervasive. The distinct difference between the United States and the United Kingdom is that these tests are based on a national curriculum that was also mandated by the legislation in 1988 (Bell, 2004). National tests are given at the completion of each key stage as summarized in Table 1.

Table 1

United Kingdom Student Testing

Year group	Age of pupils at end of year	Key Stage
Reception 1 2	5 6 7	Key Stage 1
3 4 5 6	8 9 10 11	Key Stage 2
7 8 9	12 13 14	Key Stage 3
10 11	15 16	Key Stage 4

(City of Newcastle, 2004)

Exams given when the pupils are eleven and sixteen years old, at the end of key stages 2 and 4, are particularly meaningful as the results are used to rank schools locally and nationally in the "league tables" (DeHavilland Information Services plc. [DIS], 2005). The Qualifications and Curriculum Authority is responsible for maintaining these nationwide tests (DIS, 2005), a far more expanded role than the benchmarking role of the NAEP in the United States.

The monitoring of school performance in the United Kingdom is also accomplished by the use of national-level external inspections carried out by the Office of Standards in Education, or Ofsted, a non-ministerial organization that is accountable to Parliament and inspects everything from child care to colleges (Ofsted, 2006). The following excerpt from the Ofsted strategic plan explains their inspections: "The system of inspection will entail a short and focused review of the fundamentals of a school's performance, closely related to the school's self-evaluation and improvement planning" (Ofsted, 2006, p. 10). The direct manner in which Ofsted monitors school performance is very similar to that of the state governments in the United States.

The national policies of the United Kingdom and the United States with respect to school performance, as implemented by current legislation, share similar themes regarding accountability and student achievement. The legislation in both countries contains provisions for direct intervention in schools

that consistently fail to meet performance targets. By publicly reporting performance measures, both countries effectively "productized" schools like any other consumer product. The legislation in the United Kingdom used testing "to provide the currency for accountability, simple data about schools so that parents could make informed choices" (Black, 1994, p. 194). In the United States the legislation went so far as to direct the states that they must provide transportation for students in failing schools to non-failing schools (Fritzberg, 2004). The current policies in both countries use a testing-based accountability system to monitor the performance of their public elementary and state primary schools. The use of such a system has resulted in a number of common high-profile issues for their school leaders.

Current and Continuing Issues

The use of school performance measures in the United Kingdom and the United States has resulted in a number of issues that are being debated in public, legislative, and judicial forums. This section will focus on key issues that have emerged from two aspects of their use. First, and foremost, there is the disproportionate use of performance measures as a means to ensure accountability, rather than organizational improvement. This use of performance measures is a result of the rising cost of public education in both countries. This cost, funded primarily by taxes, has created a political issue that resonates with

voters and the public at large. Calls for businesslike efficiency and accountability from policy makers are embraced in the United Kingdom and the United States, where sound fiscal practices are an intrinsic part of the societal fabric. The second aspect is the manner in which standardized test scores have emerged as a preeminent school performance measure that is publicly reported as a means to assess accountability. Due to their readily quantifiable nature, standardized test scores have overshadowed teacher assessments and learning in non-testable areas, leading many to believe they are having an adverse effect on the schools and efforts to improve their performance.

Politics and Accountability

The passing of the NCLB law in the United States has mirrored a political trend towards accountability and measurable performance for schools receiving public funds. Public budgeting for schools traditionally focused on inputs, with a view towards desired activities, but has now shifted to results and outcomes (Burke & Modarresi, 2000). This legislation was passed despite the legal difference in the role of the federal government and the states with respect to public education. The U.S. Constitution does not designate a public education role for the federal government, and responsibility for K-12 education falls to the states (U. S. Dept. of Education, 2004). As its funding share increases, the federal government will, in all probability, exercise an even greater oversight role as the

steward of taxpayer funds. Although the 1990-91 federal share of K-12 spending was 5.7 percent, by 2004 it had risen one third since then (U. S. Dept. of Education, 2004). In fact, federal funding for two main federal K-12 programs increased $9.3 billion since 2001 under the president's proposed budget for fiscal year 2005 (U. S. Dept. of Education, 2004).

The new accountability systems for schools, as exemplified by the passing of this legislation in both countries, appear to be based on several key assumptions regarding performance. These assumptions are captured in *Redesigning Accountability Systems for Education* (Fuhrman & Elmore, 2004), a source that also examined whether these assumptions were borne out in practice. The assumptions covered the areas of intent, methodology, consequences, results, and adverse impact. The assumptions were as follows:

- Performance, or student achievement, is the key value or goal of schooling, and constructing accountability around performance focuses attention on it.
- Performance is accurately and authentically measured by the assessment instruments in use.
- Consequences, or stakes, motivate school personnel and students.
- Improved instruction and higher levels of performance will result.
- Unfortunate unintended consequences are minimal. (pp. 8-9)

These assumptions form the cornerstone of an accountability system for schools in both countries that takes a quantitative approach to the delivery of public education. Assumptions with a more qualitative approach, such as schools producing well-rounded and contributing members of society with staff motivated by higher beliefs rather than consequences, would appear to be hard-pressed to find accommodation in this type of system.

The federal government made a substantial investment in standardized tests during the first year after passage of the NCLB law, appropriating $387 million to develop assessments (AFT, 2002). In doing so, the federal government has assumed a share of the financial responsibility for developing the assessments it has mandated. It should be noted that, according to the law, the states must continue to develop assessments should the federal government funding levels for this effort falter (AFT). These potential administrative costs could be problematic for states already struggling to meet educational financial costs.

The use of standardized tests as a school performance measure has also caused significant political controversy in the United Kingdom. Politicians have addressed the public clamor for better schools by promising increased pass rates on the national tests. The following excerpt from an issue brief on testing in schools described the consequences of not meeting these public expectations:

> In 1997, moreover, the Government set a target for 2002 of 80 per cent of 11-year-olds achieving Level 4 or above in the KS2 tests, a step which would result in the then Education Secretary Estelle Morris resigning due to the target's not being met. (DIS, 2005, para. 10)

The severity of these consequences can be placed in perspective for those not familiar with the government structure in the United Kingdom if considered that this would be equivalent to the U.S. Secretary of Education resigning when student proficiency targets of the NCLB Act were not met.

The National Curriculum of England has also stressed accountability in one of its aims, as described in the section entitled "To establish standards": "These standards can be used to set targets for improvement, measure progress towards those targets, and monitor and compare performance between individuals, groups and schools" (National Curriculum On-line, n.d., section 6). The media publicly reported progress on meeting these standards by ranking schools in England and Wales by their standardized test scores. These reports include detailed national newspaper inserts where schools are ranked according to their results throughout the country ("Schools Report," 2005).

The Focus on Standardized Tests

The external focus on school standardized test results, both in the United Kingdom and the United States, has caused considerable issues for schools. The

predominance of this one performance measure, and its heavy correlation to accountability instead of performance improvement, has created an environment where efforts are channeled towards this single area. Although this singular focus is not specifically advocated by either of the two governments, government messages are mixed regarding the relative importance of different performance measures. These mixed messages have resulted in an adverse impact on non-testable areas of learning as efforts and limited classroom time are targeted towards raising test scores. There has also been a negative impact on teacher morale, as teachers' traditional role of assessing student performance appears to be seconded to these tests. Perhaps the most adverse impact will be the inability of schools to develop meaningful performance improvement plans that focus on a variety of performance measures as long as the singular focus of standardized tests remains.

In the United States, this predominance of national test scores as a performance measure appears to be inconsistent with stated government goals. Only one of the six goals, as delineated in the 2002-2007 Department of Education strategic plan, is aimed at improving public education as measured by student achievement (U. S. Dept. of Education, 2002). The goals also focus on other areas:

- Create a Culture of Achievement

- Improve Student Achievement
- Develop Safe Schools and Strong Character
- Transform Education into an Evidence-Based Field
- Enhance the Quality of and Access to Postsecondary and Adult Education
- Goal Six: Establish Management Excellence (p.3)

Balanced performance measure developed for these goals could provide meaningful information regarding the accomplishment of this strategic plan. These goals, and the corresponding strategic focus of the Department of Education, do not appear to elevate test scores inappropriately.

The Department of Education appeared to send a different message with regard to the purpose and challenges of the No Child Left Behind Act of 2001, enacted just prior to the release of its strategic plan. In the information flyer found on its web site, entitled "Facts About...Measuring Progress" (U. S. Dept. of Education, 2003), there is a heavy emphasis on testing. It states,

> Testing tells parents, communities, educators and school boards which schools are doing well. If a school takes a challenging population and achieves great results, testing will show that. If a school is allowing certain groups to fall behind year after year, testing will expose that, too. (p. 1)

One is only left to wonder how school efforts towards other goals in the strategic plan will be assessed as testing progress in these non-operational areas is not

practical. One could conceive of a school that has increasing test scores but is not "doing well" due to student safety issues. A lack of progress in this area would possibly not be evident given the focus on testing measures.

The message from the government in the United Kingdom could be considered more direct with regard to the use of testing to improve schools. In an effort to evaluate and ensure accountability, publicly funded schools have been directed to meet performance goals primarily focused on student achievement as measured by standardized tests (National Curriculum On-line, n.d.). The strategy espoused by the United Kingdom counterpart to the U.S. Department of Education with respect to the use of testing is similar in its goals. In *Excellence and Enjoyment: A Strategy for Primary Schools*, promulgated by the Department of Education and Skills, the Department is very clear on the appropriate use of such tests: "use tests, targets and tables to help every child develop their potential and measure school performance" (2003, executive summary). Although performance targets in areas other than testing exist, the predominance of standardized test scores is clear.

Educational stakeholders in both countries have become concerned about the equivocal role of national standardized exams and the link to school accountability. In the United States, Monty Neill, executive director of the National Center for Fair and Open Testing (FairTest), expressed concerned about

how states will assess their schools under the NCLB Act: "Although building an accountability system based on classroom assessments makes more educational sense, most states will find it easier and less expensive to rely on standardized tests to meet the law's requirement" (2003, p. 1). The NEA has warned that the assessment of whether or not a school is performing adequately had increasingly relied on standardized test scores even before the implementation of the NCLB Act and cautioned this should only be one aspect of accountability (NEA, n.d.a). The NEA went further and proposed a balanced set of measures:

- For teachers, evaluations are a more rigorous and thorough accountability system than standardized test scores.
- For students, assessment also should take into account classroom assignments, grades, scores on teacher-developed tests and other performance measures.
- For schools, assessments should take into account graduation rates, progress on standardized tests (as opposed to just raw test scores) and other measures. (NEA, n.d.a, p. 2)

It is significant that these two stakeholders hold similar views regarding the singular focus on standardized test scores as a performance measure. Both warned against determining school performance or improvement by disproportionately weighing one such measure of school performance.

Teachers in the United Kingdom and the United States, as stakeholders very close to this issue, have risen what could be the loudest alarm. In *Where We Stand: Standards-Based Accountability and Assessment*, the American Federation of Teachers (AFT) voiced its concerns about the use of standardized tests: "The public and teachers are understandably deeply troubled that standardized tests are all too often being used inappropriately, are usurping too much instructional time, and are crowding out recognition of other important subject areas" (2003, p. 3). The AFT shared similar concerns with FairTest when members expressed their view on how states use standardized test results, commenting that "many states and local districts grossly misuse test results when they make high-stakes decisions affecting students, schools or school staff based on testing and accountability systems that do not meet professional standards" (p. 4). Teachers in the United Kingdom voiced their displeasure with the singular focus on nationwide exams by refusing as part of a union action to administer them in 1993 (Black, 1994) and almost succeeding in a similar boycott as late as 2004 (DIS, 2005).

This focus on one performance measure has raised concerns among these teachers that other learning activities are being impacted adversely. Research conducted by the National Union of Teachers (NUT) in the United Kingdom, the equivalent of the American Federation of Teachers in the United States, found

that almost five hours of classroom time was being spent each week preparing for national tests (DIS, 2005). The teachers' view of these targets based on national exams was clearly articulated by the general secretary of the union in 2003: "The Government's obsession with target setting and performance tables has been most damaging in education. Schools have been forced to jump to impossible national targets and to put on the back burner much that is valuable for children's learning" (National Union of Teachers, 2003, para. 3). A study in the United States appeared to confirm these fears. The study, conducted by the Center on Education Policy on the fourth anniversary of the NCLB, found that:

> ...71 percent of school districts reported that they had decreased the time teachers spent on subjects not specified for testing under the federal law so they could emphasize reading and math. In some cases, districts said they skipped certain subjects altogether to provide students with double reading or math time...(Davis, 2006, p. 1)

This singular focus on one performance measure, and its impact on other learning, is of great concern to teaching professionals in both countries.

This emphasis could be compared to the initial over-reliance on financial measures exhibited by members of the business sector before they found a more balanced approach to be effective. Even with this realization throughout the business sector, businesses still remain vulnerable to the adverse effects of

focusing on one performance measure. The banking industry recently learned that using performance measures as a means to assess employee success can and has led to a focus on the measure instead of the overall business (Hill, 2000). The educational sector, with less experience in this area, is especially vulnerable to these same counterproductive forces if performance measurement is not used properly. The high visibility of one performance measure, such as standardized test scores, may focus teaching and other staff efforts exclusively on raising those scores. In addition to proving detrimental to a balanced approach that may improve the organization, other adverse impacts might occur. Neglecting non-tested areas of learning is just one such problem.

School Leaders and Performance Measures

Primary and elementary school leaders in these two countries have had to consider a number of leadership issues associated with performance measures given their planned long-term use by the government. The educational leaders in the United States are under no illusion regarding their longevity as indicated by a recent survey of school leadership which revealed "almost 9 in 10 superintendents and principals (87% and 85%) believe that the push for standards, testing and accountability in their state is here to stay" (Farkas, Johnson, & Duffett, 2003, p. 20). In the United Kingdom, almost two decades of continuous use of such measures are a testament to their longevity there. One

researcher, following her study of successful head teachers, described what this meant to the profession: "Modern headship means leading a highly accountable, closely scrutinized public service" (Woods, 2002, p. 16). The obvious permanence of these measures does not obscure the larger issue in both countries. It is the use of these performance measures in the shadow of accountability that presents the most formidable challenge for these leaders.

This challenge has a distinctly personal aspect for these leaders. It has become evident that progress on these performance measures can be considered by supervisors when evaluating subordinates' personal performance. In the United States, there appears to already be a strong correlation between the two, as more than half of the superintendents who participated in a 2003 survey used test scores as a means to evaluate principal performance and more than four in ten said they were "much more likely" to remove or reassign a principal when student achievement was low in their schools (Farkas et al.). Linking the performance of these principals and head teachers so closely to operational measures such as standardized tests scores will undoubtedly have an effect on how they perceive performance measures with respect to their leadership role.

In their school leadership position, these head teachers and principals are charged with gaining acceptance by staff and parents for these measures at the point of implementation. Some insight as to the complexity of this task can be

gained from one of the first business consultants who wrote of leadership challenges associated with change in an organization: "Managing effective transitions does not allow for dealing with a single reality; it involves managing multiple realities as seen through various people's fears, hopes, and aspirations--their frames of reference" (Connor, 1992, p. 101). The positional authority these leaders have will not be enough to effect change, as one London head teacher learned: "I thought the status of being a head would reduce other people's resistance to change, but this is not the case. You still have to convince people and take them with you."(Meyers, p. 4) These school leaders will have to address the concerns of different stakeholders if they are to be successful in implementing the change associated with performance measures.

Parents can be one of the most demanding stakeholders that head teachers and principals must consider with the implementation of performance measures such as standardized test scores. Head teachers are faced with parental concerns regarding these tests, such as the "excessive parental demands for examination success" (2001, p. 5), identified by Englefield in his research of the leadership challenges of primary schools. For principals, there is also the issue of parental acceptance of the No Child Left Behind Act, as a study conducted three years after its implementation revealed:

> Those who do know enough about NCLB to have an opinion are evenly divided between those who feel favorable (39%) and those who feel unfavorable (38%) toward it. Although positive and negative feelings are nearly even, those who feel negative express a greater intensity of feeling (23% very unfavorable, versus 16% very favorable).... (Hart & Teeter, 2004, p. 2)

Even presuming in the two years since this study that the number of adults who feel favorable has increased, the likelihood is that principals may still find themselves in a position where they will have to "sell" the performance measures the NCLB legislation has placed on their schools to a significant percentage of their student's parents. Parents in the United Kingdom and the United States will have concerns that these school leaders will have to address in their leadership roles.

The use of performance measures in schools has meant head teachers and principals will face daunting challenges in their leadership role. With the accountability overtones of these measures and the link to their own personal performance, it will be a major aspect of their position. A recent study of the perceptions of 45 North Carolina principals reported that the state's accountability and testing system affected their leadership role:

> ...ABCs program had the most influence on monitoring student achievement, aligning the curriculum to the testing, providing remedial and/or tutorial opportunities, assigning teachers to grade levels or subjects, and protecting instructional time. In contrast, the instructional leadership practices that the principals believed were least influenced included dealing with student, teacher, and parent stress, evaluating teachers, and obtaining needed resources. (Lyons & Algozzine, 2006, p. 11)

The impact of testing-based accountability on the head teacher leadership role can be similar. One study that examined why head teachers left their positions early found that for some "...there was a concern for the societal change into what was seen as an alien accountability culture, particular in its link to performance management...." (Flintham, 2003, p. 6). Principals and head teachers alike will have to lead their teachers and their other staff through the change these measures engender while simultaneously providing information and assistance to parents so that they may place them in perspective and understand the impact on their children.

Summary of Literature Review

In almost every sense, current United Kingdom and United States national performance measures for public education are at the beginning of a journey the

business sector began over 50 years ago. This is true even in the United Kingdom, where their use is approaching two decades. This immaturity is reflected by the inordinate reliance on one operational performance measure, much in the same way the early use of business performance measures focused on easily quantifiable productivity measures. Post secondary educational institutions are shifting to a balanced use of performance measures and provide a model for other public education organizations.

The predominance of accountability in the legislation passed in both countries, instead of organizational performance improvement, has served to encourage this narrow focus. There are clear similarities between the two education laws that mandate a test-based accountability system for improving student achievement in schools. Clearly, the use of performance measures for evaluation and accountability purposes in both countries will, in all likelihood, continue given present political trends.

The focus on a narrow operational performance measure such as standardized test scores has created a host of common organizational and leadership issues for school leaders in both countries. The political accountability aspect, coupled with the focus on standardized tests, creates issues for head teachers and principals in their role as a staff leader and parent liaison. The effect of school performance measures on both their daily and strategic

planning is significant and will challenge them with an intensity few other issues have. The perceptions of these leaders regarding the use of performance measures in their schools can provide valuable insight for a number of internal and external stakeholders. Chapter 3 discusses the methodology that will be used to reveal their perceptions.

CHAPTER 3

Methodology

The methodology for this study had to provide data that would answer research questions that were centered on the perceptions and opinions these school leaders in the United States and the United Kingdom had regarding the use of performance measures in their schools. This qualitative study utilized Q methodology as a means to collect and analyze data on these perceptions and opinions. Q methodology is an appropriate methodology as it is able to "reveal subjective structures, attitudes, and perspectives from the person or persons being observed" (Brown, 1996, p. 564). Invented in 1935 by British physicist-psychologist William Stephenson (Brown, 1996), it has enjoyed widespread use in the field of psychology and "is most often associated with quantitative analysis due to its involvement with factor analysis" (Brown, 1996, p. 561). This quantitative aspect adds a unique rigor to this qualitative methodology.

This methodology is well suited to collect data on the perceptions and attitudes of the two participant groups in this study as it provides a means to identify any clustering of like-minded perceptions (Brown, 2004). Similarly, it

can be used to identify any clustering, or patterns of opinion, when the two groups are combined. Q methodology is a qualitative research tool that can provide answers to the research questions as it is often used for "defining participant viewpoints and perceptions" (Brown, 2004, p. 1). This methodology, using its factor analysis component, can extract factors that represent dimensions relevant to the research questions.

Design

The sort set for this study was compiled utilizing predominant themes that have emerged from a review of the literature on this issue, an accepted source from which a sort set can be elicited (Watts & Stenner, 2005). This unstructured sampling technique captured the larger issues associated with performance measures in schools, making the statements in this Q-set "broadly representative of the opinion domain" (Watts & Stenner, 2005, p. 75). Themes included the predominance of standardized test scores in school performance measures; the intended use of school performance measures by the oversight authority; parent view and use of school performance measures; the usefulness of performance measures to school leaders; the administrative impact of performance measures on schools; the use of performance measures in the non-business sector; and the use of a balanced set of performance measures when

implementing performance measurement. A list of the 62 statements in the Q-set is provided in Appendix A.

Pilot Study

The research methodology for this study included a pilot of the instrument both in the United States and United Kingdom. The pilot in the United States was conducted using students in the Educational Leadership doctoral cohort at the University of North Florida in Jacksonville. Many of these cohort members are in the public education profession, with several serving in the Duval County school system as teachers and principals.

The pilot in the United Kingdom included one head teacher from a private or independent primary school. Although a private school, the head teacher's school was in the sample county and voluntarily uses the same performance measures as state schools. The purpose of both pilots was to validate aspects of the research instrument prior to its use. Some of these aspects included time and ease associated with completing the Q-sort, the clarity of the Q-set statements, and the unbiased nature of the prompt.

The instrument pilot in both locations yielded the required feedback. The time to complete the sort, in both locations, was validated as being approximately one hour. One of the most important areas of feedback from the pilot would be regarding the 62 statements. Minor grammatical changes were

suggested, but the content and intended nuances of all the statements were understood by participants in both pilots. There was no feedback on the demographic questionnaire from the U.S. pilot of the instrument. In the U.K., it was suggested that in-service training be added under types of training as this is a commonly used term to denote professional training. The demographic questionnaire was updated accordingly prior to beginning data collection in the U.K. Participants in both pilots considered the prompt to be unbiased, a crucial result for validating the instrument.

Participants and Confidentiality

The 30 participants were equally divided between the United States and the United Kingdom. They were serving principals and head teachers of their elementary and primary schools. The principals in the United States were from the Duval County public school system in Florida, which includes 104 elementary schools. The head teachers in the United Kingdom were from the Gloucestershire County state school system, which includes 231 primary schools. The 15 participants in Duval County represented 14% of the total principal population assigned to elementary schools. The 15 participants in Gloucestershire County represented 6% of the head teachers population assigned to primary schools. It should be noted that these percentages of the larger principal and head teacher population in the two counties provided are for

contextual purposes only. These participants were not a representative sample, and Q methodology findings cannot be extrapolated to the larger population.

Somewhat different methods were used to select the participants. The populations of interest were Gloucestershire County head teachers in the United Kingdom and Duval County (Florida) principals in the United States as the groups that would complete the sort. These groups would be "representative or informative about the topic of interest" (McMillan & Schumacher, 1997, p. 171) when seeking to answer the research questions. The selection of sample participants within these population groups was accomplished differently in the United Kingdom and the United States, though the primary basis for selection in both countries was accessibility (McMillan & Schumacher). Snowball sampling through referrals was used to contact head teachers, a process which differed from the sampling used by the consultant, who relied on professional networking.

The study was designed to protect the confidentiality of all participants regarding their identities and the names of their schools. The assurance that no identifying information would be published was clearly stated in the Human Research Consent Form. These assurances were made to ensure participants would convey their candid perceptions during both the sort and interviews. This study sought to report all perceptions, regardless of their congruence with

official positions or guidelines regarding this topic, and these assurances might have been a factor in obtaining candid responses from participants. It should be noted that none of the participants displayed any concern regarding how identifying information might be reported in the study.

The design for this research study was submitted to the Institutional Review Board at the University of North Florida for review in June 2005 and approved the same month. The approval, including the certification of the principal investigator, is provided in Appendix B. The Human Research Consent Form, completed by each of the participants during the study, is provided in Appendix C.

Procedure

The method of data collection included the use of a proctor in both countries. This enabled informal interviews throughout the course of the data collection. The proctors were also available to provide assistance on both the methodology and the prompt to the participants. Initial contact with potential participants was made via telephone. A description of the research, Q methodology, and the instrument were provided during this initial contact. Those who chose to participate then met with a proctor for approximately one hour, during which the data collection occurred. The research instrument was administered to each participant individually.

Participants were first required to read the informed consent form and document their agreement to participate in the study. They were then provided with the demographic data collection sheet, written directions, the prompt, and a list of the statements. The demographic data questionnaire, Appendix D, requested the participants to provide demographic data on their location (either U. S. or U. K.), years in present position (from 1 to 7+ years), and school student enrollment. The questionnaire also contained a section querying what type of training, if any, participants had received on performance measurement. The written directions, explaining each step of the data collection effort, were then reviewed by the proctor with each participant.

The reading of the "prompt" was the first step of the Q-sort. The prompt served to provide the participant with an overview of the issue in advance of the Q-set so that the opinions postulated by the statements could be understood. This overview was designed to be neutral on the issue so as not to introduce any bias that could later be reflected in the participant's sort. The prompt only provides the context of the broader issue so that the participant can objectively consider agreement or disagreement with the sort statements as they are placed in the forced distribution. The prompt used by the participants is provided in Appendix E.

The participants were then provided with a list of the statements and asked to review them for clarity and any ambiguity. Following this review, and the clarification of any statements by the proctor, participants were readied to begin the Q-sort of the statements. The statements were provided to participants on three-and-a-half by two-inch cards, with one statement per card. The scale values were provided on similarly sized cards, arranged as the top row of the distribution, with the required statement distribution for that value in brackets under the number.

The participants were asked to accomplish the Q-sort of the 62 statements using a graduated scale that created seven groups of statements between the anchors of *most agree* and *most disagree*. The survey format forced the required distribution of the statements as depicted in Figure 1.

Most Disagree Neutral Most Agree

-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6
S	S	S	S	S	S	S	S	S	S	S	S	S
S	S	S	S	S	S	S	S	S	S	S	S	S
	S	S	S	S	S	S	S	S	S	S	S	
		S	S	S	S	S	S	S	S	S		
			S	S	S	S	S	S	S			
				S	S	S	S	S				
					S	S	S					
						S						

Where S= Statement

Figure 1. Q-sort scale and statement distribution.

The Q-sorts of the statements are a forced distribution on a scale of (+6) for Strongly Agree to (-6) for Strongly Disagree, where 0 is neutral. The statements were grouped along this response scale by the participants after they reviewed the statements and the prompt. This response scale avoids yes/no answers, or limited response scales, so as to not preclude factor analysis.

During the Q-sort, the participant was asked to consider the statements against their own personal views and rank them accordingly. As an example, consider two of the statements that express almost directly opposing opinions: "A balanced set of performance measures must be used for schools" and "Standardized test scores are the only necessary school performance measure." During the Q-sort, the participants rank-ordered these statements on the response scale according to their own beliefs. The former could be ranked to the extreme right as one of the two under Strongly Agree (+6). The latter might be ranked as one of two in the Strongly Disagree (-6) group. In both cases, only one more statement could then be rank-ordered similarly under those values, as only two statements are permitted in the distribution. The participants were permitted to re-order statements as often as they liked during the sort, as long as they maintained the required distribution.

Before, during, and after the Q-sort, the proctors collected any feedback or comments the participant provided regarding the use of performance measures

in schools and Q methodology. In the United Kingdom, these comments were recorded in a transcript by the proctor. In the United States, the principals were able to provide written comments in a section of the form used to document their sort distributions.

Data Analysis and Interpretation

In order to answer the research questions, the data were analyzed in two separate stages. The first stage was factorial analysis of sorts completed by the participants. In order to accomplish this analysis, this researcher utilized the PQMethod software to evaluate the completed Q-sorts. PQMethod is available free of charge to researchers from Kent State University and can be downloaded from the Internet. A number of commercial products are available which contain simpler interfaces but offer essentially the same functionality. The following excerpt from the PQMethod Manual describes the software's capabilities:

> PQMethod is a statistical program tailored to the requirements of Q studies, allowing easy data entry (Q-Sorts) the way they are collected, i.e. as 'piles' of statement numbers. It computes inter-correlations among Q-Sorts, which are then factor-analyzed with either the Centroid or Principal Component method. Resulting factors can be rotated either analytically (Varimax), or judgmentally with the help of two-dimensional plots. Finally, after selecting the relevant factors and 'flagging' the entries that

> define the factors, the analysis step produces an extensive report with a variety of tables on factor loadings, statement factor scores, discriminating statements for each of the factors as well as consensus statements across factors, etc. (Schmolck, 2002, p. 1)

The Principal Component method and Varimax rotation were used for the first stage of the data analysis in this study.

The sorts from each participant group were entered in PQMethod to determine any patterns of opinion within each of the respective groups. A comparison of the significant factors within each group was undertaken to determine any different perceptions within the two groups. In order to determine any similar patterns of opinion between the head teachers and the principals, the sorts from both groups were combined into one data set and entered in PQMethod. Thus, the research design produced three data sets for entry in PQMethod. These data sets were the head teacher Q-sorts, the principal Q-sorts, and an aggregate of the two.

PQMethod first produced a correlation matrix for each of the data sets. The Principal Components analysis extracted factors from each of the correlation matrices. When the factors were extracted, the VARIMAX capability of PQMethod was used to rotate the factors and maximize the number of sorts that defined each factor. VARIMAX produces uncorrelated or orthogonal factors, and

these factors indicated constructs that addressed the research questions. Following rotation, z-scores were generated for all the statements in relation to the corresponding factor. PQMethod then used QAnalyze to convert these z-scores into corresponding values in the Q-sort distribution for each of the factors.

The second stage of data analysis involved the subjective interpretation of the factor arrays. These factor arrays, or combined participant sorts, consisted of Q-set distributions for each group of participants that defined a factor. These combined sorts are a collective set of perceptions as described by the Q-set statements. As such, their interpretation allowed an understanding of the common perceptions of those participants who defined the factor. Considering the statements and their locations on the distribution allowed the emergence of themes regarding the use of performance measures in schools. When evaluating the statements, initial consideration was given to the nine statements at both the extreme right and left of the distribution. These are the statements under the positive and negative values of 6, 5, and 4. These statements provided information on the statements the group both strongly disagrees and agrees with. The statements at the center of the distribution, under -1, 0, and +1 provided insight as to what opinions the group remains neutral on.

Delimitations and Limitations of the Design

The design of this study was delimited by the focus on the perceptions of a particular population, head teachers and principals, in their school leadership roles with respect to one issue. These educators are facing the implementation of an organizational improvement process that was previously limited to the business sector. This use of performance measures, with heavy accountability overtones, has resulted in significant issues for these leaders whose early responsibilities revolved around facilities maintenance and internal academic issues such as curriculum (Catano & Stronge, 2006). The attitudes and opinions of these leaders with respect to the use of performance measures in their schools is the basis for the research questions of this study.

The study was further delimited by the selection of 15 principals in Duval County, Florida in the United States and 15 head teachers in Gloucestershire County in the United Kingdom as participants. The results obtained in this study can be attributed to these 30 participants only and it is possible that head teachers or principals that did not participate, whether in these locations or not, may have significantly different attitudes and opinions. The selection of the participants, and the inability to generalize the results of this study to larger populations, is not a limitation given the accepted attributes of Q methodology.

A limitation of the design was the statements in the Q-set derived from the literature review which are expected to represent the broad range of opinion on the issue. Some aspect of the issue may not have been sufficiently addressed by the Q-set and therefore could not have been revealed. Another limitation of the design was the candor that was expected of the participants. If the participants thought stating opinions contrary to the official view could result in retribution, the data collection effort would be adversely affected. This limitation was mitigated by assuring the participants their identities would not be revealed in the results of the study. Perhaps the most significant limitation of this design is the comparison of these two leadership groups that have a variety of cultural, statutory, and historical differences that frame this or any common issue they may face.

Summary

This two-country research study was designed to subjectively examine the attitudes and perceptions of school leaders who are at the point of implementation of performance measurement. These attitudes and perceptions will add to the body of knowledge on this current and transnational issue. The delimitations and the limitations of this study were acknowledged and given due consideration. The target number of study participants, 15 in both the U. K. and U. S., was achieved without difficulty due to outstanding cooperation from these

leaders. All participants were drawn from the target counties, Duval County, Florida in the United States, and Gloucestershire County in the United Kingdom. The participants were from a diverse sample of public and state schools which exhibited varying enrollment and economic conditions. Locations within the inner-city and in suburbia provided a rich contrast of environmental settings.

The Q-sorts were completed by the participants during the summer and fall of 2006 and were administered in the same manner both in the United States and United Kingdom. Demographic data were collected from all participants as was feedback regarding the issue. These data were examined by the researcher, and a factorial analysis of the participant sorts was accomplished using PQMethod. The results of these data analysis efforts are reported in Chapter 4.

CHAPTER 4

Presentation of Findings

This presentation of findings presents the data from each of the groups in the study and a comparison of the data between various sub-groups following analysis of the data. The first data presented is that of the head teachers in the United Kingdom. The analysis of the sort data, coupled with demographic data, was interpreted to reveal patterns of opinion as reflected by the factor arrays of significant factors. These findings are followed by the results of the elementary school principal data collection in the United States. The presentation of findings continues by examining the data when the two groups are combined into an aggregate group and the factors similarly interpreted. Comparisons within and between the groups were conducted by examining the participant sub-groups that defined the significant factors.

Head Teachers in Gloucestershire Primary Schools

In the United Kingdom, this researcher initially identified the participant sample with three referrals of state primary school head teachers in Gloucestershire County provided by the education and training liaison of a

government agency located in the county. These three initial referrals provided at least one, and as many as three, additional referrals. These referrals provided other potential participants and several offered to provide additional referrals if needed. All referrals were initially contacted via telephone to assess their desire to participate. Of all the referrals that were contacted, only one declined to participate in the study, citing a full calendar as the reason. All participants in the U. K. completed the sort and interview during working hours at their respective schools. Time to complete the sorts and interview varied between approximately 60 and 90 minutes.

Demographic data were obtained from all participants in the United Kingdom. The average enrollment of the state primary schools was 356, ranging from 420 for the largest to 90 for the smallest. Ten of the 15 head teachers had been in position over seven years, with the newest having been in position for two years. All head teachers, with the exception of one, had received some type of training in performance measurement. Of those 14, 10 considered the training to be adequate. Although the types of training received varied, all fourteen of the head teachers reported receiving in-service training or a workshop in performance measures. Computer-based training was reported by 3 of the 14; all but 3 had completed some type of self-study.

The data from the 15 sorts were labeled with a unique numeric identifier and the letters UK to denote their country and entered in PQMethod. Entering the sorts completed by the 15 head teachers resulted in the correlation matrix provided in Table 2.

Table 2

Correlation Matrix of Head Teacher Sorts

SORTS	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
15	100														
14	35	100													
13	28	54	100												
12	41	26	34	100											
11	21	51	16	18	100										
10	45	54	45	40	35	100									
9	31	22	28	50	25	39	100								
8	48	61	49	39	47	50	33	100							
7	37	58	53	26	39	51	24	57	100						
6	41	52	48	24	38	55	14	54	52	100					
5	18	51	57	12	22	39	9	38	53	47	100				
4	36	54	51	41	37	56	24	56	52	42	34	100			
3	34	49	44	41	20	43	22	40	47	47	40	43	100		
2	34	59	63	33	33	54	19	69	62	54	63	50	42	100	
1	25	37	41	24	25	47	16	37	58	37	52	38	16	48	100

Note: Decimals to two places omitted

A number of substantial correlations among the head teacher sorts were noteworthy. The most substantial correlation was that between sort 2 and sort 8. Sort 14 substantially correlated to the most sorts in the group; sorts 13, 11, 10, 8, 7, 6, 5, 4, and 2. For sort 11, this was the only substantial correlation. Sort 12 also substantially correlated to only one sort, number 9. Two sorts, numbers 15 and 3, did not substantially correlate to any of the other head teacher sorts.

A principal component analysis resulted in two significant factors, as determined by both their pre-rotational Eigen values (6.8092 and 1.4670) and their position on a Scree plot. These two factors had been rotated using VARIMAX in order to maximize the clustering of participant sorts around each. PQMethod determined z-scores for all the statements in relation to the two factors and produced a factor array, or combined sort, for each factor.

Factor A: Inadequacies of Current Measures

The sorts completed by the head teachers revealed two distinct collective perceptions as supported by the factors. The first factor, Factor A: "Inadequacies of Current Measures," was defined by 12 of the 15 participants' sorts. The statement z-scores for this factor generated the factor array, or combined sort, in Figure 2:

Most Disagree						Neutral					Most Agree	
-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6
2	15	3	8	17	4	10	7	18	5	1	5	25
20	16	46	11	19	14	22	21	29	28	9	34	36
	31	52	13	35	26	45	23	33	44	12	40	
		58	42	37	27	48	24	39	53	62		
			57	38	30	49	32	41	56			
				54	51	50	43	47				
					61	55	59					
						60						

Figure 2. Head teacher Factor A array.

This larger group of participants strongly disagreed with the use of standardized test scores as the only measure to evaluate schools (statement 2), and that school performance measures should be uniform regardless of economic and social factors (statement 31). These perceptions were also shared by the sub-group defining the second factor. A distinct perception of the Factor A sub-group was the strong disagreement that current performance measures provide a comprehensive view of their schools (statement 20). They also disagreed that their effectiveness, as perceived by their supervisor, was based largely on standardized test scores (statement 15) and perceived that performance measures had not had a positive impact on teacher morale (statement 16).

The Factor A sub-group strongly agreed that non-school-controlled social factors directly affect school performance measures (statement 25). They similarly agreed that using school performance measures to rank schools ignores other important factors (statement 36). The group further agreed that using standardized test scores as a performance measure caused teaching to the test (statement 34), and that the use of performance measures to evaluate schools has become a politically charged issue (statement 40).

Factor B: Use Balanced Measures and No Ranking

The remaining 3 participants defined Factor B, "Use Balanced Measures and No Ranking," that resulted in the combined sort in Figure 3:

Most Disagree						Neutral					Most Agree	
-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6
2	3	50	13	11	15	10	16	4	7	5	9	1
28	8	51	22	27	23	12	18	17	19	6	34	21
	31	52	30	29	37	14	24	26	20	38	62	
		55	42	43	45	33	25	35	32	39		
			48	46	53	44	40	36	56			
				47	54	60	49	41				
					57	61	59					
						58						

Figure 3. Head teacher Factor B array.

The Factor B array, defined by only 3 of the 15 participants, contained shared perceptions that were different from the larger sub-group of head teachers. Although this smaller group similarly strongly disagreed with the sole use of standardized test scores as a school performance measure when evaluating schools (statement 2), they also strongly disagreed with the use of performance measures to rank schools (statement 28). This attitude was further enforced by the belief that practice of ranking schools was disruptive (statement 39) and did not serve the public interest (statement 8). They also disagreed, as did the larger group, with elected officials setting school performance measure targets (statement 3).

This sub-group strongly agreed that teachers should play a vital role in developing performance measures (statement 62) and that the effectiveness of any measures should be objectively reviewed on a periodic basis (statement 21). This group perceives that performance measures, when developed properly and

periodically reviewed, could be used to assist school leaders in improving their schools (statement 5), a belief shared by the larger group. They also perceived that standardized test scores have overshadowed other school performance measures (statement 6) and agreed with the larger group that the predominance of this measure causes teaching to the test (statement 34).

The group of head teachers, as a whole, thought that use of performance measures in their schools had become a political issue and that important social factors must be considered in their implementation. Although the larger sub-group defining Factor A thought that performance measures were not used to evaluate their effectiveness, they perceived the measures were having an adverse effect on teacher morale. They also opined that current performance measures did not provide a comprehensive view of their schools. The smaller sub-group that defined Factor B specifically rejected the ranking of schools using performance measures while also offering that performance measures developed and reviewed properly could assist school leaders in improving their schools.

Principals in Duval Elementary Schools

Since this researcher resides in the United Kingdom, the services of a former middle school principal in the Duval County school system were engaged to proctor the instrument in the United States. This former principal utilized her professional experience within the county to select and contact a diverse range of

elementary school principals. The interviews were arranged via telephone and were conducted during normal working hours at the participants' schools.

Demographic data were obtained from all participants in the United States. The average enrollment of the elementary schools was 724, ranging from 1320 for the largest to 265 for the smallest. Eleven of the 15 participants had been the principal of their schools in excess of seven years. The most recent in position had been there for one year. All principals had received some type of training in performance measures, and all but one thought it was adequate. Every principal had received training in either a workshop or class. Computer-based training was reported by two-thirds of the principals, and 11 of the 15 reported some type of self-study.

The sorts completed by the 15 principals were entered in PQMethod to determine the correlation among the participant sorts. The sorts were labeled with a numeric identifier to identify the participant and the letters US to indicate their group. Table 3 is the resulting correlation matrix:

Table 3

Correlation Matrix of Principal Sorts

SORTS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	100														
2	35	100													
3	37	13	100												
4	53	32	48	100											
5	51	44	45	69	100										
6	30	37	36	47	49	100									
7	50	25	47	48	51	38	100								
8	39	7	37	47	25	42	43	100							
9	49	51	48	48	48	46	49	51	100						
10	40	16	47	56	55	46	31	49	32	100					
11	49	24	40	43	47	48	32	36	43	30	100				
12	51	13	48	49	33	24	30	48	37	45	40	100			
13	19	-15	20	26	13	11	18	21	-2	23	17	31	100		
14	32	13	42	48	40	36	34	33	13	52	43	46	38	100	
15	44	46	47	55	55	43	48	34	53	47	40	37	12	46	100

Note: Decimals to two places omitted

There were a number of substantial correlations among the sorts of the principals, although the significance and instances were overall less than the head teacher group correlations. Although we must be careful not to infer too much from these correlation tables, this variance may reflect their reduced exposure to national performance measures from a chronological standpoint and subsequently more varied perceptions. The most substantial correlation was between sort 4 and sort 5. Sorts 5 and 1 substantially correlated to the most sorts, with four each. Sort 2 and 8 substantially correlated to only one sort, number 9, but did not correlate substantially to each other. Three sorts, number 3, 6 and 13, did not substantially correlate to any of the other principal sorts. This was one more than the head teacher group, and sort 13 reflected the only negative

correlation to any other sorts within the two groups. A comparison of the demographic data for this participant revealed no significant differences between that principal's data and those with negative correlations.

A principal component analysis resulted in two significant factors, with pre-rotational Eigen values of 6.46 and 1.6, respectively. These two factors were rotated using VARIMAX, to maximize the number of sorts that defined the two factors. Subsequent analysis resulted in z-scores for all the statements in relation to the two factors. PQMethod then used QAnalyze to convert these z-scores into corresponding values in the Q-sort distribution for each of the factors.

Factor C: Expand and Avoid Funding Link

The sorts completed by the principals revealed two collective perceptions. Factor C, "Expand and Avoid Funding Link," was defined by 9 of the 15 sorts from the participants. The sort in Figure 4 resulted:

Most Disagree						Neutral					Most Agree	
-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6
3	2	20	10	4	7	23	30	18	17	5	9	1
16	35	46	11	8	12	24	33	21	26	6	25	36
	42	52	13	22	15	27	43	34	40	41	32	
		58	14	29	19	28	45	39	44	56		
			37	57	54	31	48	47	62			
				61	59	38	49	50				
					60	53	51					
						55						

Figure 4. Principal Factor C array.

The nine principals who defined this factor shared attitudes with the other sub-group regarding elected officials setting school performance measure targets (statement 3) and the use of standardized test scores as the only school performance measure (statement 2). The Factor C sub-group further disagreed that the use of performance measure in schools had a positive impact on teacher morale (statement 16) and with linking school funding to progress on performance measures (statement 42). This group of participants also concurred in their strong disagreement that current school performance measures provided a comprehensive view of their schools (statement 20), but accepted them as part of their educational system for the foreseeable future (statement 20 and 35). They strongly disagreed that standardized test scores would emerge as the only performance measure to evaluate schools (statement 52).

This larger sub-group of principals agreed that school funding based on performance measures must consider social factors in the community (statement 9) while agreeing with the smaller group that performance measures could be used to assist school leaders (statement 5). This sub-group of principals decried the ranking of schools using performance measures, believing this practice ignores important factors (statement 36). The principals in both sub-groups strongly agreed that parents are using performance measures to differentiate between the effectiveness of schools (statement 41).

Factor D: True Believers

The remaining 6 sorts defined Factor D: "True Believers," and resulted in the combined sort depicted in Figure 5:

Most Disagree						Neutral					Most Agree	
-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6
22	2	3	61	12	10	4	9	15	1	41	5	23
35	27	28	13	29	11	8	16	18	6	38	32	56
	46	39	24	33	20	14	45	21	7	25	40	
		43	30	37	42	26	48	36	17	31		
			55	50	52	34	54	49	19			
				53	59	44	58	51				
					60	47	62					
						57						

Figure 5. Principal Factor D array.

In reviewing the representative sort of this smaller sub-group of principals, there were similar perceptions with the larger sub-group. This smaller sub-group shared the same opinion as the larger group regarding the use of performance measures to measure student learning (statement 46). However, this Factor D sub-group strongly disagreed with not using performance measures to rank schools (statement 28) and did not agree that public ranking of schools based on performance measures caused disruptive competition between schools (statement 39). They further disagreed with not using standardized test scores as a school performance measure (statement 22), although they perceived performance measures were more appropriate for businesses (statement 27). This

sub-group also perceived that the effectiveness of school performance measures had not been lessened by their status as a political issue (statement 43).

The statements of agreement for this sub-group focused on their personal effectiveness with respect to performance measures. Like the larger group of principals, they agreed that school performance measures, including standardized test scores, could be used to assist instructional efforts (statement 32). They also perceived that performance measures, when used with other techniques (statement 56), could be helpful for schools. There was a strong belief that their personal evaluations were linked to progress on performance measures. These principals strongly agreed that showing progress on their school performance measures was a major concern for them (statement 23), as their effectiveness was assessed on these measures (statement 38).

The principal group as a whole perceived that performance measures should be apolitical and that standardized test scores should be used as part of a balanced set of performance measures. They also held a realistic view that performance measures were an integral part of their future and would not be abandoned in the near future as a means to evaluate schools. The larger sub-group of principals perceived that linking performance on these measures to school funding was wrong and that there should be no linkage between the two. They also opined that other factors must be considered when implementing

performance measures. There were different attitudes between the two sub-groups regarding the use of standardized test scores, ranking schools, and the link to personal assessments of their effectiveness. The smaller sub-group thought the use of standardized tests as a performance measure acceptable, but did agree that balanced performance measures must be used to evaluate school performance. These six principals strongly perceived that the assessment of their effectiveness was linked to progress on performance measures and that this was a major concern for them.

Head Teachers and Principals

The sort data from both the head teacher and principal groups were combined into one data set and entered into PQMethod, and the correlation matrix is provided in Appendix F. Eleven sort pairs negatively correlated; however, it should be noted that 9 of these pairs included the same sort. That sort was principal sort 13, the same sort that was the only negatively correlated sort in the sub-groups. Sort 13 negatively correlated with one principal sort and 8 head teacher sorts. The most substantial correlation was between sort 9 from a principal and sort 27 from a head teacher. Sort 9 substantially correlated to the most sorts, 7 head teachers and 2 principals. There were more substantial correlations that consisted of head teacher pairs than of principal pairs. Nine pairs of substantially correlated sorts contained one sort from each sub-group.

Three principal sorts and 1 head teacher sort did not substantially correlate to any other sorts.

A principal component analysis resulted in the emergence of two significant factors once again, as had occurred in the principal and head teacher groups. The two factors had pre-rotational Eigen values of 11.667 and 3.0733, respectively.

Factor E: Test Skeptics

Factor E was defined by 14 of the 30 sorts, of which only 2 were from the principals in the United States. Figure 6 provides the factor array of Factor E: "Test Skeptics," defined predominantly by head teachers.

Most Disagree					Neutral					Most Agree		
-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6
2	58	42	46	35	51	49	32	33	39	40	34	36
16	3	52	57	37	54	50	43	41	53	9	6	1
	20	15	8	38	61	55	45	44	56	12	25	
		31	11	4	10	59	48	47	62	28		
			13	14	17	60	21	18	5			
				19	27	7	23	29				
					30	22	24					
						26						

Figure 6. Principal and head teacher Factor E array.

This sub-group strongly disagreed with the use of standardized test scores as the only performance measure when evaluating schools (statement 2). They did not agree that standardized test scores would emerge as the only performance

measure for schools (statement 52) or that they should be used as a means to obtain more funding from the government (statement 58). They did not agree their effectiveness was measured against standardized test scores (statement 15) and that performance measures must consider the social and economic factors of schools (statement 31). This sub-group agreed with the sub-group that defined the other factor with respect to role of elected officials in setting school performance targets (statement 3) and in the permanence of performance measures (statement 35).They also held similar beliefs regarding funding (statement 42) and the ability of performance measures to measure student learning (statement 46).

The Factor E sub-group strongly agreed that the use of standardized test scores as a performance measure causes teaching to the test (statement 34) and ranking schools using performance measures ignores other important factors (statement 36). They also thought this would have an adverse impact on school learning (statement 12) and that ranking based on performance measures should be avoided (statement 28). They did agree with the sub-group defining the second factor with regard to the political nature of the issue (statement 40) and that performance measures could be helpful for schools when used with other assessment techniques (statement 56).

Factor F: Fair Use

The second factor, Factor F, was defined by 16 of the 30 participant sorts. Of these 16 sorts, only 3 were head teachers in the UK. Figure 8 provides that the factor array for Factor F: "Fair Use":

Most Disagree						Neutral						Most Agree
-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6
23	35	42	37	39	33	48	34	49	36	40	32	56
3	46	55	11	43	53	50	44	51	38	9	41	5
	28	13	16	54	57	54	45	17	62	23	1	
		22	20	61	58	59	47	18	6	25		
			27	24	60	4	7	26	21			
				29	10	8	15	31				
					14	12	19					
						30						

Figure 7. Principal and head teacher Factor F array.

The participants who defined the second factor did not agree that showing progress on performance measure was a major concern of theirs (statement 23) but disagreed that performance measures should not be used to rank schools (statement 28). They strongly disagreed with not using standardized test scores as a performance measure (statement 22), and did not agree that the use of school performance measures diverted serious effort from assessing schools (statement 55). These participants also agreed that these measures, including standardized test scores, could assist teachers in their instructional efforts (statement 32).

They did agree with the sub-group that defined Factor E in that the purpose of performance measures was not to reduce costs (statement 13) and that funding should not be linked to performance measures (statement 42). This sub-group agreed with those defining Factor E that performance measures could be used to assist school leaders (statement 5) and that balanced performance measures must be used for schools (statement 1). They also agreed with this group that non-school controlled factors could directly affect school performance measures (statement 1).

Relationships of Data to Research Questions

The sorts completed by the participants provided data as to each participant's perceptions regarding the use of performance measures in their schools, and in doing so, answered the first research question. The factorial analysis of these sorts, both within the groups and as an aggregate, allowed the emergence of significant factors and their associated arrays. The interpretation of these factor arrays, representing common perceptions or viewpoints, allowed this researcher to recognize themes that could be both examined and contrasted between the various sub-groups. The discussion and conclusions based on the data are this researcher's interpretation of the factors that emerged. The names given to the factors are an attempt to capture the predominant theme that differentiated the common perceptions and opinions of the sub-groups. These

names are not meant to be all-inclusive as they do not, and cannot, simplistically capture the different perceptions of all the participants that defined the factor.

The emergence of significant factors, both for each participant group and as an aggregate, indicated identifiable clusters of opinions for these participant groups. Demographic data assisted in answering the research question by providing a means to establish some context of the participant responses, both within each group and during the comparison between the two groups.

Head Teacher Perceptions

The interpretation of these two representative head teacher sorts reveals both similar overarching opinions and unique perceptions on the use of performance measures in their schools. The two factors that emerged were:

Factor A: Inadequacies of Current Measures

Factor B: Use Balanced Measures and No Ranking

The entire group disagreed with the use of standardized test scores as the only performance measure when evaluating schools and thought their use caused teaching to the test. The larger group of head teachers defining Factor A perceived that social and economic factors affect their school performance measures and are not appropriately considered. As a result, they thought the current performance measures were inadequate in assessing the progress their schools were making. This was borne out by interview feedback from six of the

head teachers who were in the sub-group that defined this factor. Two mentioned that the economic and social background of the children was not considered. One of these head teachers, when providing the name of a potential participant, elaborated on the wealth of resources this head teacher had due to the surrounding affluent community when compared to the resources at his disposal. Another head teacher of a school in a challenging area commented on a specific student who, due to issues at home, was often left to care for her younger siblings and whose school work suffered as a result.

Although agreeing on this larger view of performance measures, the smaller group that defined Factor B seemed to share a perception that the use of balanced performance measures in schools had some value. They opined they could be used to rank schools if developed by teachers and divorced from the political stage. Their agreement that their effectiveness, as perceived by their supervisors, was based on school progress on performance measures indicates a certain degree of acceptance with regard to their use.

This shared perception that performance measures might have some value when developed and used in an inclusive and balanced manner was borne out by discussions with the three head teachers who were in the sub-group that defined Factor B. One of these head teachers remarked that the standardized tests were subject to inconsistent marking and decried the fact that passing or

failure could be determined by only one question. It is noteworthy that none of these head teachers were critical of performance measures in general, but rather of the use of standardized test scores as a performance measure.

Another head teacher in this group repeatedly advocated the use of performance measures when applied in a balanced manner. This head teacher sought clarification of some sort statements that advocated such a use. The third head teacher had restructured the curriculum with less focus on the test scores, and more on non-testable areas. These 3 head teachers all emphasized the merits of balanced performance measures as opposed to reliance on standardized test scores as the sole measure. This group also perceived ranking schools was disruptive, did not serve the public interest, and that performance measures should not be used for this purpose. The demographic data for these 3 head teachers did not differ significantly from the larger group.

Principal Perceptions

In reviewing the interpretation of these the two representative sorts, there were common perceptions regarding the impact of performance measures on teacher morale and that they should be apolitical. The two significant factors that emerged, Factor C and D, were described as follows:

Factor C: Expand and Avoid Funding Link

Factor D: True Believers

There were also unique perceptions in each of the two groups that focused on different areas. In the larger group that defined Factor C, there was a clear perception that social factors have an affect on performance measures and should be considered. These opinions were similar to those expressed by the head teachers who defined Factor A: Inadequacies of Current Measures. It is noteworthy that the most heavily defined factor in both the head teacher and principal groups addressed the deficiencies of current performance measures. The principals who defined Factor C also shared a strong belief that school funding should be divorced from performance measures.

One of the unique perceptions of the sub-group that defined Factor D was benefit in the use of performance measures for schools when not inordinately focused on standardized test scores and when used in conjunction with other performance assessment techniques. This smaller group also perceived that showing progress on performance measures was important, as they thought their own performance was assessed against such progress. The 2 principals with the least time in position, one with a year and the other with 3 years, were in this group of 6 principals. This may have influenced their beliefs regarding the use of performance measures, including standardized scores, as a means to assess their personal performance.

Comparing Head Teacher and Principal Perceptions

Prior to comparing the perceptions of the head teachers and the principals, a comparison of their demographic data is in order. This comparison revealed several similarities and differences that bear mentioning. One of the most obvious differences, acknowledged by this researcher prior to beginning the study, was the enrollment of the schools with respect to the country. The average enrollment of the head teachers' schools was 356, compared to 724 for the principals' schools. The schools in Duval County were effectively twice as large on average as those in Gloucestershire. The traditionally smaller schools in Gloucestershire County result from a policy that allows smaller villages and towns to retain their own schools rather than consolidating in larger schools that encompass a wider geographic area. There is a greater degree of consolidation in urban areas such as the city of Gloucestershire, where school enrollment for the participants from these schools was in some cases almost five times that of the village schools in this study.

Another difference was the number of head teachers, when compared to the principals, who thought their training on performance measures had been inadequate. Four head teachers, representing almost one-third of the group, thought so compared to only one principal in 15. Although the type of training in performance measures for both groups was for the most part similar, twice as

many principals had received computer-based training than their U. K. counterparts.

With respect to similarities between the two groups, years in position as head teacher or principal was the most obvious. Each group had 10 or more participants in these leadership positions for over 7 years. Although the type of training in performance measures for both groups was for the most part similar, twice as many principals had received computer-based training than their U. K. counterparts. A large portion of participants in both groups utilized self-study to increase their knowledge of performance measures.

In comparing the representative sorts from each of the two groups, the most evident difference is the number of principals or head teachers who defined each of the factors. When comparing the number of sorts that defined each factor in each of the two participant groups, it appears as if a larger number of head teachers shared a common perception with respect to the use of performance measures in schools. Twelve of the head teachers' sorts supported the most significant factor, as opposed to 9 of the 15 for the principals. It could be inferred from this loading that longer use of performance measures in the United Kingdom has resulted in a more commonly held belief system among these head teachers regarding their use.

There was a common belief among a majority of the participants in both groups that the social and economic factors of the school community must be considered when using performance measures. A similar majority opined that standardized test scores could not be the only performance measure for evaluating schools. The belief that performance measures should be apolitical and that political leaders should not set targets was also prevalent. There appeared to be no obvious disagreement among the head teachers and the principals on these aspects of school performance measures.

There were differences between the some participants in each of the two groups regarding their personal evaluations with respect to performance measures. The majority of the head teachers strongly agreed that their performance was not assessed against current performance measures. The principals who defined the second factor admitted that showing progress on performance measures was a major concern of theirs. It might be inferred that these head teachers are not as concerned given their relatively longer experience with these national-level performance standards.

Another aspect of the principal beliefs was an apparent acceptance of performance measures in schools as evidenced by their strong disagreement with the statement that they were only useful in a business environment. Although the head teacher group perceived that balanced performance measures could be

useful to school leaders, their advocacy and support appeared more guarded than that of the principals. The principals agreed with more statements that advocated the correct use of performance measures, both to measure progress and assist in teaching. This seemed to present the principals as more accepting of performance measures, although they clearly perceived that there should be no linkage to school funding.

Head Teacher and Principal Perceptions as a Group

In order to reveal common opinions and beliefs of the entire group of participants, the sorts of the head teachers and principals were entered in PQMethod as one group. Two factors emerged during the analysis, each defined by sub-groups that contained both principals and head teachers. The two factors that emerged were:

Factor E: Test Skeptics

Factor F: Fair Use

The sorts that defined the two factors, for the most part, followed national boundaries. Eighty-five percent of the sorts that defined Factor E were head teachers from the United Kingdom and similarly, eighty-one percent of the sorts that defined F were principals from the United States.

The theme of Factor E: Test Skeptics was aptly named. These 11 head teachers and 3 principals opined the manner in which standardized test scores

should be used was extremely limited. They clearly perceived that scores from these exams should not be the sole indicator of a school's performance and that they would never attain that status. Further, this group did not agree that their personal performance was measured against these exam results. The narrow utility they envisioned for standardized tests was further illustrated by their strong belief that performance measures must consider social and economic factors, a property not normally attributed to such exams. This group also shared common beliefs regarding the adverse effect of standardized tests, believing they cause teaching to the test.

The group of principals and head teachers that defined Factor F: Fair Use shared strong opinions that advocated the balanced use of performance measures, even for purposes of ranking schools. This group of 13 principals and 3 head teachers perceived that standardized test scores could be used as a performance measure and that they could assist teachers in their instructional efforts. Their view that school performance measures did not divert serious effort from assessing schools inferred the balanced use of such measures could be an integral part of the school processes. It is noteworthy that this group's strong belief in the fair and balanced use of performance measures included equally strong beliefs that there should be no linkage to school funding.

Chapter Summary

The results clearly answered the research questions that formed the framework of this study. A number of factors emerged both within the two groups and the aggregate group that provided insight as to the perceptions of these school leaders on this transnational issue. In the next chapter, these results will be discussed in their leadership context and their applicability for stakeholders. In addition to drawing conclusions in these areas, the next chapter will also address areas for potential research and the suitability of Q methodology in educational research.

CHAPTER 5

Discussion and Conclusions

The first chapter of this study described the purpose and research questions that provided the framework for this study examining primary and elementary school leader perceptions of performance measures. Four of the five research questions focused on the perceptions of a group of United Kingdom head teachers and United States principals that participated in the study. The last research question addressed the validity of Q methodology to collect the data. The chapter also defined terms unique to each country, those associated with the methodology, and others that were common in the literature. The chapter ended by describing the significance of this research to various stakeholders in both countries.

Chapter 2 contained a literature review that provided background and context to the issue of school performance measures in the United Kingdom and the United States. An historical overview of the use of performance measures in both countries examined their development and use in both business and education. The current policy of school oversight implemented in both countries by national-level legislation was also provided. The literature review continued by exploring key issues associated with this implementation, such as

accountability and the use of standardized tests. Finally, the chapter examined the role of the school leader and the impact of these performance measures on the leader's role.

The methodology and findings were provided in Chapters 3 and 4, respectively. The description of the design covered the methodology, selection and confidentiality of participants, procedures, and data analysis. The results of pilot studies conducted in both countries were also provided. A section that covered the delimitations and limitations of the design was also provided at the end of Chapter 3. The next chapter provided the findings from the study. Participant demographic data were discussed for similarities and differences in training, time in position, and school enrollment. Sort data were analyzed for the head teacher group, the principal group, and an aggregate group composed of all the participants. Two significant factors emerged in each of the groups, and their arrays were used to interpret the meanings. The composition of the sorts that defined each factor was examined, and a comparison between the two groups was provided. Chapter 4 ended by discussing the relationships of the data to the research questions, thereby setting the stage for the following major conclusions and recommendations.

Major Conclusions of the Study

This study revealed the perceptions of 30 school leaders on an issue that is at the forefront of educational policy in the United Kingdom and the United States. An overall examination of these perceptions, both in separate geographic groups and as one aggregate group, plainly indicated shared opinions and beliefs among these head teachers and principals. Although these perceptions cannot be generalized to a larger population, this does not diminish the message they convey. The following conclusions are the substance of that message and can be useful for those interested in understanding how school performance measures have impacted this leadership role in both countries.

The Adverse Effects of Focusing on One Performance Measure

The predominance of standardized test scores in assessing school performance, coupled by their use in an accountability-based policy by both countries, was a major theme that emerged in the literature review and the participant sorts. Sorts completed by the head teachers and principals lamented this predominance of one measure and the overall political nature of school performance measures in general. This common perception of test scores, given the chronological difference of their implementation as a component of accountability-based systems in both countries, bears consideration by policy makers. The message that current policies convey to these school leaders by

weighing these scores so heavily is one of political posturing rather than a sincere desire to improve schools. As long as this perception persists, the organizational improvement benefit of performance measures for these primary and elementary schools will be hampered.

Although the participants accepted the political reality of performance measures, there was an equally strong opinion that this type of reality is inappropriate and adversely affects the manner in which measures are developed and reported. Several head teachers in the United Kingdom voiced their dissatisfaction with the unofficial ranking carried out by the media following release of stage test scores. Another group of the participants indicated a preference for teachers to play a vital role in the development of performance measures. Clearly, these participants were of the opinion that the political nature of school performance measures made their proper development and implementation problematic. The prevalence of political targets and their associated media coverage would make it challenging to include education professionals and inform the public of the proper context in which operational measures such as test results should be considered.

One of the most strongly held opinions among both the head teachers and the principals was that social and economic factors must be considered when implementing performance measures in schools. The unbalanced focus on

standardized test scores makes consideration of these factors essential. Schools in affluent areas will have more resources at their disposal to increase student test scores. One principal lamented the fact that she would have responded differently to some of the sort statements had she still been posted in an inner-city school in an economically disadvantaged area.

Those schools with a large English as a second language enrollment will be challenged to prepare their students for nationwide exams that cannot, by the nature of their broad application, take diverse social factors into consideration. One head teacher described a question on a recent reading and writing exam that did just that. She described a student from a very conservative culture who was asked to write about the conversation between himself and his parent if he wished to stay up late. The student's response, expected by exam graders to be a prolonged discussion, instead consisted of the student asking the question once and being told no by the parent. This student's culture did not allow for such discussions with one's parent and as a result, his test score suffered.

Advocacy for Balanced Performance Measures

There is little doubt that balanced performance measures are a proven organizational tool to improve effectiveness. Many successful companies utilize the Balanced Scorecard approach, which includes "measures on customer satisfaction, internal processes, and the organization's innovation and

improvement activities" (Kaplan & Norton, 1992, p. 1) instead of just productivity measures. The realization that this approach was effective took the international business community many years to reach. The business sector, driven by competitive forces, evolved its use of performance measures to a more holistic model.

It can be argued that the public education sectors in the United Kingdom and the United States, while at different points of adoption when compared, are both in the same early stages of performance measurement that the business sector experienced. The predominant focus on one measure to assess performance improvement and the heavy accountability overtones suggest a parallel to those early stages of business performance improvement. Narrow financial measures were used to define successful and failing companies. An inordinate focus was placed on units of production in much the same way that focus is now on standardized test scores. Efforts directed solely at increasing units of production in business ignored quality, the workforce, and customer relationships. That focus ultimately proved the undoing of several companies in the competitive global marketplace and may have similarly adverse consequences for schools if policy makers use that approach. The question may be whether there will be sufficient motivation in the educational sector,

comparable to competition in the business sector, to shift the focus away from narrow performance measures.

One of the prevalent opinions that emerged from both the head teacher and principal sorts addressed the manner in which school performance measures should be used. There was a perception that properly developed and implemented performance measures can provide benefit to elementary and primary schools. The strength of this perception is significant, given the adverse effects associated with the manner in which they are currently implemented in the United Kingdom and United States. Despite the skewing of this tool towards accountability, these leaders have managed to appreciate the contribution balanced performance measures can make. External stakeholders and policy makers should appreciate the dichotomy of these leaders endorsing the overall use of performance measures despite the current narrow focus and heavy accountability overtones.

Leadership at the Point of Implementation

The challenges school leaders face with the implementation of performance measures in schools, given their current focus and use, are daunting. They will have to address staff concerns in this area, convey the meaning of exam results to parents, and meet targets set by supervisory bodies, all while trying to improve the organization. The leaders who participated in this

study opined that the use of performance measures did not contribute to teacher morale in their schools. Employee and, indeed, customer satisfaction--areas critical to organizational growth and success--may be impacted negatively by the emphasis on one performance measure. Prospective and current leaders in both countries will have to consider they may be faced with the same situation.

School leaders who wish to use performance measures for organizational improvement will face a number of obstacles. In a resource-constrained environment, the comparative effort to develop and use non-operational measures can prove significant if no externally provided resources such as standardized tests are available. Coupled with the widely held perception that a school's success is based on this one operational measure, convincing staff to dedicate effort on other measures will most likely require concerted leadership effort. A narrow focus of performance measures could prevent schools from allocating the resources and effort to truly benefit from a process of continuous improvement using performance measurement.

Limitations and Delimitations of Study

The delimitations discussed in Chapter 3 addressed the refinement of both the sample group and the issue that was the subject of this research study. The data collected from these 30 participants, both the individual sorts and the factorial analysis of group sorts, addressed all of the research questions. The

sample size of 15 head teachers and 15 principals proved adequate for purposes of the research questions and the use of Q methodology. The participant patterns of opinion and factors revealed during the study were similar to schools of thought on the issue that emerged during the literature review.

With the exception of one limitation that was discussed in Chapter 3, none of the other limitations appeared to affect the collection or analysis of the data, and no new limitations emerged during the course of the study. The one limitation that appeared to manifest itself in a noticeable manner was the collection of non-sort data regarding perceptions in the two countries. The manifestation took the form of less of this data from the principals when compared to the head teachers. Although the participant Q-sort did not appear to be affected by the use of a different proctor in the United Kingdom than in the United States, the collection of this non-sort data may have been affected by the use of different data collection methods. The principals in the United States, in lieu of interviews, provided free response data relating to the issue in the form of written comments. The head teachers, on the other hand, were interviewed and their comments transcribed by the proctor. This difference in data collection meant that the principals were somewhat constrained in expressing their viewpoints as a result of having to write their own comments down when compared to the head teachers. Given the significance of this non-sort data in

corroborating Q-sort data, this limitation most likely hindered the thematic interpretation of the factors defined by the principals.

Recommendations

The recommendations of this study, based on the findings and conclusions, can be focused on three general groups. The first group is those policy makers charged with the development and implementation of school performance measures in both the United Kingdom and the United States. The second group is those individuals who, due to their professional development plan, could be considered prospective principals and head teachers in the two countries. The last category of recommendations is for a slightly less differentiated group, educational researchers.

For Educational Policy Makers

Notwithstanding the inability to generalize the results, this study has provided valuable insight with respect to school leaders' perceptions of an issue that increasingly knows no national boundaries and is at the forefront of educational policy. Two recurring perceptions resonated in both the head teacher and principal groups. Both of these perceptions should be considered by policy makers as they develop and modify policies regarding school performance measures.

The first perception concerned the nature of school performance measures. School leaders in these two groups perceived the measures had become a high-profile political issue for both their governments. The support of performance measures in these primary and elementary schools is hampered by the perception they are being used more as a political issue rather than to actually improve schools. Although the participants accepted the political reality of school performance measures, there was an equally strong perception that this is inappropriate and adversely affects the manner in which they are developed and reported. Several head teachers in the United Kingdom voiced their dissatisfaction with a political climate that fosters the unofficial ranking carried out by the media following release of stage test scores.

The second prevalent belief is one that emerged in both the head teacher and principal representative sorts. This was the belief that performance measures, when properly developed and placed in perspective, can provide benefit to elementary and primary schools. Given the accountability overtones of their initial implementation, the ability of these leaders to see the possible organizational benefit of these measures should be noted by policy makers. This indicates that the damage done to the practice of performance measurement in schools by their inappropriate use might yet be undone by implementation of more balanced measures.

The results of this study can be particularly useful for those in a position to influence the manner in which performance measure are perceived and utilized. That these two groups of participants, with such different chronological exposure to nationally mandated performance measures, might share common perceptions regarding the political nature of such measures and the manner in which they could be helpful is worthy of consideration by policy makers. For those policy makers whose overriding goal is school improvement, the advice of these leaders could not be more apparent.

For Prospective Principals and Head Teachers

For those education professionals in the United Kingdom and United States whose career path may lead to the top leadership position in primary and elementary schools, the opinions and beliefs of the leaders revealed by this study may have reinforced existing perceptions or provided food for thought. Whatever the case, a number of recommendations can be drawn from these leaders' perceptions that are appropriate for those aspiring to these positions.

First, it is apparent there are a number of stakeholder equities in a school oversight system based for the most part on accountability instead of improvement. The data from this study indicated that the head teachers and the principals understood the different manner in which their teachers, supervisors, and parents viewed performance measurement data. Aspiring leaders could

prepare for their leadership roles by learning about and understanding those equities as much as is possible prior to assuming the position. Understanding what each group of internal and external stakeholders views as important will aid communication of policy, school acheivements, and desired efforts from staff.

The administration of mandated standardized tests is a good example. The different ways in which this testing, and the subsequent results, will be viewed will be at least as varied as the number of stakeholder groups. For a parent audience that is inundated with school rankings in the media based on these scores, school leaders may wish to convey that this is not a comprehensive view of the school. One head teacher in this study, whose school had not distinguished itself in the league tables, related how she advised prospective parents to visit the school rather than rely on such reports when selecting the school in which to enroll their children. School leaders who can communicate performance measurement issues such as this in the context of individual stakeholder equities are more likely to be successful in conveying their messages.

The majority of the school leaders in this study understood the benefit of balanced performance measures despite current policies that focus on accountibility and one measure. Prospective school leaders should ensure they understand the manner in which such measures can be developed and used so they can be used to their organization's advantage. It would be understandable

for school leaders to avoid the use of anything but mandated performance measures due to their heavy accountability overtones. However, in doing so, school leaders would deny themselves an important tool that has been demonstrated to improve organizational performance. The school leaders in this study understood that other social and economic factors must be considered when developing and implementing performance measures. For those who aspire to be head teachers and principals, distinguishing incorrect application of performance measures from their beneficial use could be key to accurately reporting school progress and focusing limited resources.

For Educational Researchers

For educational researchers, this study's recommendations were twofold. First, the study provided further validation of Q methodology as an educational research tool. The richness of data from this methodology, when compared to a Likert-type survey with no forced distribution, is startling. Clearly, this methodology can be a very useful tool for educational researchers who are conducting qualitative studies. The second recommendation of this study concerns potential areas for further study of internal and external stakeholder perceptions on this extremely relevant issue.

The use of Q methodology in this study provided remarkable insight as to the perceptions and opinions of these leaders. The range of educational research

issues that could benefit from such a revealing qualitative methodology is vast. Simplistic surveys that do not explore the full range of participants' beliefs on these issues will be hard-pressed to provide similarly meaningful data. This methodology can be especially useful in understanding the human component of the complex issues that face the educational community.

Q methodology challenges the participant to consider many dimensions of a particular issue and their own beliefs in a remarkably unique way. In addition to weighing the opinion statements against their own beliefs, participants must weigh the statements against each other. This instrument, by forcing the participant to rank order the statements, allows meaningful and interdependent data to emerge. The format forces participants to consider their own views against a myriad of other viewpoints on that issue. Following the factor analysis, the data allow for substantive qualitative interpretation by the researcher. Themes can be examined on an individual or group basis. For analyzing a groups' sorts, any clustering of opinion can provide a revealing window into shared perceptions and beliefs.

Just as the perceptions of these principals and head teachers have provided valuable insight into the perceptions school leaders have regarding performance measures, similar benefit may be obtained from examining other stakeholders within the educational systems of both countries. Teacher sample

groups that completed sorts with a more applicable Q-set of statements might provide these leaders, among others, with useful insight for tailoring the use of performance measures to their unique organizational climate. Parents, another important stakeholder in the use of performance measures in schools, might provide valuable insight as to how they view their use, and in doing so, reveal areas that should be addressed. A number of studies have examined the effects of accountability based systems for school oversight. An examination of the blame associated with failing schools in New Zealand and England (Thrupp, 1998) was one such study that contrasted the school inspection regimes in the two countries. The effect of these accountability-based systems on stakeholder opinions and perceptions could be further examined in these efforts to explore the issue of school performance measures.

Conclusion

The results of this study provide insight as to how these leaders perceive performance measures should, and should not be, utilized so the measures might be useful in improving their schools. The message their perceptions send is clear: use a balanced set of performance measures and they will be accepted and supported by school leaders and their staffs. The disproportionate weight given to standardized test scores for purposes of evaluating and ranking schools is categorically rejected by these leaders. A link between performance measures of

any kind and funding is similarly rejected. These leaders do agree that the use of standardized test scores, in concert with other performance measures, can improve both teaching efforts and school performance.

The head teachers and principals in this study are under no misconceptions regarding the permanence of school performance measures and accept that they are firmly ensconced as part of their school governance system. With this acceptance comes a belief that further effort in their development and implementation is needed so that they may achieve their full potential as a leadership and organizational tool.

APPENDIX A: Q-SET

1. A balanced set of performance measures must be used for schools.
2. Standardized test scores should be used as the only school performance measure when evaluating schools.
3. Elected officials should set school performance measure targets.
4. School performance measures should be used primarily for evaluating schools.
5. Performance measures can be used to assist school leaders in improving their schools.
6. Standardized test scores have overshadowed other school performance measures.
7. Teacher work-place satisfaction should be equally weighted with other school performance measures.
8. Ranking schools using school performance measures serves the public interest.
9. School funding based on performance measures must consider social factors of the community.
10. School performance measures should be equally weighted in importance.
11. Performance measures for schools assure accountability to taxpayers.
12. School performance measures can have an adverse impact on student learning.
13. The underlying goal of school performance measures is to reduce cost.
14. Parents believe standardized test scores are the only performance measure necessary for student achievement.
15. My effectiveness, as perceived by my supervisor, is based largely on one school performance measure, standardized test scores.

16. The use of performance measures in schools has had a positive impact on teacher morale.
17. School attendance rates should be a major school performance measure.
18. Ranking schools using balanced performance measures would serve the public interest better than current systems.
19. Performance measures are used primarily to improve schools.
20. Current school performance measures provide a comprehensive view of the school.
21. The effectiveness of school performance measures should be objectively reviewed on a periodic basis by the originating authority.
22. Standardized test scores should not be used as a school performance measure.
23. Showing progress on school performance measures is a major concern of mine.
24. Collecting performance measure data places a burden on school resources.
25. Non-school controlled social factors can directly affect school performance measures.
26. Students should not have to attend schools with unsatisfactory performance measures.
27. Performance measures are more appropriate for businesses than schools.
28. Performance measures should not be used to rank schools.
29. The weight of each school performance measure should not be disproportionate to the others.
30. Public perceptions affect the number and relative weighting of school performance measures.
31. Performance measures should be uniform for all schools regardless of economic or social conditions.

32. School performance measures, such as standardized test scores, can assist teachers in their instructional efforts.
33. School performance measures place an administrative burden on teachers.
34. The use of standardized test scores as a performance measure causes "teaching to the test."
35. The use of performance measures in schools is a passing "fad" that will soon be replaced by a new management concept.
36. Using performance measures to rank schools ignores other important factors.
37. Using more than a few school performance measures to evaluate performance is counterproductive.
38. My effectiveness, as perceived by my supervisor, is based largely on my school's progress on performance measures.
39. Public reporting of school ranking, based on performance measures, causes disruptive competition between schools.
40. The use of school performance measures to evaluate schools is a politically charged issue.
41. Parents use performance measures as a way to differentiate between the effectiveness of schools.
42. Funding for schools should be linked to progress on performance measures.
43. School performance measures are not effective because they have become a political issue.
44. Community economic data must be considered when reviewing school performance measures.
45. Performance based funding increases the administrative reporting burden for schools.
46. Student learning cannot be measured by performance measures.

47. Parents and teachers should have a larger voice in deciding what school performance measures should be used.
48. The effectiveness of elected officials should not be linked to school performance measures.
49. Schools not making progress on performance measures should receive additional funding.
50. School leaders are not being engaged in the development of school performance measures.
51. The use of performance measures to evaluate schools will increase in the future.
52. Standardized test scores will emerge as the only performance measure used to evaluate schools.
53. Parents have not been engaged in the development of balanced performance measures for schools.
54. The goals of performance measurement in schools are clearly articulated to school leaders.
55. Using school performance measures to evaluate schools diverts serious effort from assessing schools.
56. School performance measures, when used with other performance assessment techniques, can be helpful to schools.
57. Performance measures can be useful in the fiscal administration of schools.
58. School performance measures should be used as a means to obtain more funding from government.
59. Performance based funding for schools will be too difficult for government to effectively monitor.

60. School performance measures would be more meaningful when combined with performance measures from other public agencies serving the same community.
61. The public drive for accountability in schools has limited the benefit of school performance measures.
62. Teachers should play a vital role in the development of balanced school performance measures.

APPENDIX B: UNF IRB APPROVAL

05-04

University of North Florida
Division of Sponsored Research and Training

REQUEST FOR REVIEW BY INSTITUTIONAL REVIEW BOARD FOR THE PROTECTION OF HUMAN SUBJECTS

Principal Investigator: Rene Velez, Doctoral Candidate

Faculty Advisor (if student project): Dr. Kathe Kasten, Dissertation Chair

College/Dept.: College of Education and Human Services Campus Address & Phone: 904-620-1789

Project Title: Perceptions of School Performance Measures: A study of Principals and Headmasters in the United States and the United Kingdom using Q Methodology

Project Type: ☒ Non-funded (or student research) ☐ Externally Funded Research

Supporting Agency (if any) _____ Deadline: _____

Dated Submitted to DSRT for Review: 6/20/05 Project Terminatinn Date: 12/31/05

Type Review Requested: ☒ Exempt/Category # 2 ☐ Expedited/Category # _____ ☐ Full IRB

IRB USE:

Regulatory/Ethical justification for type of review: _____

Type of Review: ☒ Same as Requested ☐ Other than Requested (Type) _____

Subjects at Risk: ☐ Yes ☐ No ☒ Minimal

If yes, potential benefits justify proceedings ☐ Yes ☐ No

Informed Consent Required: ☒ Yes ☐ No If yes, forms approved: ☒ Yes ☐ No

Modifications Required: _____

Second Review Needed: ☐ Yes ☒ No

1st Review By: _____ 10/12/05 (signature and date)

☒ Approved ☐ Conditionally Approved, Pending Modifications ☐ Referred to Full IRB

IRB meeting date (full IRB review) _____

2nd Review By: _____ (signature and date)

☐ Approved ☐ Conditionally Approved, Pending Modifications ☐ Referred to Full IRB

IRB meeting date (full IRB review) _____

Right to Appeal: ☐ Make appointment with Vice President for Academic Affairs

RCVD PM 3:38 JUN 15'05

Rev. 08/03

05-104

University of North Florida
Division of Sponsored Research and Training

CERTIFICATION OF PRINCIPAL INVESTIGATOR

VIII. Certification of Principal Investigator

I have read and understand the U.S. Department of Health and Human Services procedures concerning research involving human subjects as stated in the June 1991, *Federal Register* announcement of policy, and I will abide by them. In addition, I accept the following responsibilities:

A. Principal Investigator:

1. I will obtain approval from the IRB prior to instituting any change in project protocol.
2. I will bring to the attention of the IRB the development of any unexpected risks or adverse effects.
3. I will keep signed consent forms, if required, from each experimental subject for the duration of the project.
4. I will submit a status report at twelve (12)-month or shorter time intervals (as indicated on the approval letter) attesting to the current status of the project.
5. If applicable to my project, I have attached a copy of the informed consent form(s) and a copy of the test instrument(s) for my project.

I accept the responsibilities indicated above:

Signature Deleted

Signed: _____ 6/13/05
Principal Investigator Date

B. Faculty Advisor (if student project)

I have collaborated in the development of the research proposal described in the attached and have reviewed all of the information enclosed and will oversee the work described. I will endeavor to ensure that all of the responsibilities described in "A" above are fulfilled. Confidential material and completed informed consent forms will be maintained in the Department or under its control.

Signature Deleted

Signed: _____ 6/21/05
Faculty Advisor (if student project) Date

UNF IRB Number: 05-104
Approval Date: 6/10/05
Revision Date: _____

05-104

University of North Florida
Division of Sponsored Research and Training

UNF IRB Number: 05-104
Approval Date: 10/12/05
Revision Date:

GUIDE TO FILLING OUT APPROVAL REQUEST

I. Objectives of Investigation

The objective of this study, using Q methodology, is to determine if there are any patterns of perception on the research issue. This methodology will be used to examine the attitudes and perspectives principals and headmasters have with regard to the use of performance measurement in their schools. Although these two groups have a similar level of leadership responsibility in that elementary and primary schools are roughly equivalent, national level performance measures have been in use substantially longer in the United Kingdom. This difference, coupled with the obvious geographic and curriculum differences, may reveal useful knowledge of similarities or differences both within and between these two leadership groups. The null hypotheses is that no patterns of perception will emerge from these participants.

II. Explanation for Use and Description of Subjects

The selection of the participants will be based on access and willingness to participate. In the United States, the researcher will enlist the services of Dr. Judy Poppel for data collection. Dr. Poppel is a retired Duval County elementary school principal who will utilize convenience sampling to select participants, contacting them initially via telephone. The data collection in the United Kingdom will be accomplished by the researcher. The participants will be selected based on an affirmative response to a letter requesting their participation. The sample letter is attached. The pool from which the participants are to be recruited is primary schools in Gloucestershire County, United Kingdom and elementary schools in Duval County, Florida, United States. The total number of participants will be thirty, divided evenly between the United Kingdom and the United States.
There are no experimental aspects of this study. This study will utilize the standard Q methodology to collect data. Bias in selecting participants is not a factor for this research methodology.

III. Methods or Procedures

The instrument begins by asking the participant to provide demographic data on their location, years in present position, school enrollment, and what type of training, if any, they have received on performance measurement. Detailed directions, explaining each step of the data collection effort, will be used in conjunction with the proctor to guide the participant. A "prompt", explaining the context of the statements, is the first step of the actual q-sort. The participant is asked to review the statements, provided in a list and on separate cards, in no particular order so as to gain an understanding of the themes. Following this review, participants will be asked to conduct a q-sort of the statements. The q-sort of the statements will be a forced distribution and rank ordered on a scale of +6 for Strongly Agree to -6 for Strongly Disagree, where zero is neutral. The statements will be grouped along this response scale by the participant after they consider the statements as compared to their own opinions and perceptions. A proctor will record the responses and conclude the interview by recording feedback on the research methodology and the statements from the participant.The data will be analyzed using software designed for Q methodology. The software and data will be stored on the Principal Investigator's personal computer, where only he will have access to it. Data will be reported in an aggregated and anonymous manner.

IV. Assessment of Risks and Benefits to Human Subjects

Participants will have an opportunity to learn more about the Q methodology as a research tool. The findings of this international study will also provide the participants with insight of their peer's attitudes, both in the United States and the United Kingdom, towards a current issue for schools.

There are no known risks in this study and no known risk of injury.

V. Measures to Protect Human Subjects

This research methodology presents no risks to the safety of the participants.

VI. Methods of Obtaining "Informed Consent" from Subjects

The proctor will arrange an interview with one or more participants. At the start of the interview, the proctor will provide a verbal overview of the research objectives and methodology. Participants will then be required to read the informed consent form and agree to participate in the study. Consent will be documented using the attached informed consent form. They will then be provided with the research instrument.

VII. How Results Will be Used

Data from this study may be published or used in publications. However, the name and identifying information of participants will not be published to protect confidentiality. Data may be used subsequently and will be stored indefinitely by the principal investigator on his personal computer.

VIII. Certification of Principal Investigator

I have read and understand the U.S. Department of Health and Human Services procedures concerning research involving human subjects as stated in the June 1991, Federal Register announcement of policy, and I will abide by them. In addition, I accept the following responsibilities:

A. Principal Investigator:

1. I will obtain approval from the IRB prior to instituting any change in project protocol.
2. I will bring to the attention of the IRB the development of any unexpected risks or adverse effects.
3. I will keep signed consent forms, if required, from each experimental subject for the duration of the project.
4. I will submit a status report at 12-month or shorter time intervals (as indicated on the approval letter) attesting to the current status of the project.
5. If applicable to my project, I have attached a copy of the informed consent form(s) and a copy of the test instrument(s) for my project.

I accept the responsibilities indicated abo**UNF IRB Number:** 05-104

Signed: _____ **Approval Date:** 10/12/05

Principal Investigator **Revision Date:** _____ Date

B. Faculty Advisor (if student project)

I have collaborated in the development of the research proposal described in the attached and have reviewed all of the information enclosed and will oversee the work described. I will endeavor to ensure that all of the responsibilities described in "A" above are fulfilled. Confidential material and completed informed consent forms will be maintained in the Department or under its control.

Signed: _____ _____
Faculty Advisor (if student project) Date

UNF IRB Number: 05.104
Approval Date: 10/12/05
Revision Date: _____

Rev. 1/30/2002

APPENDIX C: CONSENT FORM

05-104 R

UNIVERSITY OF NORTH FLORIDA

Human Research Consent Form

Title: Perceptions of School Performance Measures: A study of Principals and Headmasters in the United States and the United Kingdom using Q Methodology

Investigators: Rene Velez, Doctoral Candidate; Dr. Kathe Kasten, Dissertation Chair

Affiliations: None

Contact Information: Rene Velez - riveravelez@onetel.net, 011-44-124-222-7406;
Dr. Kathe Kasten- kkasten@unf.edu, 904-620-1789

Approved By Institutional Review Board: Not yet submitted

This is an important form. Please read carefully. It tells you what you need to know about this research study. If you agree to take part in this study, you need to sign this form. Your signature means that you have been told about the study and what the risks are. Your signature on this form also means that you want to take part in this study.

Your participation in this research is entirely voluntary. Refusal to participate in this research will involve no penalty or loss of benefits to which you otherwise are entitled.

You may discontinue participation in this research study at any time without penalty or loss of benefits you are otherwise entitled to.

What is the purpose of this study?
The purpose of this study is to examine the attitudes and perspectives of elementary school principals and primary school headmasters have with regard to performance measurement. Although these two groups have a similar level of leadership responsibility in that elementary and primary schools are roughly equivalent, national level performance measures have been in use substantially longer in the United Kingdom. This difference, coupled with the obvious geographic and curriculum differences, may reveal useful knowledge of similarities or differences both within and between these two leadership groups.

UNF IRB Number: 05-104
Approval Date: 10/12/05
Revision Date: _____

How many participants will take part in this study?
Approximately thirty.

What will happen in this study?
Participants will complete a "q-sort" as part of the Q methodology to collect data. The q-sort consists of a forced distribution and rank ordering of statements on a scale of +6 for Strongly

IRB #
Status
Approval Date

Agree to -6 for Strongly Disagree, where zero is neutral. The statements should be grouped along this response scale compared to the participants own opinions and perceptions. Participants will also be asked to provide limited demographic data. There are no experimental procedures associated with this study.

How long will I be in the study?
Forty-five minutes to one hour.

UNF IRB Number: 05-104
Approval Date: 10/17/05
Revision Date: _____

Are there reasons I might leave the study early?
Taking part in this study is your decision. You may decide to stop at any time. You should tell the director of the study that you wish to stop. In addition, the director of this study may stop you from taking part if it is in your best interest.

What are the risks of the study?
There are no known risks in this study and no known risk of injury.

What happens if I am injured because I took part in this study?
N/A

(If more than minimal risk, will I be compensated if injured? Will medical treatment be provided? If so, what will it consist of? Where can I get further information on this matter?)
N/A

Are there benefits to taking part in this study?
You will have an opportunity to learn more about the Q methodology as a research tool. The findings of this international study will provide you with insight of your peers' attitudes towards performance evaluation. If you desire, you will be provided with an abstract of the completed study.

What other choices do I have if I do not take part in this study?
N/A

Are there any monetary or other compensation or inducements for my taking part in this subject?
No.

IRB #
Status
Approval Date

UNF IRB Number: 05-104
Approval Date: 10/12/05
Revision Date: _____

Are there any financial costs to me to take part in this study?
No.

What are my rights if I take part in this study?
You do not have to take part in this study; but if you do, you may stop at any time.
You do not give up any of your rights by taking part in this study.

What about confidentiality?
Data from this study may be published or used in publications. However, your name and identifying information will not be published.

Explain your method further
Before beginning the q-sort you will be provided an opportunity to review background information on the issue and the opinion statements that will be used for the sort. The data collection is accomplished by reviewing and sorting sixty-two opinion statements on a graduated scale. These statements of opinion will address the use of performance measures in schools. The principal investigator or his assistant will record your responses on a form.

Will there be audiotaping or videotaping? If so, will I get to view them before they are used? Who will review tapes besides the researchers? Who will have access to the tapes? When will they be destroyed?

(Note – If tapes are to be used outside of the research project, a separate release form should be obtained.)
There will be no audiotaping or videotaping during this study.

Who can answer my questions?
You may talk to Rene Velez at any time about questions and concerns you may have about this study. You may contact Rene Velez at his home in the United Kingdom, 011-44-124-222-7406, 33 Redgrove Park, Cheltenham, Glos. GL51 6QY. Alternatively, you may also contact his Dissertation Chair, Dr. Kathe Kasten, at the University of North Florida, kkasten@unf.edu, (904) 620-1789.

You may get further information about UNF policies, the conduct of this study, the rights of research subjects or if you suffer injury related to your participation in this research project from the Chair of the Institutional Review Board, Dr. Kathaleen Bloom, at (904) 620-2684.

IRB #
Status
Approval Date

I have had an opportunity to have my questions answered. I have been given a copy of this form. I agree to take part in this study. I am over 18 years of age.

I am at least 18 years old. _____ (initials)

I have had the study that I am agreeing to participate in explained to me to my satisfaction. _____ (initials)

I have had the opportunity to ask any questions that I may have had regarding this study. _____ (initials)

I agree to participate in *(study name)* Perceptions of School Performance Measures: A Study of Principals and Headmasters in the United States and the United Kingdom using Q Methodology being conducted by
(PI) Rene Velez and the University of North Florida.

Date	Printed Name of Participant
	Signed Name of Participant
Date	Printed Name of Individual Obtaining Consent
Date	Signed Name of Individual Obtaining Consent

UNF IRB Number: 05-104
Approval Date: 10/12/05
Revision Date: _____

IRB #
Status
Approval Date

APPENDIX D: DEMOGRAPHIC QUESTIONNAIRE

Demographic information (please circle one)

School location: US UK

Years in present position: 1 2 3 4 5 6 7+

Do you feel your training on performance measures has been adequate? Yes No

The training was: (circle all that apply)

Workshop Self-Study Class Computer-based

In-Service Training

The student enrollment of my school is approximately: _____.

APPENDIX E: SORT PROMPT

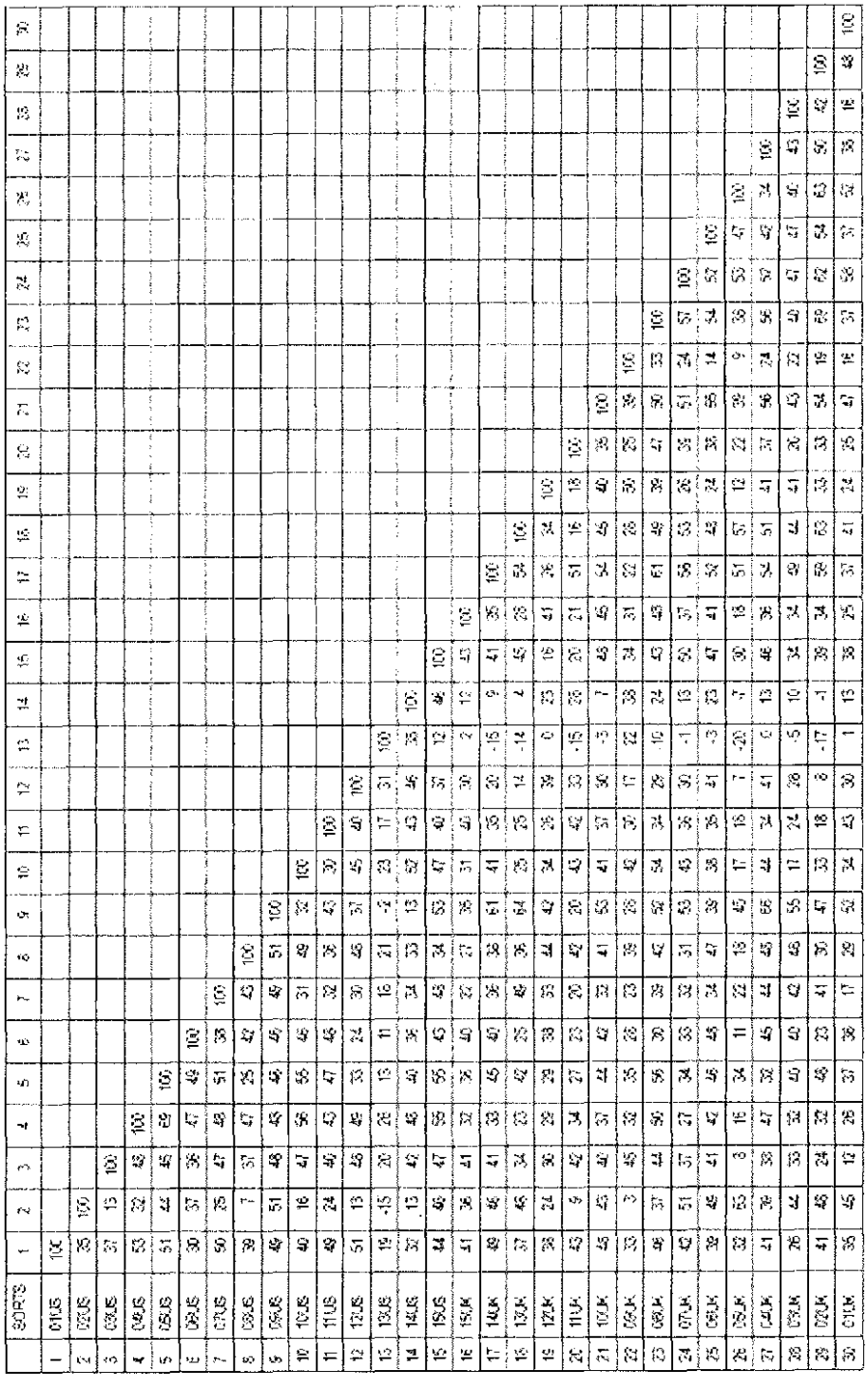
Performance measures have been used, in varying degrees, by the business sector both in the United States and United Kingdom for over forty years. The business sector initially focused on easily measured financial performance measures to measure success or improvement. There were also instances where massive data collection efforts were undertaken to provide information on numerous performance measures. In recent years, the business sector has learned that a narrow or broad use of performance measures may not accurately measure performance or provide useful feedback for improvement. Many organizations in the business sector now use a limited and equally weighted number of performance measures such as customer data, employee satisfaction, and internal business processes along with financial performance. This balanced approach has proved successful for a number of organizations and has become the predominant use of performance measures in the business sector.

The use of performance measures has spread outside the business sector to non-traditional areas such as publicly funded education. The appeal of these easily understood indications of output and success to government and the public has led to such use in the United Kingdom and United States. Government agencies have used performance measures as a means to evaluate schools, enforce accountability, and in some cases allocate funding. The public

has used these measures as a means to measure student achievement and as useful information when selecting schools for their children. These performance measures were also intended to provide school leaders with useful information in the administration of their schools. In the United Kingdom, schools are ranked nationally based on performance measures (test results) that are published in league tables. In the United States, State accountability systems and provisions of the No Child Left Behind Act have accelerated the use of such performance measures nationwide.

In much the same way as financial data initially dominated performance measures in the business sector, standardized test scores have emerged as the major indicator of student achievement, and by extension, school performance. The business sector eventually shifted from such a singular focus as this limited the potential benefit of performance measurement. A balanced approach, using a small number of equally weighted performance measures, may provide a more accurate assessment of school performance and provide useful information for school leaders. Some proposed performance measures that may provide a more balanced approach in education include teacher satisfaction, teacher professional growth, student attendance, and internal administrative processes. Measures reflecting student advancement and engagement could also be useful in this balanced approach.

APPENDIX F: HEAD TEACHER AND PRINCIPAL SORTS CORRELATION



REFERENCES

American Federation of Teachers. (2002). *Questions and answers about student testing/accountability.* Retrieved January 15, 2005, from http://aft.org/pubs-reports/reportlist.htmr

American Federation of Teachers. (2003). *Where we stand: Standards-based accountability and assessment.* (Item number 39-0246). Education Issues Department.

Anderson, P., Cuellar, N. C., & Rich, K. (2003). Performance improvement in higher education: Adapting a model from health care agencies. *Journal of Nursing Education, 42*(9), 416-420.

Arif, M., & Smiley, F. (2003). Business-like accountability in education: Commentary on Morrill to Baldrige. *Education, 123*(4), 740-762.

Bell, L. (1999). Back to the future: The development of educational policy in England. *Journal of Educational Administration, 37*(3), 200.

Black, P. J. (1994). Performance assessment and accountability: The experience in England and Wales. *Educational Evaluation and Policy Analysis, 16*(2), 191-203.

Brown, S. R. (1996). Q methodology and qualitative research. *Qualitative Health Research, 6,* 561-567.

Brown, B. (2004). *Illuminating patterns of perception: An overview of Q Methodology.*

[Technical Note CMU/SEI-2004-026].Hanscom AFB: HQ ESC/XPK.

Burke, J. C., & Modarresi, S. (2000). To keep or not to keep performance funding: Signals from stakeholders. *The Journal of Higher Education, 71*(4), 432-453.

Catano, N., & Stronge, J. H. (2006). What are principals expected to do? Congruence between principal evaluation and performance standards. *NASSP Bulletin, 90*(3), 221-237.

City of Newcastle. (2004). *The national curriculum explained.* Retrieved February 2, 2005, from http://www.newcastle.gov.uk/educlibnew.nsf

Connor, D. R. (1992). *Managing at the speed of change.* New York: Villard Books, Random House.

Davis, M. R. (2006, April 5). Study: NCLB leads to cuts for some subjects. *Education Week,* Retrieved May 7, 2006, from http://www.edweek.org

DeHavilland Information Services plc. (2005). *Debate – issue briefs: Testing in schools.* Retrieved on March 1, 2005, from http://www.politics.co.uk/issues/testing-in-schools-$2757075.htm

Deming, W. E. (1995). *Out of the crisis.* Cambridge, MA: MIT Center for Advanced Engineering Study.

Department for Education and Skills. (2003) *Excellence and enjoyment: A strategy for primary schools.* Retrieved on January 10, 2005, from http:// www.standards.dfes.gov.uk/ primary/ publications/ literacy/ 63553/

pns_excell_enjoy037703v2.pdf

Education Editor. (2005, February 19). Do better next term. *The Economist, Britain Edition, 107*, 53-54.

Education Week. (n.d.). *Glossary*. Retrieved October 7, 2006, from http://www.edweek.org/rc/glossary/accountability.html

Elmore, R. F. (2002). Testing trap: The single largest—and possibly most destructive—federal intrusion into America's public schools. *Harvard Magazine, 105*(1), 35-37, 97.

Englefield, S. (2001). *Leading to success: Judging success in primary schools in challenging contexts*. OFSTED, NCSL evidence base, www.ncsl.org.uk

Farkas, S., & Johnson, J., & Duffett, A. (2003). *Rolling up their sleeves; Superintendents and principals talk about what's needed to fix public schools*. New York: Public Agenda.

Flintham, A. (2003). *When reservoirs run dry: Why some head teachers leave headship early*. (Full Practitioner Enquiry Report). Nottingham, United Kingdom: National College for School Leadership.

Fritzberg, G. J. (2004). No child left behind? Assessing President Bush's assessment law. *Educational Foundations, 18*(3/4), 7-23.

Frost, B. (2000). *Measuring performance: Using the new metrics to deploy strategy and improve performance*. Dallas, TX: Measurement International.

Fuhrman, S. H., & Elmore, R. F. (Eds.). (2004). *Redesigning accountability systems for education.* New York: Teachers College Press.

Harbour, J. L. (1997). *The basics of performance measurement.* Portland, OR: Productivity Press.

Hart, P. D., & Teeter, R. M. (2004). *Equity and adequacy: Americans speak on public school funding* (Research Report Executive Summary). Princeton, NJ: Education Testing Service.

Hill, D. (2000). Cooperation or competition? What messages are our measurements sending our staffs about working together? *Strategic Finance, 81*(8), 53-57.

Hoff, D. J. (2006, October 18). Big business going to bat for NCLB. *Education Week,* Retrieved October 23, 2006, from http://www.edweek.org .

Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard: Measures that drive performance. *Harvard Business Review,* [Reprint 92105], 71-79.

Lyons, J. E., & Algozzine, B. (2006). Perceptions of the impact of accountability on the role of principals. *Education Policy Analysis Archives, 14*(16), 1-16

Massy, W. F. (2003). *Honoring the trust: Quality and cost containment in higher education.* Bolton, MA: Anker Publishing Company.

McKeown, B., & Thomas, D., (1988). Q Methodology. In M.S. Lewis-Beck (ed.) *Series: Quantitative applications in the social sciences.* Newbury Park, CA:

Sage Publications.

McMillan, J. H., & Schumacher, S. (1997). *Research in education: A conceptual introduction* (4th ed.). New York: Addison-Wesley Educational Publishers.

Meyers, K. (2001). *Leadership in action: Talking heads.* NCSL evidence base, www.ncsl.org.uk

Morgan, J. (2000). How effective are your measurement systems? *Purchasing, 129*(10), 26-28.

National Center for Education Statistics. (n.d.). *Assessment of educational progress.* Retrieved January 28, 2005, from http://nces.ed.gov/nationsreportcard

National Curriculum On-line. (n.d.). *About the national curriculum.* Retrieved January 30, 2005, from http://www.nc.uk.net/nc_resources/html/valuesAimsPurposes.shtml

National Education Association. (n.d.a). *Accountability and testing.* Retrieved on March 1, 2005, from http://www.nea.org/accountability/index.html

National Education Association. (n.d.b). *Adequate yearly progress.* Retrieved on March 1, 2005, from http://www.nea.org/esea/eseaayp.html

National Education Association. (n.d.c). *The nation's report card.* Retrieved on March 1, 2005, from http://www.nea.org/accountability/naep-accountability.html

National Institute of Standards and Technology. (2005). *Baldrige National Quality*

Program: Education criteria for performance excellence. Washington, DC: Author

National Union of Teachers. (2003). *NUT calls for government target setting in education to end*. Retrieved March 1, 2005, from http://www.teachers.org.uk/story.php?id=2783

Neill, M. (2003). High stakes, high risk: The dangerous consequences of high-stakes testing. *American School Board Journal, 190*(2), 4-6.

Ofsted. (2006). *Strategic Plan 2006 to 2007*. Retrieved October 29, 2006, from http://www.ofsted.gov.uk/assets/Internet_Content/Shared_Content/Files/stratpl0607.pdf

Olson, L. (2004). England refines accountability reforms. *Education Week, 23*(34), 1, 20.

Scherkenbach, W. W. (1991). *The Deming route to quality and productivity: Road maps and roadblocks*. Rockville, MD: Mercury Press.

Schmolck, P. (2002). *PQMethod - 2.11* [Computer manual]. Retrieved February 1, 2005, from http://www.rz.unibw-muenchen.de/%7Ep41bsmk/qmethod/pqmanual.htm

Schools Report. (2005, 13 January). *The Times (London, UK), daily ed.*, Secondary Schools Report, 1-20.

Siesfeld, T., & Pape, J. P. (2003). See the forest and the trees: Gaining visibility

into the real drivers of value. *Mercer Management Journal, 16*, 51-59.

Thrupp, M. (1998). Exploring the politics of blame: School inspection and its contestation in New Zealand and England. *Comparative Education, 34(2)*, 195-208.

U.S. Department of Education. (2002a). *No Child Left Behind Act of 2001: Executive summary*. Retrieved January 28, 2005, from http://www.ed.gov/nclb/overview/intro/execsumm.html

U.S. Department of Education. (2002b). *U.S. Department of Education strategic plan 2002-2007*. Retrieved on January 15, 2005, from http://www.ed.gov/about/reports/strat/plan2002-07/index.html

U.S. Department of Education. (2003). *The facts...about measuring progress*. Retrieved January 27, 2005, from http://www.ed.gov/nclb/accountability/ayp/testing.html

U.S. Department of Education. (2004). *10 Facts about K-12 education funding*. Retrieved January 27, 2005 from http://www.ed.gov/about/overview/fed/10facts/index.html

Walton, M. (1986). *The Deming management method*. New York: Perigree Books.

Watts, S., & Stenner, P. (2005). Doing Q methodology: Theory, method, and interpretation. *Qualitative Research in Psychology, 2*, 67-91.

Winkworth, I. (1999, December). *Performance measurement of U.K. university*

libraries (Issue 207). Washington, DC: Association of Research Libraries.

Woods, R. (2002). *Enchanted headteachers: Sustainability in primary school headship.* OFSTED, NCSL evidence base, www.ncsl.org.uk

VITA

René Velez

Educational and Professional Experience

Academic Degrees

Doctor of Education (in progress), University of North Florida, Jacksonville, FL.

Leadership and Adult Education, specialization in performance measurement

- Coursework completed with 3.82 GPA
- Anticipated Graduation Date: December 2006

Master of Arts: Adult Education, University of Rhode Island, Kingston, RI (1994)

Bachelor of Science, United States Naval Academy, Annapolis, MD (1982)

Professional Experience

Program Manager and Training Lead, US Department of Defense, United Kingdom (October 2004-Present)

Deputy Training Director, US Department of Defense, Maryland (October 2002-September 2004)

Executive Assistant to Special Assistant to the Chief of Naval Operations for Information Operations, Pentagon, Washington (October 2001- September 2002)

Program Manager, Defence Procurement Agency, Abbeywood, United Kingdom (May 1999-September 2001)

Director, Combat Systems Training Group, Mayport Naval Station, Florida. Designated Afloat Master Training Specialist. (October 1995-April 1999)

Curriculum Lead, Department Head Division, Surface Warfare Officer's School, Newport, Rhode Island. Designated Master Training Specialist. (September 1992-March1995)

Unrestricted Line (Surface Warfare) and Restricted Line (Special Duty, Cryptologic) Officer, (May1982- June 2003). Retired at the rank of Commander, United States Navy.