Expanding Data Analysis Skills in Educational Leaders:
Implications for Preparation Programs

Barbara Polnick, ***Sam Houston State University***
Stacey L. Edmonson, ***Sam Houston State University***

Abstract

The National Policy Board of Educational Administration reflects the belief that principals should be taught processes for experimenting and learning from real world data to meet the challenges of the work environment. This study of practicing principals yielded a content analysis of 482 responses, which reflect the need for principals to effectively use data available on their campuses. This study more clearly defined the content and strategies that should be used in the training of educational leaders, both practicing and prospective principals**,** which should be of significance to educational leadership preparation programs seeking to improve the relevance of their coursework.

Introduction

In the era of *No Child Left Behind* (United States Department of Education, 2002), principals nationwide have the responsibility of improving achievement for all students on their campuses. While this expectation to increase performance on mandatory accountability assessments for all students is a problem, no one solution that works best for every school has been found. Thus, principals are faced with exploring many solutions to meet the specific needs of their campuses. The ability to accurately and appropriately use the data made available to them through local, state, and national accountability measures is critical to principals' effectiveness at ultimately improving student achievement (Price & Burton, 2004; Yeagley, 2001). As Englert, Fries, Goodwin, Martin-Glenn, and Michael (2004) reported, "If

schools are not actively engaged in effectively using accountability data, generating the increases in student achievement required by this legislation [No Child Left Behind] becomes unattainable" (p.1).

The work of school principals has become increasingly more complex and demanding. As we move through the 21st century, principals face demands not only to be effective leaders but also to operate successfully in an environment of continuous change (Hoyle, English, & Steffy, 1998). To prepare principals who can meet these challenges, the National Policy Board for Educational Administration (NPBEA) created a training guide, *Principals for our Changing Schools: Knowledge and Skill Base* (Thomson, 1993). This publication reflects the belief that principals need to be taught the processes for experimenting and learning from real world data to be ready to meet the daily challenges of the work environment. While the NPBEA standards require principals to look at statistics and data analysis, very little training on how to gather and analyze data to make informed decisions is provided in the training manual or in many preparation programs. Nor are principals prepared well enough to effectively analyze and report their findings to their stakeholders, especially in the current age of data-driven accountability (Creighton, 2001; Holcomb, 2004; McNamara, 1994). In fact, Holcomb (2004) found that lack of proper training is one of the six barriers that prevent school leaders from effectively using data. In her words, the ability to effectively use data is a skill "that too few school leaders have had the opportunity to acquire in their graduate work or have seen modeled in their own experiences" (p. 27).

The 21 domains found in the NPBEA typology represent a convenient classification system one can use to better examine preparation strategies for school principals. The domains in the NPBEA typology are not discrete but rather interrelated, with 11 process- or skill-oriented domains and 10 content-focused domains. This typology reflects the belief that principals need to be taught a process for experimenting and learning from real-world data to be ready to meet the daily challenges they will face in the work environment. Thus, the overall goal of *Principals for Changing our Schools* (Thomson, 1993) was to develop professionals who have the understanding and skills necessary to address routine as well as emergent problems of practice.

An essential expectation elaborated in the NPBEA training guidelines was the need for practicing principals to develop basic statistics and data analysis skills that will assist them in their day-to-day operations of the school. The seven NPBEA functional domains address the organizational processes and techniques by which the mission of the school is achieved; in simpler terms, the data analysis skills in these domains are necessary for the principal to be effective on the job. The six NPBEA programmatic domains reflect instruction, learning environment, curriculum, student guidance centers, staff development, program evaluation, and resource allocation; the need for principals to possess basic statistical and data analysis skills is clear in these domains as well.

While two of the 21 domains in the proposed NPBEA training guidelines deal indirectly with data, information collection, and measurement, no domain deals specifically and explicitly with statistics and data analysis, which is a necessary skill for practicing principals. The standards do require principals to look at statistics and data-analysis within each domain, but provide little training on how to gather and analyze the evidence necessary to make successful and informed decisions.

Theoretical Perspectives

This exploratory study is grounded in the vast theoretical basis of the NPBEA's standards for principals. Based on these perspectives, five specific research questions were used to guide this inquiry:

(1) What are the most common situations where principals use data analysis skills to accurately define problems encountered in their work?

(2) What are the most appropriate data analysis strategies principals need to know to solve the typical problems encountered in their work?

(3) How can graphic displays be used to report the results of problem defining and data analysis so that these findings can be understood by school stakeholders who are not likely to have specialized statistical training?

(4) How can principal training programs effectively prepare principals to use the data analysis and statistical skills essential to solving problems on their campuses?

(5) What critical statistical concepts and strategies need to be included in a formative evaluation of staff development for principals to address the need for essential data analysis and statistical skills?

McNamara and Thompson (1996), as well as Holcomb (2004), observed that statistics courses as they are usually taught in graduate schools of education are not designed for the school principal. Specifically, these courses tend not to use real-world data when teaching principals the knowledge and skills needed to use statistics in their work environment (Bradshaw & Phillips, 2002). In addition, they note that "many statistics service courses taught in graduate schools of education direct a major portion of their syllabus toward the use of inferential statistics as a tool in conducting academic research projects and dissertations" (p.381). Thus, far less time is devoted to survey methods, estimation techniques, exploratory data analysis, and statistical graphs for reporting the findings of practical inquiries, which are the essential statistics and data-analysis skills principals need to be successful on the job.

McNamara and Thompson (1996) proposed an alternative strategy for developing data analysis skills in principals by emphasizing data analysis as a process

of discovery. This model, while focused on how educator preparation programs should teach data analysis skills to leadership students, includes the following guidelines:

(1) Emphasizing data analysis: Statistics is a set of methods used to analyze real-world data. Emphasizing data analysis to allow practitioners to focus on discovery is vital for producing accurate results that can inform school improvement efforts.

(2) Using real world data: Basic statistics courses should be taught as an integral part of the principal preparation program using real-world data that principals encounter in problem-solving and decision-making tasks in their job performance.

(3) Focusing on descriptive statistics: Principals typically use data on all students to solve pressing problems and to make decisions for their current academic year. This is very different from using data as sample evidence for comparisons with other schools, as educational research projects might do to influence future school years.

(4) Using accurate descriptions: This guideline recognizes all three essential properties needed to accurately describe a univariate distribution. These properties are the measure of center, measure of spread, and the shape of the distribution.

(5) Learning exploratory data: This guideline capitalizes on the fact that viewing data using the open-ended assumptions reveals truth about random fluctuations, error and other confusion often encountered in school data.

(6) Using graphic displays: This guideline emphasizes the importance of using data graphics in all aspects of real-world data analysis. Principals need to enhance their skill development in constructing and interpreting data graphics. Using well- designed data graphics provides a simple but powerful tool to interpret and enhance the report on real data.

(7) Reporting outliers: This guideline emphasizes why a principal should learn to analyze and report outliers. Using statistical graphic displays shows if the real-world data distribution is well-behaved or ill-behaved. This is a very important consideration for data analysis in schools since many real-world data distributions do not follow the idealized normal distribution often assumed in the application of many formal statistical methods. Principals must not only be prepared to recognize outliers that result in ill-behaved data distributions, but also must be prepared to take outliers into account so that their data analysis yields valid and reliable conclusions.

Methods and Procedures

This exploratory study was used to create an inventory of situations that describe opportunities for principals to use statistics and data analysis methods at the campus level. The process involved four steps: (1) identification of opportunities for data analysis, (2) development of a framework to guide content analysis, (3) creation of a set of categories for the concept "administrative tasks" and "data collection procedures," and (4) description of emerging trends from this analysis.

Participants in the Study

Participants were drawn from three different Prekindergarten-12th grade school districts, classified as urban, suburban and rural, according to the Texas Education Agency. These districts represented three different regions in Texas: Southeastern, Central, and Northern. Nineteen principals with a variety of experiences and background knowledge in both school improvement and statistics and data analysis volunteered to participate in the study. The schools in which the principals worked represented a wide variety of student demographics in terms of ethnicity, socio-economic levels, and English Language Learner status.

Data Collection

The first step in this phase was devoted to collecting opportunities for data analysis at the campus level. Three primary methods of data collection were used to accomplish this task. Specifically, opportunities for data analysis were entries recorded on personal logs, statements provided by 19 principals who agreed to participate in the study, and appropriate situations described in the literature on the school principalship. Taken collectively, this process produced 482 individual opportunities for conducting meaningful data analyses that could be used to inform decision making at the campus level.

Data Analysis

This step was completed when each of these 482 opportunities for data analysis was recorded on separate index cards. Putting each opportunity on a single index card provided a convenient set of input data to be used in all subsequent data analysis tasks undertaken in steps two through four. The second step was used to create a theoretical framework to guide all content analysis activities to be undertaken in steps three and four. A preliminary analysis of the 482 cards in the data set suggested that a two dimensional matrix would provide a meaningful way to synthesize the evidence assembled in these cards. The first dimension of this theoretical matrix was designed to specify the actual administrative task indicated on each card. This theoretical dimension was used to guide the content analysis activities in step three. The second dimension of the theoretical matrix was designed to identify the primary data analysis task

reflected on each card. This theoretical dimension was used to guide the content analysis activities in step four.

The third step was to create a meaningful set of categories for the concept *administrative task*. This step was undertaken using standard content analysis procedures detailed in Erlandson, Harris, Skipper and Allen (1993). Also helpful in constructing and implementing the step three content analysis were guidelines offered in Supovitz and Klein (2003), Lancy (1993), Merriam (1988) and Patton (1990). These content analysis procedures and guidelines as they were used in step three can be summarized as follows. First, the actual administrative task indicated on each card was identified and highlighted. Second, using concept coding, administrative tasks reflecting the same general concept were grouped together. Third, each group of cards was examined to ensure that each card was properly classified. Fourth, all initial groups of cards were then compared to see if some group categories should be merged or subdivided to better represent specific concepts. When appropriate, new groups were formed. Fifth, all groups of cards (both initial and newly formed groups) were examined again to ensure that there were no misclassifications. Finally, when all misclassifications were settled, each group of cards was given a common label (individual concept) that best reflected all of the administrative tasks contained in the group.

This iterative process yielded 21 emergent categories. While these 21 categories were not always completely independent, they were judged to be a meaningful and convenient way to describe the wide array of critical administrative tasks principals are expected to perform on the job. These 21 categories included (in alphabetical order): Budgeting; Campus Planning; Curriculum; Discipline; Extracurricular and Student Activities; Facilities and Auxiliary Services; Grants; Instructional Materials and Supplies; Instructional Setting and Design; Parent and Community; Personnel and Staffing; Professional Development for Administrators; Professional Development for Teachers; Program Evaluation; Special Programs; Student Assessment and Achievement; Texas Assessment of Academic Skills (TAAS); Tardies, Attendance and Enrollment; Teacher Observation and Evaluation; Technology; and Transitions Between Schools.

The fourth step was used to create a meaningful set of categories for the concept *data analysis procedures*. A review of basic statistics and data analysis texts (for example see Kitchens, 1997, and McNamara, 1994), indicated that four essential data analysis procedures can be used to create a meaningful set of categories. These four data analysis procedures (categories) are *defining problems*, *analyzing data*, *proposing solutions and reporting findings*. With these four theoretical categories in place, each of the 482 cards were placed into the one category that best reflected the primary intent indicated in the card statement detailing the opportunity for data analysis. The result yielded a univariate distribution for the concept *data analysis procedures*. The distribution has four specific frequencies. Each of these four frequency counts indicates the number of times one specific category was mentioned in the complete set of 482 data cards.

The resulting empirical matrix used to synthesize the data set is a 21 by 4 data matrix represented in Table 1. Inspection of this table indicates that 21 specific administrative task categories (displayed in alphabetical order) are used in the first dimension and four specific data analysis tasks (or procedures) are used in the second dimension. In addition, Table 1 also provides 84 specific frequency counts, one for each cell in the data matrix. Accordingly, the 84 entries in this bivariate distribution indicate both the specific administrative task and the primary data analysis procedure (intent) for each of the 482 cards in the data set. This bivariate distribution also provides a complete set of information needed to determine trends.

Results and Conclusions

The final step in this study was dedicated to describing trends that emerged from the analyses conducted in the four previous steps. Four specific trends emerged. Each trend is described below.

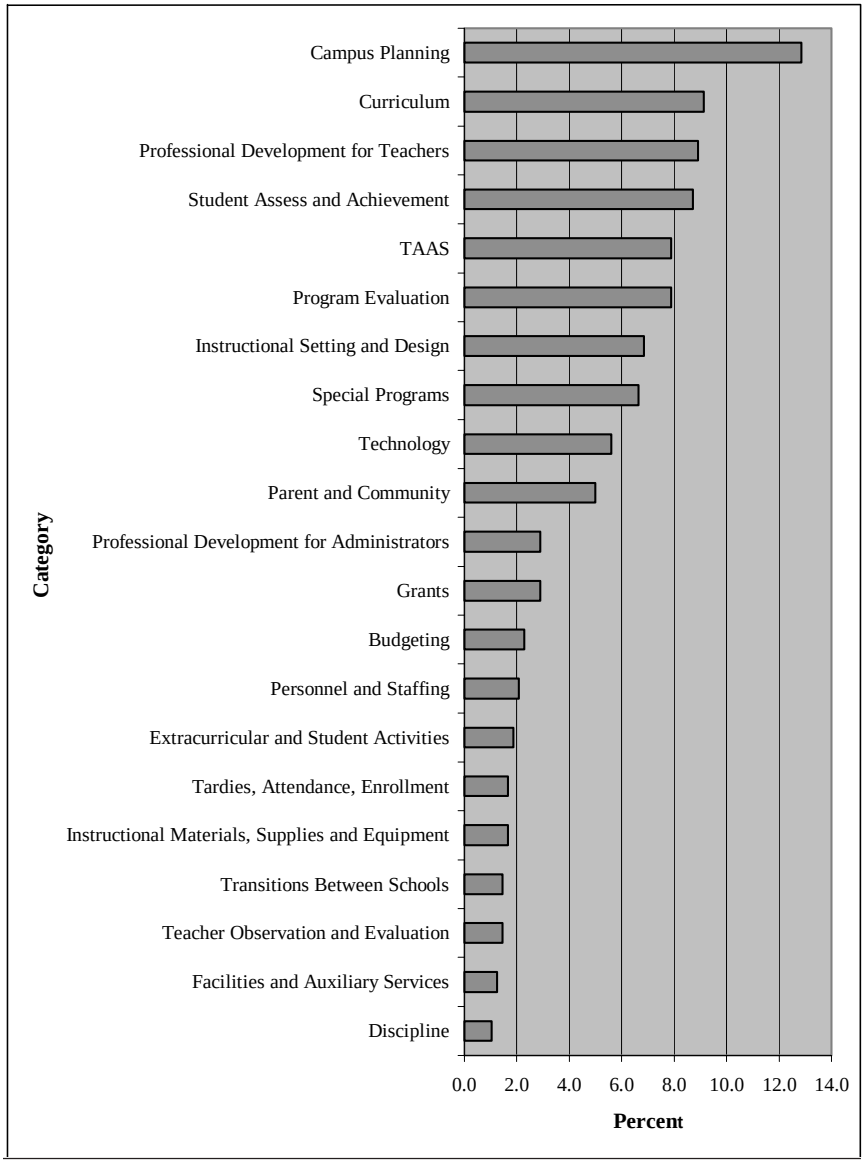
Table 1. Summary of Categories

Category	Defining Problems	Proposing Solutions	Analyzing Data	Reporting Results	Total
Budgeting	1	2	8	0	11
Campus Planning	9	11	34	8	62
Curriculum	8	9	22	5	44
Discipline	0	0	5	0	5
Extracurricular and Student Activities	4	2	3	0	9
Facilities and Auxiliary Services	2	0	3	1	6
Grants	5	1	8	0	14
Instructional Materials, Supplies and Equipment	2	2	1	3	8
Instructional Setting and Design	9	6	16	2	33
Parent and Community	9	4	5	6	24
Personnel and Staffing	1	1	7	1	10
Professional Development for Administrators	0	4	6	4	14
Professional Development for Teachers	21	5	11	6	43
Program Evaluation	19	3	13	3	38
Special Programs	9	7	12	4	32
Student Assess and Achievement	26	1	10	5	42
TAAS	8	2	22	6	38
Tardies, Attendance, Enrollment	1	2	4	1	8
Teacher Observation and Evaluation	3	0	3	1	7
Technology	20	2	2	3	27
Transitions Between Schools	4	0	2	1	7
TOTAL	161	64	197	60	482

Trend One: Administrative Tasks

Figure 1 is used to display the administrative task categories and their corresponding percents in rank order with the largest percent as the first rank and

Figure 1. Summary of Categories for All Responses, Sorted by Frequency



the smallest percent as the twenty-first rank. Inspection of these percents indicates that there is a large variability among the administrative tasks. Viewed from this perspective, a clear trend emerges. Four administrative categories contain approximately 40 percent of the 482 opportunities for data analysis at the campus level. These data yield the following summary information, as indicated in Table 2.

If one expands the highest priorities to include the top seven categories, approximately 62.23 percent of the 482 opportunities for data analysis fall into these seven categories. The fifth through seventh highest frequencies are described in Table 3.

Thus, just one-third of the administrative task categories account for more than three out of every five opportunities for data analysis. This trend suggests the following summary statement: While there is a wide range of opportunities for data analysis at the campus level, a small subset of administrative tasks (seven to be exact) account for a clear majority of these data analysis opportunities.

Trend Two: Data Analysis Tasks

Figure 2 provides a rank-ordered distribution of the data analysis procedures (skill areas) designated for the 482 index cards in the data set. This figure provides the evidence needed to declare the second trend. Better than two out of every five opportunities for data analysis fall into just one of the four data analysis skill areas. Specifically 197 index cards (40.87 percent) indicated the primary data analysis task of interest was analyzing data. The percentages and their distribution among categories can be seen more clearly in Table 4 below.

Notice also that combining the data for the first and second ranked data analysis task categories suggests that almost three-fourths of the data-analysis skills designations are either analyzing data or defining the problem.

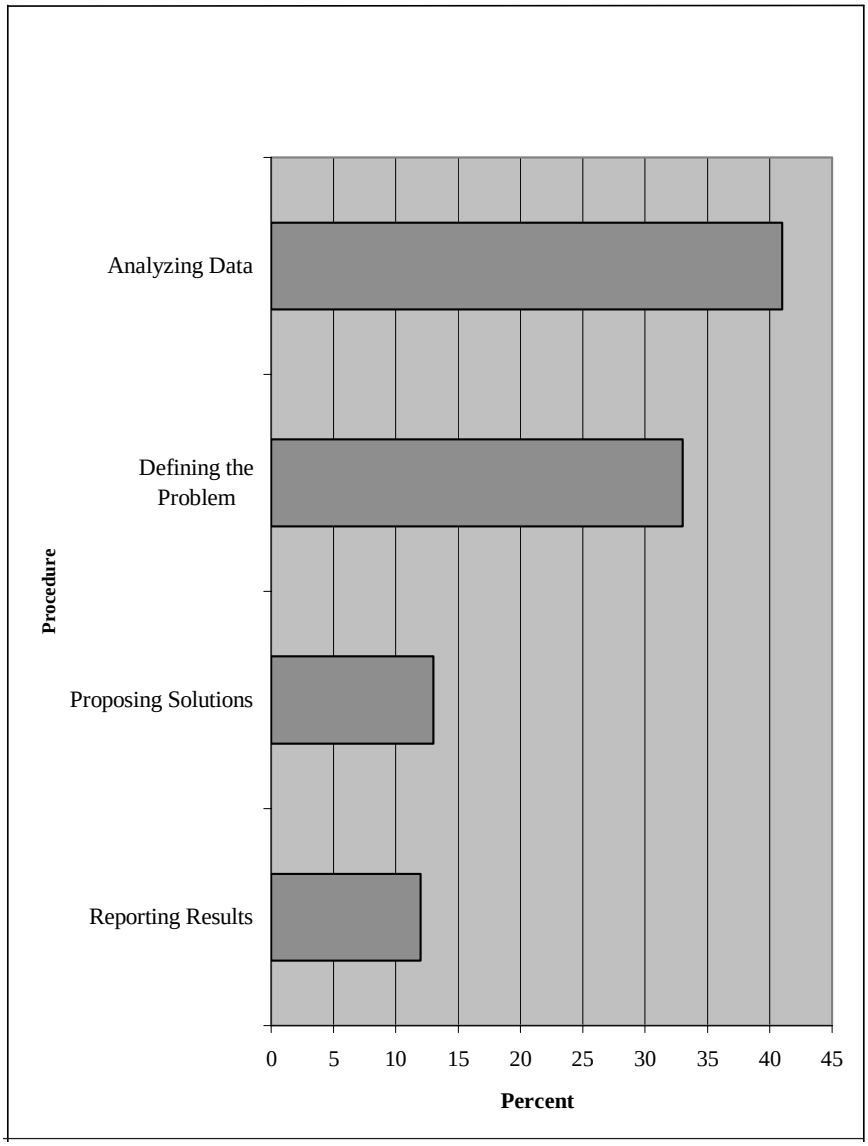
Table 2. Frequently Occurring Administrative Categories, First through Fourth

Administrative Task Category	Number	Percent
Campus Planning	62	12.86
Curriculum	44	9.13
Teacher Professional Development	43	8.92
Student Assessment and Achievement	42	8.71
Total	191	39.62

Table 3. Frequently Occurring Administrative Categories, Fifth through Seventh

Administrative Task Category	Number	Percent
TAAS	38	7.88
Program Evaluation	38	7.88
Instructional Setting and Design	33	6.85
Total	109	22.61

Figure 2. Summary of Data Analysis Procedures



This summary information yields the second trend of interest. Specifically, while there is interest in all four major data-analysis procedures (skill areas), most interest centers on analyzing data and defining problems rather than on proposing solutions and reporting results.

This trend has a "common sense" ring to it in that reporting results is a

Table 4. Frequently Occurring Administrative Categories, All Categories

Administrative Task Category	Number	Percent
Analyzing Data	197	40.88
Defining the Problem	161	33.40
Proposing Solutions	64	13.28
Reporting Results	60	12.44
Total	482	100

meaningful task only if the problems defined by principals are accurate and the corresponding data analysis procedures used in problem defining and problem solving are both reliable and valid.

Trend Three: Working With Existing Data

A content analysis of all 482 index cards indicated that better than nine out of every ten index cards reflected opportunities for data analysis using existing data available on the campus. Thus, less than ten percent of the opportunities for data analysis depend on collecting new data. Since existing data at the campus level almost always includes information on all individuals, programs and support services, the third trend provides clear evidence that the primary data analysis skills needed by principals reside in descriptive rather than inferential statistics.

Trend Four: Instructional Concerns

Inspection of the summary information provided in Figure 1 provides the evidence needed for the fourth trend. Specifically, the top seven administrative task categories in Figure 1 focus almost exclusively on instructional issues and concerns. This trend suggests that efforts to develop statistics and data analysis skills for school principals would be most effective when initial examples used to illustrate relevant data analysis procedures come from instructional situations encountered on the campus. Taken collectively, the four insights identified in this reflective exercise provided specific information the authors used in the third phase of this record of study, which turns attention toward the development of recommendations.

Recommendations

Guidelines offered in the McNamara and Thompson (1996) model suggest that teaching statistics in principal preparation programs should (1) emphasize data analysis methods that allow principals to focus on discovering actual relationships, (2) illustrate statistical concepts using real-world data principals are likely to encounter in the schools, (3) focus primarily on descriptive rather than inferential statistical methods, and (4) stress the use of graphics in both data analysis and report preparation.

Recommendation One

This recommendation draws on the profile of traditional graduate school of education statistics courses given in McNamara and Thompson (1996). They suggested that statistics courses as they are usually taught in graduate schools of education are not designed for the practitioner. Hence, these courses tend not to use school-based data when teaching principals the knowledge and skills necessary to use statistics in their work environment. With this in mind, findings from this study endorse the relevance of the McNamara and Thompson model for developing statistics and data analysis skills in principal preparation programs. This emphasis on inferential statistics clearly runs counter to the need documented in this study for concentrating on descriptive statistical methods as the most likely data analysis skills principals must use to be successful on the job.

Recommendation Two

A second recommendation from this study suggests that alternative graduate statistics course designed specifically for practicing and prospective school principals should begin by requiring all students enrolled to identify the actual data that individual campuses must collect and forward to the central office for preparing school district reports. Once these data are identified, students should then be asked to specify how these data could be used by the principal to inform planning and decision making at the campus level. Principals who understand how to apply the concept of rank ordering to the frequencies of a univariate distribution should be in a better position to prepare meaningful reports. In more specific terms, knowing how to rank order data on a variable of interest helps policy makers and practitioners to focus on developing priorities, which is an essential skill in both planning and decision making.

Creating this perspective at the beginning of the course will make a significant contribution to reinforcing the idea that basic statistical concepts have direct and immediate application in the schools. Once students are familiar with the data available at the campus and understand its potential value to inform campus decisions, focus should turn immediately toward illustrating how computer technology (specifically statistical data analysis programs) can be used effectively and efficiently to generate relevant data analysis and statistical graphics.

Recommendation Three

The formal introduction to basic statistical concepts in an alternative graduate statistics course designed specifically for practicing and prospective school principals should include a focus on these three skills development areas: (1) how to construct basic statistical graphics for univariate and bivariate distributions, (2) how to identify specific statistical concepts that allow one to discover patterns and trends that can be incorporated into graphic representations, and (3) how to use these graphic representations effectively in both data analysis and report preparation.

Furthermore, the development of data analysis skills needed to prepare an action research report should include a focus on statistical graphic methods, such as bar charts and box plots.

Recommendation Four

An alternative graduate statistics course designed specifically for practicing and prospective school principals should emphasize throughout the course these two key ideas: (1) the need to understand the important distinction between data and information and (2) the need to use this important distinction as a theoretical guide for designing effective data analysis procedures that answer critical decision-oriented questions encountered on the job. When properly implemented, this recommendation will provide the necessary first step that allows principal and their professional colleagues to move beyond the unfortunate situation where schools are data rich, but analysis poor.

Conclusions

This study more clearly defined the content and strategies that should be used in the training of educational leaders, both practicing and prospective principals. Furthermore, this study identified the most common opportunities for data analysis encountered by principals in their day to day jobs. Once these opportunities are identified, training programs can then facilitate the development of principals who are able to meet the data-driven challenges of school improvement and student achievement as suggested by the McNamara and Thompson (1996) model. Recognizing a principal's daily opportunities for data analysis and data reporting, including the use of graphic displays, supports the need for emphasis in these areas in principal training programs. Principals must be able not only to identify and use available data from their campuses to make effective, informed decisions, but they must also be able to effectively and accurately communicate this information to their stakeholders.

References

Bradshaw, L. K., & Phillips, J. C. (2002). Learning what they need to know: School leaders' lessons of experience. In G. Perreault & F. Lunenburg (Eds.), *The changing world of school administration* (pp.182-198). Lanham, MD: Scarecrow Press.

Creighton, T. (2001). Data analysis in administrators' hands: An oxymoron. *School Administrator, 58*(4), 52-57.

Englert, K., Fries, D., Goodwin, B., Martin-Glenn, M., & Michael, S. (2004). *Understanding how principals use data in a new environment of accountability*. Aurora, CO: Mid-Continent Research for Education and Learning.

Erlandson, D. A., Harris, E. L., Skipper, B. L., & Allen, S. D. (1993). *Doing naturalistic inquiry: A guide to methods.* Newbury Park, CA: Sage.

Holcomb, E. L. (2004). *Getting excited about data: Combining people, passion, and proof to maximize student achievement.* Thousand Oaks, CA: Corwin Press.

Hoyle, J. R., English, F., & Steffy, B. (1998). *Skills for successful 21st century school leaders.* Arlington, VA: American Association of School Administrators.

Kitchens, L. J. (1997). *Exploring statistics: A modern introduction to data analysis and inference.* Pacific Grove, CA: Duxbury Press.

Lancy, D. F. (1993). *Qualitative research in education: An introduction to the major traditions.* New York: Longman.

McNamara, J. F. (1994). *Surveys and experiments in education research.* Lancaster, PA: Technomics Publishing.

McNamara, J. F., & Thompson, D. P. (1996). Teaching statistics in principal preparation: Part one. *International Journal of Educational Reform, 5*(3), 381-389.

Merriam, S. B. (1988). *Case study research in education: A qualitative approach.* San Francisco: Jossey-Bass.

Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, CA: Sage.

Price, W. J., & Burton, E.M. (2004). Leadership skills for supporting learning. *Leadership, 34*(2), 20-22.

Supovitz, J. A., & Klein, V. (2003). *Mapping a course for improved student learning: How innovative schools systematically use student performance data to guide improvement.* Philadelphia, PA: Consortium for Policy Research in Education.

Thomson, S. D. (Ed.). (1993). *Principals for our changing schools: Knowledge and skill base.* Lancaster, PA: Technomic Publishing.

United States Department of Education (2002). *No Child Left Behind.* Retrieved February 15, 2003, from http://www.nclb.gov/

Yeagley, R. (2001). Data in your hands. *School Administrator, 58*(4), 12.

About the Authors

Barbara Polnick is an assistant professor in Educational Leadership and Counseling in the College of Education at Sam Houston State University, Huntsville, Texas. She earned her Ed. D. from Texas A & M University, College Station, TX and has extensive experience in the public schools as a teacher, regional education service center specialist, K-8 supervisor, Grade 7-8 assistant principal, and K-12 curriculum director. Her research interests and expertise lies in school improvement, curriculum design and evaluation of instructional programs, teacher leadership, and gender studies.

Stacey Edmonson is an associate professor and coordinator of doctoral studies in Educational Leadership at Sam Houston State University. Her areas of interest include qualitative methodologies, burnout and efficacy among educators, and school law. She has served Texas public school districts as a central office administrator, principal, and teacher. She is the current president of the Texas Council of Professors of Educational Administration and an executive board member of the National Council of Professors of Educational Administration.