educational considerations

The National Education Finance Conference®

NEFC

2013 Conference
Hyatt Regency
Indianapolis, Indiana
May 1-3, 2013

The Mission of the National Education Finance Conference (NEFC) is to provide a forum truly focused on school finance to enhance education for both young and adult learners. NEFC provides a venue for collaboration among legislators, post-secondary education, school district and state agency personnel, professional organizations, and researchers concerned with the importance of equity, adequacy, and efficiency concepts that affect state, local and federal revenue generation, distribution, and expenditures.

www.nationaledfinance.com

R. Craig Wood, Chair
Carlee Poston Escue, Deputy Chair
info@nationaledfinance.com

TABLE OF CONTENTS

Educational Considerations Design/Layout by
Mary Hammel, Kansas State University

Editors' Note

In this issue of *Educational Considerations*, we are pleased to offer readers extended, in-depth discussions of two critical issues for educational leaders and policymakers: Cost-effective factors that have the potential to improve student achievement and effective preparation programs for education leaders. We are honored to have two distinguished scholars to provide theory- and research-based insights into these topics which have challenged researchers, policymakers, and practitioners for decades. The first article, "A Theory of School Achievement: A Quantum View," by James L. Phelps, extends his research on class size reduction which was showcased in a special issue of *Educational Considerations* last fall. From that foundation, he has developed and operationalized a comprehensive theory of student achievement. His mathematical model provides researchers with a fresh approach to thinking about this important line of inquiry.

In the second article, "Doctoral Programs in Educational Leadership: A Duality Framework of Commonality and Differences," Perry A. Zirkel has collected and synthesized several decades of a wide range of literature related to the ongoing dialogue and debate on whether the Ed.D. or the Ph.D. best serves the needs of preK-12 educational leaders, in particular, school district superintendents. The organization of this wealth of information into a coherent framework is meant to assist not only those involved in the design and delivery of educational leadership doctoral programs, but also the practitioners who will enroll in them.

David C. Thompson
Chair, Board of Editors

Faith E. Crampton
Executive Editor

A Theory of School Achievement: A Quantum View[1]

James L. Phelps

Introduction

What is reality? In order to make predictions, all concepts in a scientific study and subsequent theory must be accurately represented by mathematical principles, and those concepts and principles must embody reality. Because there is no single universal concept and principle, complementary concepts and principles must be combined in order to comprehensively embrace everything observed and measured.

Early science was directed toward moving objects (e.g., balls down an inclined plane, orbits of planets) and the associated concepts, principles, and predictions were extremely accurate. Later, much different concepts and principles were accurately applied to the movement of electrons and photons within the atom. Now there are concepts and principles regarding people (e.g., personality traits and the learning curve) and organizations (e.g., effectiveness and cost-effectiveness).

Given this context, which point of view listed below better represents reality as schools seek to improve student achievement?

- Schools as moving objects: When the circumstances of the average school are known and changed in a specific way, achievement gains are certain because all schools react in the same predictable way—schools are identical.
- Schools as people or organizations: Individual schools behave distinctively and respond to changes of circumstances differently, so achievement gains can never be predicted with certainty and must be predicted by probabilities—schools are unique.

James L. Phelps set his early sights on composing music for movies, but he also had a keen interest in mathematics and science. Receiving a B.A. and M.A. in Music Education from the University of Michigan, he taught music in junior and senior high schools. He returned to the University of Michigan where he received a Ph.D. in Educational Administration in 1970. His career took an unexpected turn when he served as staff to the Governor's Commission on Educational Reform in Michigan and later became Educational Assistant to the Governor. Because of his interest in school finance, he became associated with the American Education Finance Association where he served as President. He served as Deputy Superintendent of the Michigan Department of Education, retiring in 1995. Currently he sings in two choirs, plays string bass in an orchestra, and continues to compose and arrange both instrumental and vocal music.

Only with the second point of view of reality can concepts and mathematical principles emerge to describe, explain, and predict individual school achievement, i.e., a theory.

The following excerpt, from a 1929 lecture by the quantum physicist Werner Heisenberg (2011,155) at the University of Chicago, illustrates the challenges involved in theory development:[2]

> The experiments of [education] and their results can be described in the language of daily life. Thus if the [educator] did not demand a theory to explain his results and could be content, say, with a description of [the relationships between various achievement and explanatory variables], everything would be simple and there would be no need of an epistemological discussion. Difficulties arise only in the attempt to classify and synthesize the results, to establish the relation of the cause and effect between them—in short, to construct a theory. This synthetic process has been applied not only to the results of scientific experiments, but, in the course of ages, also to the simplest experiences of daily life, and in this way all concepts have been formed. In the process, the solid ground of experimental proof has often been forsaken, and generalizations have been accepted uncritically, until finally contradictions between theory and experiment have become apparent. In order to avoid these contradictions, it seems necessary to demand that no concept enter a theory which has not been verified...at least to the same degree as the experiments to be explained by the theory.

Physical laws are established based on certain concepts and mathematical principles. There is a set of concepts and principles explaining with great accuracy the movement of objects, planets around the sun, and the moon and satellites orbiting earth, as follows:

- If the initial position and momentum are known, the position of the object in the future can be determined with great certainty; predictions are deterministic.
- The location of the object is continuous; an object such as a satellite can orbit any distance from earth.
- The concept applies without limits; an object can be anywhere in the entire universe.
- The only error in prediction is due to the restrictions of the measuring instruments.

In the late 19th and early 20th centuries, there were discoveries challenging these concepts and principles. The first discovery was that the speed of light was fixed, followed by Einstein's modification of Newton's formulation of planetary motion to what is known as the theory of relativity (Cox and Forshaw 2009, 87-89). Another discovery, Planck's quantum, led to concepts and principles fundamentally different than those of Newton and Einstein (Hawking 2011, ix). His discovery was not concerned with the macro world of space, but with the micro world of the atom. The quantum concepts and principles are substantially different. Below are some examples:

- Instead of objects moving through space the objects are electrons moving around a nucleus.
- Electrons behave both as a particle and a wave.
- An electron can only be in a shell a certain integer distance from the nucleus.
- The number of shells is limited.

- No matter how accurate the measurement instruments, there will always be uncertainty as to the position and momentum of the particle.
- The position and movement of the particles can only be measured by probabilities.

These discoveries and subsequent theory are known as quantum mechanics.

In "The Atomic Theory of Matter," Planck (2011, 42-43) described the difference between the macro- and micro-worlds. According to Planck, the macro-observer sees a gas. The only analytic method the macro-observer has to determine the position of an object is measurements from a substantial number of observations and calculation of the probability by finding the mean value, concluding that the mean value of a sufficiently large number of throws with a six-sided die is three and one-half. In contrast, the micro-observer sees only an individual atom. Therefore, this observer's interest is only in the probability of the position of an electron within the atom, and so concludes the probability of the one side of the die is one-sixth. If there are numerous observations plotted by X- and Y-coordinates, each with their unique location, the macro method to determine position requires calculation of the average of the X- and Y-values in order to find the average point. The probability of the average is the probability of the X-value times the probability of the Y-value ($1/2 * 1/2 = 1/4$). In contrast, the micro-method requires the calculation of every observation, each with its own probability. A unique probability for each and every observation is fundamental in quantum theory.

In most school achievement research, the relationships between achievement and explanatory variables follow the Newton and Einstein concept/principle and the viewpoint of the macro-observer: Deterministic measures based on the mean value of a sufficiently large number of schools. What if the relationships between achievement and explanatory variables followed Planck's quantum concept/principle and the viewpoint of the micro-observer; that is, the non-deterministic measurement of individual schools, each with its own probability? What influence would a quantum theory of school achievement have on research, training, and practice?

There is no set of generally accepted concepts or mathematical principles underlying the multiple diverse studies estimating the relationships between school achievement and various explanatory variables; in short, there is no comprehensive theory of school achievement. In this article, the purpose of the analyses and thought experiments, culminating in a series of postulates,[3] is to define the fundamental concepts and mathematical principles of such a theory. These issues are addressed in this article through discussion of the following:

- Why achievement measures are quantum in nature: discrete integer values with upper- and lower-limits requiring probabilistic measurements.
- Why normal curve statistics commonly used in achievement research are based on continuous variables with no upper- and lower-limits and implied deterministic measurements.
- How normal curve statistics can accommodate the quantum nature of achievement by considering the relationships between achievement and explanatory variables as nonlinear, nondeterministic, and probabilistic.
- How nonlinear relationships allow for the calculation of achievement levels and probabilities unique to each individual school (Planck's microview).
- How the nonlinear interpretation leads to a calculation of cost-effectiveness.
- How conceptually and statistically related variables can be combined to measure their collective influence on achievement.
- How normal curve statistics and combinations of explanatory variables can be used in a comprehensive theory of school achievement and mathematical model simulating how changes in individual school policies could influence the probability of improved achievement.

The Nature of Achievement: A Thought Experiment

Assume two students take a one-question test, on which in previous trials half the students got the question correct, a 50-50 chance. One student answers the question correctly and the other incorrectly for a scorecard of (1,0). These students then participate in a special program for which research predicts an increase of achievement score of .5. On a comparable single-question test, what are the predicted results? Will the first student increase his score? No, she is already at the limit and a score of 1.5 is impossible. What about the second, will his score be .5? Obviously no, scores come only in increments of 1. The scorecard remains (1,0). Moreover, there is no way to calculate an average. The average of (1, 0) is 1/2, an imaginary number because it is not a quantum integer number. If the requirements of limits and quantum measures are ignored, then the scorecard is (1.5, .5). If the projected increase of score is 1, then by the same logic the new score card reads (1,1), and further increases are not possible.

Now the same situation is interpreted with quantum probability measures. The probability of both students achieving a correct answer starts at .5, a 50-50 chance and a scorecard of (.5, .5). If research suggests an improvement increment of .1 the scorecard is (.6, .6). The average of .6 is a real number. Further increases are possible. The inconsistencies of the first interpretation are eliminated. The numbers change as more students and questions are added, but the underlying principles remain:

- Achievement answers come only in discrete, quantum values—correct or incorrect—and answers cannot be subdivided.
- There is an upper-limit and a lower-limit—all correct and all incorrect.
- The chance of being correct or incorrect is calculated by probabilities.

Organization of the Article

This article is divided into eleven sections, as follows:

I. Mathematics of Achievement and Coin Tossing
II. Statistical Interpretations Based on the Normal Curve
III. First Epistemological Interlude
IV. Return to Statistical Interpretations
V. Cost-Effectiveness
VI. Special Circumstances of Statistical Measures
VII. Second Epistemological Interlude
VIII. Attempts to Classify and Synthesize: The Principle of Complementarity
IX. A Theory of School Achievement

The final section presents implications for research, professional training, and practice.

I. The Mathematics of Achievement and Coin Tossing

Achievement testing is an art as well as a science. In test development, there are two potentially conflicting objectives to be balanced. First, tests should reflect the material covered in the instructional process, but second, tests should be constructed to have substantial variation in individual scores in order to distinguish achievement proficiency among students. Ideally, the instructional process would result in all students achieving a perfect score, an indication of effective schooling. This is easily achieved by making the test items extremely easy to answer correctly. In contrast, the test could be constructed to identify those students who can answer questions well beyond the initial instruction, for example, by requiring a synthesis of the presented material. This is also easily achieved by making the items extremely difficult to answer. In the first instance, the distribution is skewed to the right (many achieving high scores), and in the second the distribution is skewed to the left (many achieving low scores). If items were selected so the chance of getting each item correct were 50-50, both objectives would be balanced.

Binomial Distribution and Probability[4]

The early interest in probability was associated with games of chance and flipping coins was a logical starting point. The chances of flipping a head or a tail, is calculated by the coefficients of the binomial expansion $(p + q)^n$ where p is the chance of a head, q the chance of not being a head, and n is the number of coins involved. The descriptive statistics of the binomial expansion are:

- Mean = np;
- Variance = npq;
- p + q = 1.

When flipping coins, p and q equal .5; that is, a 50/50 chance. As the value of n becomes larger, the pattern representing the chances of flipping the number of heads is represented by what is known as "Pascal's Triangle" after the mathematician Blaise Pascal.[5]

(See diagram below.) The probability for each combination is the respective coefficients divided by the sum; therefore, the sum of the probabilities always equals 1.

Pascal's Triangle
(n = 0, 1, 2, 3)

```
        1
      1   1
    1   2   1
  1   3   3   1
```
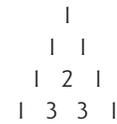
The probability of each number of heads is depicted by a histogram taking the shape of a bell-shaped curve. (See Figure 1.) The sum of the probabilities represented by the bars and the area under the curve equals 1.

Binomial Distribution and Achievement Testing

Achievement testing and coin tossing are similar because of the correct/incorrect heads/tails symmetry. The probability, the value of p for an achievement test, is estimated by what items are included in the achievement test. The mean (np) is the anticipated mean for a student population. The mean is also calculated after the fact when the anticipated and actual means converge as the number of trials increases. Likewise, the variance (npq) is estimated by test construction and confirmed after multiple trials. The anticipated variance is at the maximum at p = .5 where the placement of individual student performance is at a maximum. Changing the value of p, and therefore the mean and variance, has critical impact on the expected outcome of the achievement results. At the extremes, if p is set at 1, all students would be expected to achieve a perfect score; the expected mean would be the parameter n and the variance would be expected to be 0. In contrast, if p were set at 0, the all students would be expected to get all questions incorrect with a mean of 0 and a variance of 0. Figure 2 illustrates the effect of changing p-values.

The geometry of these limits is instructive. As the value of p increases (or decreases), the shape of the distribution changes. When the p-value is .5, the distribution is symmetrical and bell-shaped. As the p-value increases (or decreases), the distribution

Figure 1
Probability of Number of Heads
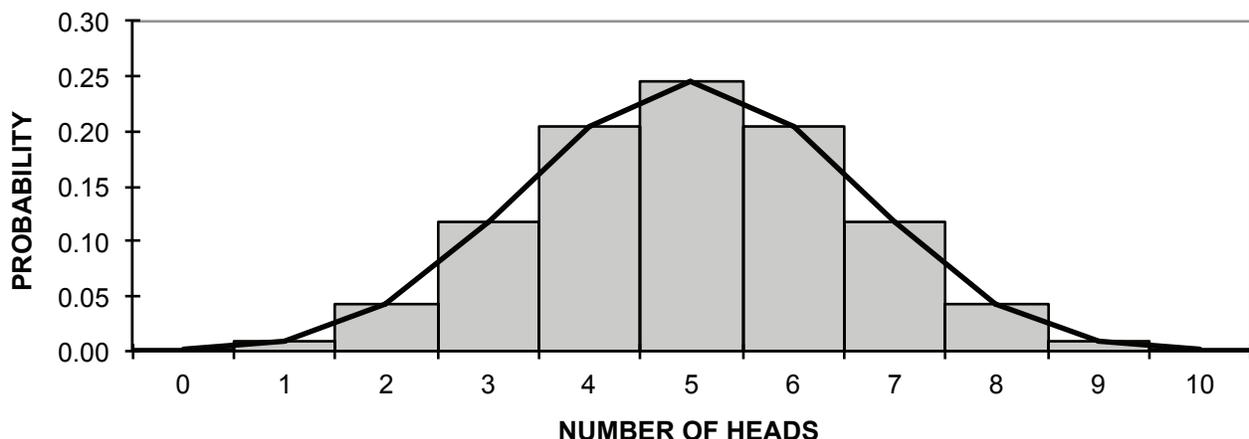
Figure 2
Effect of Changing P-Values on Binomial Statistics
(n=100)
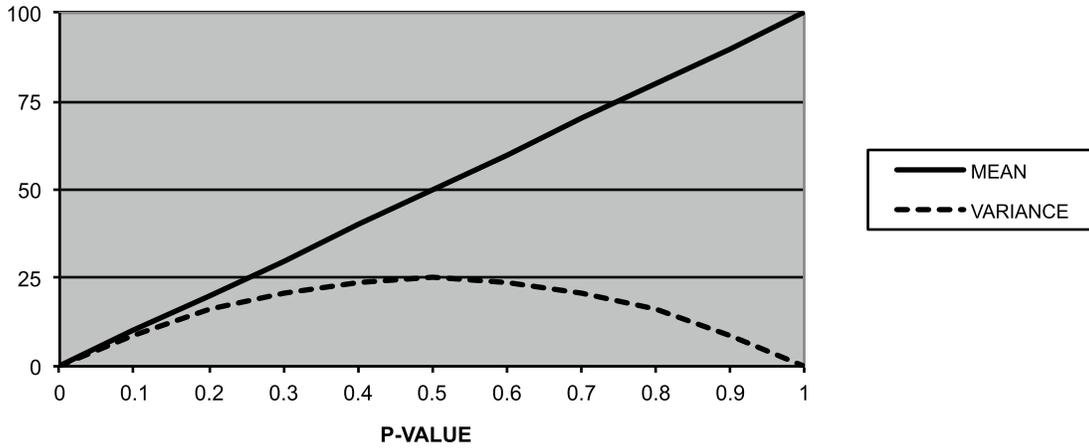
becomes increasingly skewed. The reason is obvious; the upper and lower limits (all correct or all incorrect) prohibit the distribution from remaining bell-shaped. Figure 3 illustrates the change of shape of the distribution and the probability limit as the p-value changes.

Calculating Probabilities: The Normal Curve[6]

At the time of the original inquiry into probability, there were no computers, so doing the calculations for the binomial expansion was tedious. A more practical solution was sought. As more coins were included (n became larger), the histogram resembled a continuous bell-shaped curve. If a mathematical function representing this bell-shaped curve could be developed, the calculations would be easier. One universal curve with an easy method of calculating the probabilities was the goal based on a fundamental probability theorem, as follows:

Probability Theorem: Probabilities are additive with the sum of all possibilities equal to 1. The total area under the curve equals 1, so the area between any two points on the curve equals a probability.

As Newton's and Leibniz's calculus became more sophisticated, a solution emerged. The concept is straightforward although the mathematics is rather sophisticated. Each point of the histogram

(e.g., see Figure 1) is converted to an X-value with the height of the histogram as the Y-value. Once this step is accomplished, two principles of calculus are applied. The first principle is for the intervals on the X-axis, the discrete integers, to become increasingly subdivided (noted as dx).[7] At every X-value, the Y-value (dy) is calculated and (dy/dx) is the slope at that point. The second principle is for the points on the X-axis to be extended in both directions to infinity; that is, an infinite number of coins or questions and for the coins or questions to be infinitely subdivided. There was one more obstacle—how to measure the mean and variance. With the value of n set to infinity, the binomial formula for the mean does not work (infinity times p). In order for the new bell-shaped curve to be universal, a standard measuring convention was developed. When the mean ($\underline{X}$) is set to 0, and the variance (δ^2) set to 1, a universal system emerges. This transformation, $((x - \underline{X})/\delta)$ is now known as a standard score or a Z-score. The calculus notation for these steps is, as follows:[8]

$$dy / y = ((-x - q dx) dx) / \delta^2 + (x + dx) q \, dx$$

As dx approaches 0, this becomes:

$$dy / y = -x \, dx / \delta^2$$

Figure 3
Effect of Changing the P-Value on the Shape of the Distribution
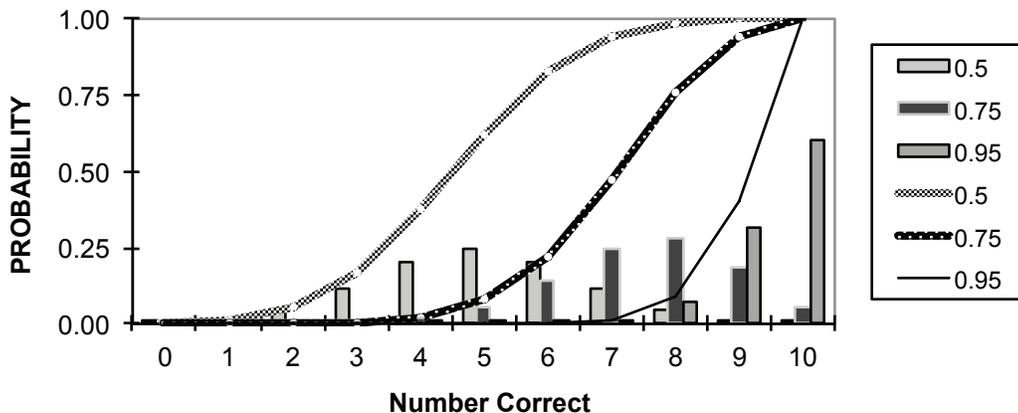
Again relying on calculus, the X-values and Y-values were integrated (summed) from minus infinity to positive infinity.

$\ln(y) = -x^2/(\delta^2) + \text{constant}$
$y = e^{\wedge}(-x^2/(\delta^2) + \text{constant})$
$y = A\, e^{\wedge}(-x^2/(2\delta^2))$

Because the slope of the line is ever-changing, the result is what we now call the normal curve.[9]

The final step is to make the area under the curve equal to 1. With the Z-score as the exponent of the normal probability curve and the area under the normal curve equal to $\sqrt{2\pi}$, the goal is achieved—a universal function describing probabilities. The formulas for the normal probability curve are, as follows:

$$y = \frac{e^{\wedge}-z^2/2}{\sqrt{2\pi}}$$

When Z =0, the mean, the value of y is at the maximum point:

$$y = \frac{1}{\sqrt{2\pi}} = .3989$$

By changing the conditions as required by the calculus, the normal probability curve is not identical to the binomial distribution. Because the normal probability curve extends to infinity in both directions and is continuous (i.e., can be subdivided), any increment can be added to the observations, and while the mean of the distribution will change in the amount of the increment, the variance and Z-scores remain unchanged. The shape of the normal probability curve and the respective probabilities remain unchanged. Figure 4 is a comparison of the cumulative binomial and cumulative normal probability distributions,[10] with the number of questions being 10. In both cases, the area under the curve equals 1. As this number of questions increases, the distributions become closer, becoming practically identical when the number (n) becomes infinite.

The normal probability distribution is a theoretical mathematical construct. It is based on the binomial distribution, another theoretical construct, and not on some natural phenomenon although many distributions in nature are bell-shaped. The purpose of the normal probability distribution is to easily calculate probabilities. Statistical analysis is based on the probabilities determined by this and other mathematical distributions. To repeat, the normal probability distribution and the binomial distribution are founded on different assumptions:

- The binomial distribution is a discrete integer-based histogram while the normal probability distribution is a continuous curve.
- The binomial distribution has upper- and lower-limits while the normal probability distribution extends to infinity in both directions.
- The binomial distribution changes shape if the parameter p (thus the mean and variance) changes, while the shape of the normal probability distribution does not change shape if the parameters (mean and variance) change because it is always measured in Z-scores.

The slope of the curve in Figure 4 is different at every Z-score with the slope approaching but never reaching 0 (and never a cumulative probability of 1) for the normal probability distribution but actually reaching 0 (and a cumulative probability of 1) for the binomial distribution. Above a Z-score of 2, the cumulative probability, the potential gain in probability, and the slope reduce rapidly as demonstrated in the Table below. Clearly, the chance of an observation with a Z-score above 3 is minuscule.

Table
Changes as Z-Scores Increase

Z-Score	2	2.5	3	3.5	4
Cumulative Probability	0.97725	0.99379	0.99865	0.99977	0.99997
Potential Gain in Probability	0.02275	0.00621	0.00135	0.00023	0.00003
Slope	0.05399	0.01753	0.00443	0.00087	0.00013

When applying these findings to school achievement, two postulates can be formulated:

Postulate 1: Every student, classroom, and school has a different probability for increasing or decreasing achievement depending on their previous standing measured in Z-scores.

Figure 4
Comparison of Cumulative Binomial and Normal Probability Distributions
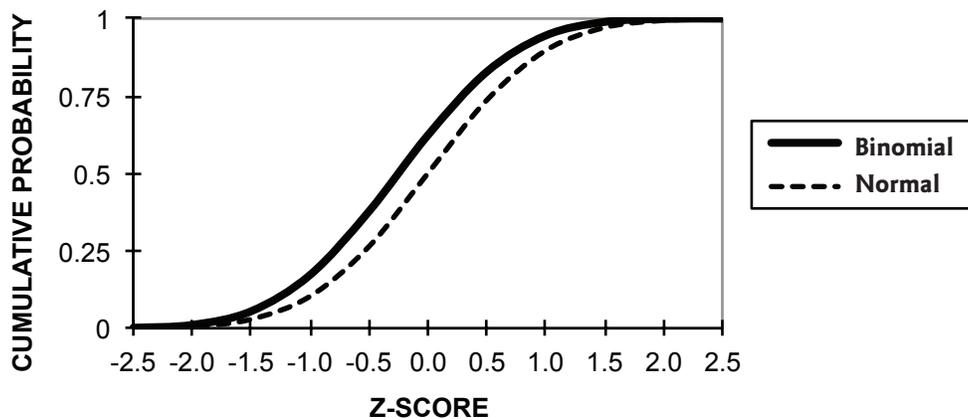
The Probability/Percentile Duality

The probability can be calculated for every expected achievement score measured either as the number or percent of correct answers using the area under the normal curve. In addition, when a test is administered and statistics are calculated on the population, every score can be converted to a percentile ranking, i.e., how a particular score compares to the entire population. Specifically, the cumulative normal probability distribution for any Z-score provides dual information regarding both the percentile ranking (comparative score) and the probability of obtaining the score. The normal probability distribution provides information only about the probability. Figure 5 is a comparison of the two distributions.

For any Z-score, the percentile and probability of the score can be calculated. Above the mean in the cumulative curve where the slope is decreasing, the probability of increasing is less than the probability of decreasing. Below the mean, where the slope is increasing, the relationship is reversed. This is commonly called regression to the mean, indicating that nature tends to prefer the state with the highest probability—the mean.

II. Statistical Interpretations Based on the Normal Curve

The standard or Z-score is the fundamental metric of the normal probability distribution, and it is also the fundamental metric in estimating the magnitude of relationships between achievement and explanatory variables.

Some basic descriptive statistics are required for statistical analysis: the mean—the center point of the distribution; the variance—the area parameter of the normal curve; and the standard deviation (the square root of the variance)—the width parameter. With these three parameters—mean ($\underline{X}$), variance (δ^2), and standard deviation (δ)—the necessary information is present to convert each observation of the distribution (x) into a Z-score by the function: $(x - \underline{X}) / \delta$. As a result, the standard deviation (δ) equals 1, the variance (δ^2) equals 1, and the area under the curve equals 1. The standard deviation is the Z-score unit length on the X-axis, and the variance is the area under the curve (length squared is area).

Linear Interpretation:
Correlation Coefficient and Standard Partial Coefficient

The magnitude of the relationship between achievement and an explanatory variable is frequently called the effect size. For a single explanatory variable, the correlation coefficient (r) represents the magnitude of the relationship. It is the slope of a regression line when achievement and the explanatory variable are measured in Z-scores. It is analogous to Planck's macro-observer based on an average (the average squared distance from the mean, or least squares). The common interpretation of effect size is Newtonian: If the initial position and momentum are known, the future position is known with great certainty. For every increase of one unit in the explanatory variable, the achievement variable is predicted to increase by the value of the effect size.

More frequently there are multiple explanatory variables. Multiple regression analysis accommodates this situation. When explanatory variables are correlated, as usually the case, the correlation coefficients (the various r-values) are not the measure of relationships. A new variable is calculated adjusting the coefficients to compensate for the correlations among the explanatory variables. This adjustment variable is the Beta (β), the standard partial correlation coefficient. It is called standard because all variables are measured in standard or Z-scores, and partial because it accounts for the correlation among the explanatory variables. Partial also means that if one variable changes, the other control variables remain constant. Frequently, these measures are converted from Z-scores back to actual scores, i.e., the number or percent of questions answered correctly.

Figure 5
Normal and Cumulative Normal Probability Distributions
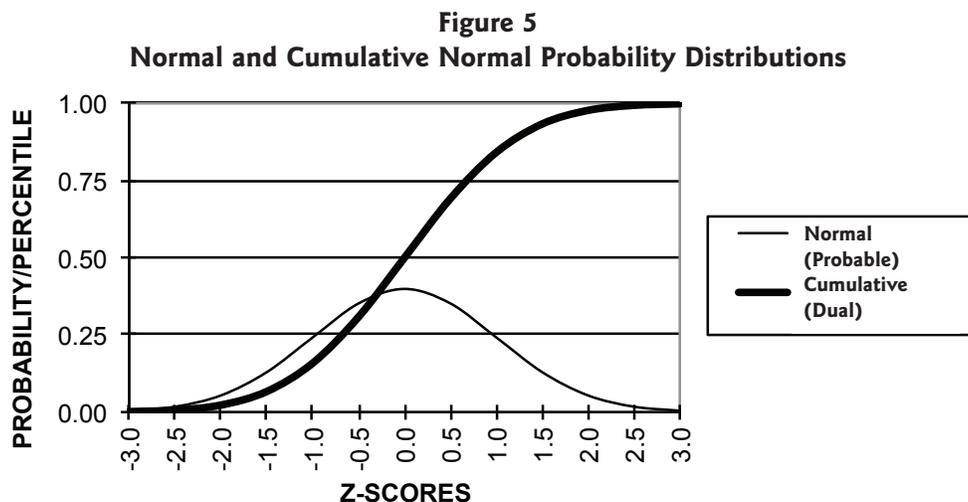
The coefficient β also has a linear relationship with achievement and is commonly interpreted as being reasonably certain. To the contrary, at every point on the regression line, there is a distribution describing a probability range. A more precise interpretation is: For every increase of one unit in the explanatory variable, the achievement variable is predicted to increase within a range defined by the value of β as the average and the standard error of estimate as the probability range. Therefore, for any given Z-score there is no information regarding the unique position of any observation within the distribution. Rather, the position for all observations is considered to be the mean; and no information is provided regarding the probability of any single observation.

> Postulate 5: The correlation coefficient (r) and the standard partial coefficient (β) are measures of average relationships and carry no information regarding the position or probability of any single observation.

Nonlinear Interpretation: Explained Variance

Here, a short review is in order. Variance (δ^2) is the area parameter of the normal curve. Second, the cumulative normal probability distribution represents the sum of the probabilities and is equal to 1; and, third, the probability and percentile ranking can be calculated for any Z-score from the cumulative normal probability distribution.

Regression analysis provides a statistic called the coefficient of determination, the R^2, or the explained variance where:

- The explained variance statistic represents the proportion of area under the normal probability distribution attributable to all the explanatory variables.
- The explained variance for each individual explanatory variable is the product of the basic statistics r and β (r*β); e.g., the variable explains 50% of the variance.
- When the explained variance attributable to each explanatory variable (r*β) is summed, it is the total explained variance or R^2. When added to the unexplained variance, the total is 1.
- The mean of an explanatory variable predicts the mean of the achievement variable; that is, all curves intersect at the mean of the X- and Y-axes.

Figure 6 illustrates the two statistical interpretations: the Newtonian nature of the linear deterministic and the quantum nature of the nonlinear probabilistic. The Y-axis is duality of probability and percentile for the nonlinear interpretation and the percent correct for the linear. To focus full attention on the interpretations, the values of β and R^2 are 1, total prediction.

Distinction between Linear and Nonlinear Interpretations

The discussion has focused on two measurement concepts, predicting a school achievement score and estimating the probability of obtaining a score. The most obvious miscalculation for the linear interpretation is the prediction of 120% and -20% percent achievement at the extreme Z-scores. In addition, the liner interpretation provides no information regarding the probability for any individual school. Because of the percentile/probability duality, the nonlinear interpretation provides information regarding the predicted score (in percentiles) and the probability of obtaining the score because there is a unique slope associated with every school's Z-score. In the linear interpretation, the initial position has no impact on the magnitude of increase because the increase will be the same for all observations. With the nonlinear interpretation, the initial position is critical for it has a direct impact on the magnitude and probability of the increase. A graph of a learning curve is so similar to the probability/percentile curve its inclusion would not be instructive. However, the existence of such a learning curve adds credence to the nonlinear interpretation of achievement.[11]

Diminishing Returns: A Thought Experiment

Diminishing returns is a fundamental principle in many disciplines, such as economics and business: As an input increases beyond a certain point, the rate of increase of the output gradually decreases. In order to illustrate the principle, it is not necessary to collect and analyze data; rather, a thought experiment suffices. Assume a study concluded that the number of available textbooks had a relationship with achievement. Remember, the number of books has no direct relationship with achievement; instead, it is more related to the amount of time the books are in use. One book for 50 students produces one level of achievement, two books a higher level, and, as the number of books continued upwards, so would achievement. At what point is the diminishing returns reached? If

Figure 6
Graphic Representation of Linear and Nonlinear Interpretations
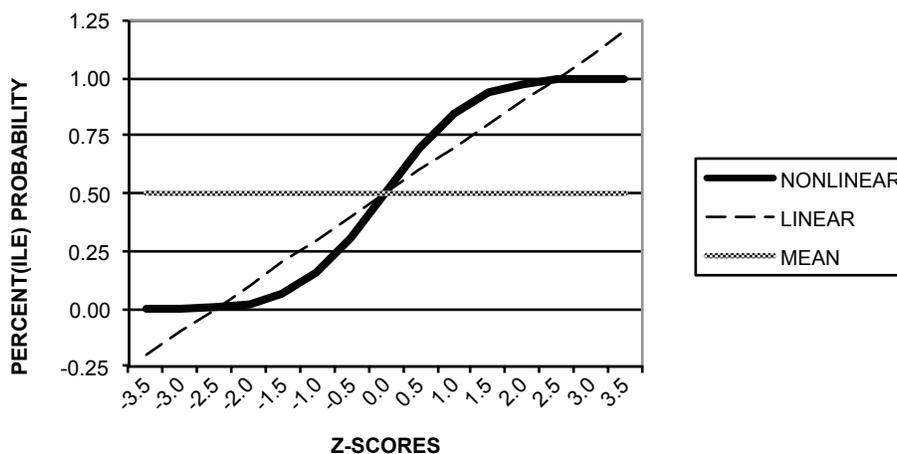
the number of books is divided by the number of students, the sequence of fractions give some idea of the answer: 1/50, 2/50, etc. At 25/50 or one book for every two students, it is feasible for students to share. The investment to double the number of books so every student had their own book would not double the achievement. More than one book per student would be illogical. There is a point where common sense concludes a reasonable point has been reached. If a thought experiment results in diminishing returns, then the obvious conclusion is that the mathematical function is nonlinear. In the case of learning, the principle of diminishing returns is a part of the learning curve.

The Nature of Achievement—Deterministic or Nondeterministic: A Thought Experiment

Assume 11 students take a ten-question test and they score 0 to 10 respectively, for an average of 5. Through some intervention, the average score is predicted to increase to 6. What will be the new scores? The student with 10 correct must stay at 10, while the student scoring 9 correct would move to 10. The rest must move up an average of 1.11 in order for the new average to be 6. How does the student who achieved the perfect score know they will get the same score? They don't. The probability might be high, but they cannot be sure due to regression to the mean. How does the student who achieved 9 correct know they will improve by 1 and not 1.11? How do the rest of the students know the scores of the top two students so they can improve their performance the exact amount to raise the average to 6? They cannot. There is no way, short of cheating, that the students can know their future score and how much they must improve in order for the results to exactly equal the predicted value. In contrast, according to Galileo, objects know exactly at what velocity to fall. According to Newton, planets know exactly their path through the sky and the tides know exactly when to shift. According to Einstein, light knows exactly how to travel through space-time. Einstein called this "spooky action at a distance" because gravity determines exactly how all objects behave (Cox and Forshaw 2011, 140). There is no "spooky action at a distance" determining how students answer questions; there is only the probability of how they might answer.

Further assume that the intervention was a reduction in class size from 25 to 20 students. Surely achievement scores would not increase immediately when five students leave the room (although the average might change). For there to be an improvement in achievement for the remaining 20 students, there must be a change in behavior by the teacher and the students; after all, achievement can only be improved by a change in behavior.

The thought experiments can be classified into either of two mathematical functions: (1) Linear, continuous returns, and deterministic; or (2) nonlinear, diminishing returns, and probabilistic, as follows:

(1) Linear Achievement = $\beta f(z)$, where β is the coefficient and $f(z)$ is the linear achievement function.
(2) Nonlinear Achievement = $R^2 f(z)$, where R^2 is the explained variance and $f(z)$ is the probability/percentile duality function.

Postulate 6: The nonlinear interpretation gives accurate information regarding maximum and minimum scores and provides information regarding probability. The linear interpretation gives inaccurate information regarding maximum and minimum scores and provides no information regarding probability.

Postulate 7: Because of the duality of the nonlinear interpretation—percentile and probability—as the Z-score moves to either side of 0 (the mean), the returns to scale become increasingly smaller.

Postulate 8: The probability of achievement change is predicated on the likelihood of a change in behaviors.

Postulate 9: The initial condition with the nonlinear function is crucial in determining the magnitude and probability of change.

III. First Epistemological Interlude

The interpretation of Figure 6 prompts an epistemological discussion, as suggested by Heisenberg (2011), regarding the purpose of knowledge and how an understanding of reality influences the interpretation. After the experiments and analysis revealed the structure of the atom, there was a difference of opinion regarding the underlying interpretation of quantum theory. The research evidence and the mathematical proof by Heisenberg of an uncertainty principle supported a nondeterministic, probabilistic interpretation, and Bohr (2011), one of the originators of the theory, was an ardent advocate. Bohr based his thinking on two arguments: (1) The interpretation should only be concerned with what is actually observed and measured, in other words, reality; and (2) the interpretation should favor the mathematical function containing "all the possible information" (Hawking 2011, 445).

Einstein, who wrote one of the seminal papers leading to the quantum movement and his Nobel Prize, agreed with the experimental findings and mathematics, but could not agree with the nondeterministic, probabilistic interpretation (Einstein, Podolsky, and Rosen 2011). He replied to Bohr with the now famous quote, "God does not play dice," arguing for a deterministic interpretation consistent with his theory of relativity, for which he did not receive a second prize (Cox and Forshaw 2009, 190). He could not give an alternative explanation only to say a yet undiscovered variable was missing to make the explanation deterministic (Einstein, Podolsky, and Rosen 2011). Bohr and Einstein exchanged a series of papers trying to convince the other their interpretation was correct. Focusing on the importance of accurately representing reality, Bohr (2011, 471) wrote: "The extent to which an unambiguous meaning can be attributed to such an expression as 'physical reality' cannot of course be deduced from *a priori* philosophical conceptions, but must be founded on a direct appeal to experiments and measurements."

In essence, Bohr was telling Einstein that it is not what you believe, it is what experiments and mathematics tell you. Einstein, in turn, was saying, he knew that the experiments and mathematics were correct, but he still couldn't believe them, that something was missing. To the issue at hand, the mathematics and logic presented

above weigh in favor of the nonlinear percentile/probability interpretation because it provides accurate information regarding the reality of both achievement limits and the probability of obtaining specific levels of achievement—all the possible information. In contrast, the linear interpretation provides inaccurate information regarding achievement limits and no information regarding probability, thus founded more on beliefs.

IV. Return to Statistical Interpretations

Linear Interpretation of Multiple Explanatory Variables

The Beta (β) coefficient is the common multiple regression statistic. When multiple variables are included in an analysis, the linear and implied deterministic interpretation represents a theory of substitution; that is, any variable can substitute for any another in order to attain an achievement target. This is because the position on the regression line makes no difference in the prediction since the slope is the same for all schools. The difference is the amount of the increase necessary in the explanatory variable to reach the target. This is evident in Figure 7.

> Postulate 10: Because the linear interpretation is based on the Beta's—partial correlation—all explanatory variables cannot move simultaneously; the Z-score of one variable may move while the Z-scores of the others must remain unchanged. If all variables move simultaneously, the limit would be reached sooner.

Nonlinear Interpretation of Multiple Explanatory Variables

When predicting achievement with the combination of explanatory variables, the explained variance is consistent with the quantum nature of achievement—probability/percentile measures with limits. The explained variance is calculated by summing the product ($r*\beta$) for the variables, not by summing the r-values or the β-values.

> Postulate 11: When dealing with multiple explanatory variables, the respective variances ($r*\beta$) can be added with the sum being the explained variance (R^2); the explained variance plus the unexplained variance equals 1.

The normal probability curve can be subdivided, with each subdivision attributable to a single explanatory variable and measured as the percentage of the area under the curve. Percentage of area under the curve is equivalent to the percentile/probability duality. Hence, the R^2 is the percentile range on the normal probability curve accounting for or explained by a combination of explanatory variables. The R^2 of several hypothetical explanatory variables is illustrated in Figure 8. Because the mean of the explanatory variable predicts the mean of the achievement variable, all curves intersect at a Z-score of 0, the 50th percentile. When viewed as probabilities, it demonstrates the principle of regression to the mean; that is, the probability of moving to the mean is greater than moving to the extremes. The R^2, it should be emphasized, is built on a non-substitution theory. No input can be substituted for another because the position on the curve for every explanatory variable is unique for every student, classroom, and school.

Comparison of Statistical Interpretations

The two preceding figures represent equations. There are two solutions to the linear equation: (1) If all schools were at the mean (Z-score of 0), all schools would be predicted to achieve at the mean; and (2) If every school invested unlimited resources into every variable, all students in all schools would have better than perfect achievement scores. There is one universal solution for all schools because every school is assigned the same linear coefficients. These are misleading solutions because the interpretation does not represent a common understanding of reality. With the nonlinear interpretation, if a school were at the mean, the achievement results would be at the mean--the same as the linear interpretation. More importantly, because there is a unique position (Z-score) on every variable for every school, there would be a unique allocation of resources among the variables in order to achieve the best possible increase in achievement. Again, the interpretation depends on the perception of reality, i.e., best possible achievement or better than perfect achievement.

> Postulate 12: When each explanatory variable is measured by the variance ($r*\beta$), each variable represents the unique contribution to the total explanation of achievement.

> Postulate 13: Because each variable is based on the variance ($r*\beta$), the Z-score of every variable may move without ever exceeding a percentile limit.

Figure 7
Beta Coefficients with Various Slopes Moving Upwards from 0 to .5
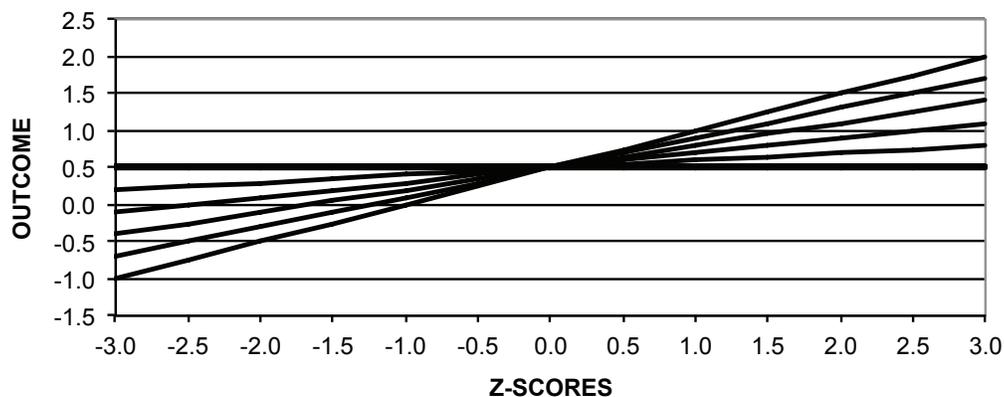
Figure 8
Explained Variance with Various Variances, Moving Upward from 0 to 1.0
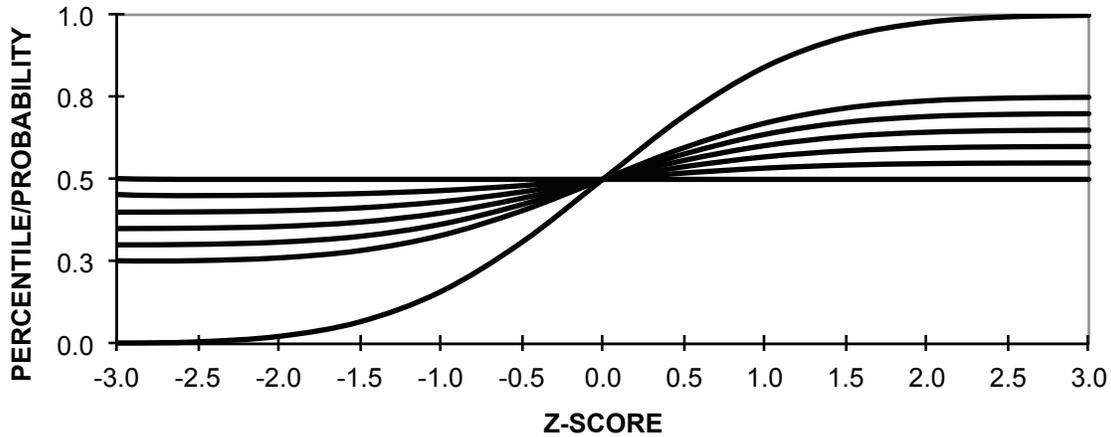
V. Cost-Effectiveness

Financial cost is a major consideration when making policy decisions.[12] An adjustment can be made to the effect size in order to compare the cost-effectiveness of various explanatory variables. The cost-effect-size is calculated by dividing the effect size by the cost of increasing the explanatory variable Z-score by one unit. In essence, cost is equally important as the effect size when considering the impact on achievement. If the unit cost of one variable was one-quarter the cost of another, the effect size of the second variable must be four times as large for the two variables to be equally cost-effective. Figure 9 illustrates cost-effectiveness curves for various effect sizes (in R^2). Of note is the following:

- The unit cost is per Z-score; the range is $\pm$ 3, the practical maximum and minimum.
- The "Percentile per $" is based on one dollar per Z-score. If the unit cost increases, the percentile per $ metric decreases proportionally.
- The maximum of the cost-effective curve is at a Z-score of 1.13 or 4.13 units of cost. At this point, .688 of the total funds (practical maximum at Z = 3) will yield .869 of the potential achievement.
- While predicted achievement will continue to increase with additional funding, it will be at a reduced rate.

When cost-effectiveness is considered, the difference between the linear and nonlinear interpretations becomes even more striking. Once the most cost-effective variable is identified for the linear interpretation, there is every reason to invest all available funds into that single variable. In contrast, the nonlinear interpretation provides a thought-provoking alternative: Funding continues linearly; but the effect size is nonlinear. So there is a point of maximum cost-effectiveness for every variable. The sensible goal is to pursue the maximum cost-effective point on all variables.

> Postulate 14: With the nonlinear micro-interpretation, there is a unique cost-effectiveness curve for every explanatory variable and a unique position (Z-score) on the curve for every school. Therefore there is a unique and optimal solution to the allocation of financial resources to achieve the optimal level of achievement for each school.
>
> Corollary: With the linear macro-interpretation there is only one most cost-effective variable applicable to all schools—one universal solution.

VI. Special Circumstances of Statistical Measures

There are special circumstances influencing the uncertainty of statistical measurement, such as a lack of clear definitions,

Figure 9
Cost-Effectiveness Moving Upwards from .05 to .20
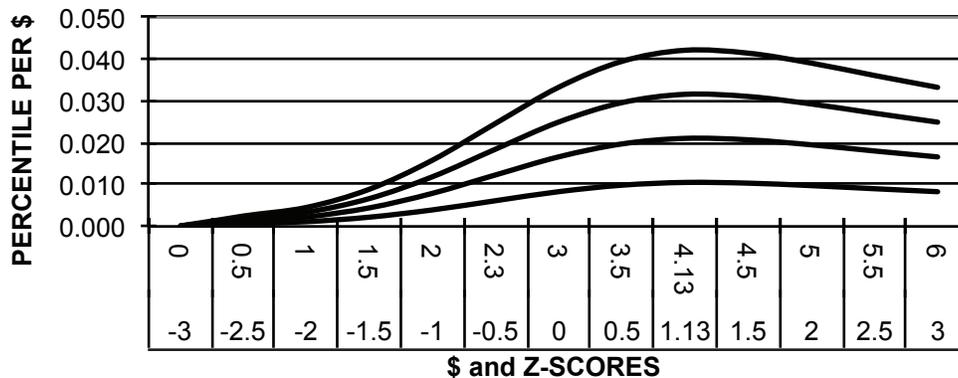
unavailability of data, and substantial correlation among explanatory variables. These issues substantially determine the accuracy of predicted achievement and the coherence of an explanation.

Socioeconomic Status (SES)

Previous research has demonstrated the substantial influence SES has in predicting achievement, so it must be included in a school achievement theory and prediction model. The prediction model is:

$$A_i = \beta * SES_i + e \ \ or \ A_i = (\beta*r) * SES_i + e$$

Where A = achievement, A_i and SES_i = individual schools, e = error, and the coefficient β applies equally to the variance (β*r). The prediction must use the same variables (Achievement and SES) that were used in the analysis to determine the weighting (β). Because the achievement prediction is made for a future event, the best estimate of the true value of β(*r) is the average of previous events (Taylor 1982, 117). Therefore, the prediction of future achievement must meet four conditions:

(1) Achievement and SES must be defined and measured consistently over time.

(2) Because achievement and SES are defined and measured consistently over time, the coefficient β(*r) is a constant, invariant over time.

(3) If the definition and measure of achievement (A) changes (e.g., from reading to mathematics), then β(*r) will change because the definition and measure of SES remains consistent.

(4) The coefficient β(*r) is selected to maximize the prediction of achievement by SES.

Socioeconomic status requires special consideration when analyzing school achievement because no universal definition exists, so no single data variable exists. Instead, a single index number representing SES must be constructed from available data serving as proxy variables. The proxy variables for SES are generally comprised of student, family, and community characteristics, which are usually substantially correlated. SES proxy variables sometimes include education and income levels but, in the context of school achievement, it does not follow that a student's achievement will automatically increase when family income increases or parents graduate. More likely, families with higher education and income levels, or any of the proxy variables, encourage a set of behaviors related to achievement, but the desired behaviors are not absolutely determined by these measures. The behaviors can be fostered to some degree anywhere. Unfortunately, these behaviors are not well defined nor are data available. Researchers do their best to collect proxy variables representing student, family, and community behavioral traits.

> Postulate 15: SES is a combination of proxy variables representing unobserved student, family, and community behavioral traits.

After potential proxy variables are identified, there is another consideration: How to select the final variables and weightings. In essence, how do we define and measure SES? The revised prediction model is:

$$SES_i = V_{1i} * \underline{W}_1 + V_{2i} * \underline{W}_2 +... \qquad (Equation \ 1)$$

$$A_i = \beta(*r) * (V_{1i} * \underline{W}_1 + V_{2i} * \underline{W}_2 +...) + e \qquad (Equation \ 2)$$

Where
V = proxy variable
and
$\underline{W}$ = average weighting

The same conditions apply to the proxy variables and weightings, as follows:

- The variables (V) and weightings ($\underline{W}$-values) must be invariant over time.
- The sum of the terms (V * $\underline{W}$) represents SES (equation 1) and must be defined and measured consistently over time.
- The variables (V) and weightings ($\underline{W}$) must also be consistent across achievement measures, averaged weightings over time and across achievement measures.
- The variables and weightings (V * $\underline{W}$) are selected to maximize β(*r) so that the prediction of every achievement measure is maximized.

There is no unambiguous method to divide the shared variance among the correlated proxy variables. Because the correlated proxy variables all contribute to the same behavioral trait, the proxy variables are combined into a single number index. This is a fundamental principle of factor theory. Establishing a factor is consistent with equations 1 and 2:

$$SES_i \ Factor = V_{1i} * \underline{W}_1 + V_{2i} * \underline{W}_2 +...+ e$$

$$A_i = \beta(*r) * V_{1i} * \underline{W}_1 + V_{2i} * \underline{W}_2 +...+ e$$

> Postulate 16: SES should be constructed as a factor with the same variables with weightings averaged across achievement measures and over time.

> Postulate 17: SES cannot be defined and measured at the same time the relationship between achievement and SES is measured. Two complementary analyses are required.[13]

> Postulate 18: Measuring the relationship between achievement and explanatory variables depends sub-stantially on how well SES is measured. A larger relationship between achievement and SES will tend to increase the relationship between achievement and the other explanatory variables (Phelps 2011c).

> Postulate 19: Because the definition of SES and the available data vary due to state data collection, measurements of SES are unique to states.

Other Factors

Phelps and Addonizio (2006) applied the above method to the SES proxy variables and formed an SES factor, but this method was not applied to other groups of statistically and conceptually related variables such as staff characteristics (experience, training, age, salary) or staff roles (teachers, instructional support staff, teachers aides, administrators). Because of small changes in the correlation matrix, there were chaotic results for these explanatory variables across years.[14] The results were confusing and impossible to explain. Surely, the various staff characteristics work together rather than

separately to influence staff behavior just as the various SES proxy variables work together to influence student behavior. Similarly, the various staffing roles work together as a team to influence student achievement. In a later study when the factor principle was applied to develop single number indices for staff characteristics and staff roles, the confusion disappeared, and the results were easily explained (Phelps 2009).

Factor analysis is a valuable tool in cases where conceptually- and statistically-related variables occur. There are three components of statistical variance: (1) common or shared by many variables; (2) unique, present in only one variable; and (3) error. Factor analysis groups explanatory variables sharing common variance. Because there is no unambiguous method to partition the shared variance among correlated variables, a reasonable solution is to combine the related variables into a factor, a single number index representing the concept and the explanatory influence of the entire group (Phelps 2011e).

> Postulate 20: When explanatory variables are conceptually and statistically related, combining them into factors produces a more coherent explanation and avoids chaotic statistical results.

Effective Use of Resources and Measurement of Unobserved Variables: A Thought Experiment

The uncertainty of measurement, i.e., the uncertainty of how human and financial resources are transformed into achievement by a school, is the major reason why achievement is better defined by probabilities. Assuming a statistical analysis predicts future achievement of three schools with reasonable accuracy, this thought experiment follows the results of the predictions over several years. The results are analyzed to determine how closely predicted achievement compares with actual achievement. The hypothetical results are: (1) The average actual achievement of one school was significantly higher than the average predicted achievement; (2) The average achievement of the second school was almost exactly what was predicted; and (3) The average achievement of the third school was significantly lower than what was predicted. There are two possible conclusions: The differences are entirely due to random measurement error, or something unobserved has systematically taken place in each school having a substantial influence on achievement levels. The latter explanation is what economists call the fixed or school effect and can be considered as a measure of effectiveness, i.e., how human and financial resources are transformed into achievement (Wooldridge 2000). The fixed or school effect is obtained by averaging the residuals over time and is described in many econometric textbooks.[15, 16] In the three-school hypothetical, one school was effective, a second was neutral, and the third was ineffective. The reason for the level of effectiveness cannot be due to any of the variables included in the original prediction. The difference is likely due to organizational behaviors (Levin 1997). The magnitude of the school effect is substantial (Phelps 2009).

> Postulate 21: It is possible to estimate the influence of unobserved variables on achievement by the econometric technique of fixed or school effect. The school effect factor represents the school's unique operational behavioral characteristics.

VII. Second Epistemological Interlude

Once more the underlying question is: What is reality? The quantum view starts with the nature of the atom, from the Greek word *atomos*, meaning indivisible, or the smallest piece,[17] but acknowledging that the atom is a component of something larger. In chemistry, organic elements bond into acids and then into DNA. Achievement test construction combines individual skills like addition, subtraction, division, and multiplication, into something called numeration. Psychology combines individual abilities and preferences into traits or characteristics. The elementary building blocks of most phenomena are combined into larger concepts. Are the explanatory variables of school achievement somehow different and hence cannot or should not be combined? A case can be made against combining only if the explanatory variables were conceptually and statistically unique. Then a single conceptual and statistically unique variable would be a factor. Regarding the reality of schools, the presumption is that schools have distinctive characteristics or traits which can be identified and measured as combinations of variables—factors. Guilford (1965, 470) addressed this point, as follows:

> It is usually easy enough to apply a measuring instrument and to obtain some numerical data. In the physical sciences the meaning of the numbers that are used to describe phenomena is usually well established... In the behavioral sciences, however, the connection between a number and the thing, or things, for which it stands is not nearly so obvious.

In the social sciences, the thing or things are measures of individual or group characteristics or traits. Because schools are comprised of people, the behavioral trait concept is more compelling than the object notion associated with the physical sciences. In the case of schools, the factors are best considered as measures of organizational traits whereby each school has its own personality, chemistry, or DNA. In the final analysis, it is not the number of objects that deterministically cause achievement; rather, it is the traits, what the numbers represent, that influence the probability of success. The final observation of Guilford (1965, 480) is instructive: "On the whole, there is much more to be gained in increasing the R^2 by discovery or identification of new factors than there is by increasing the loadings [weightings] for already known factors."[18]

VIII. Attempts to Classify and Synthesize: The Principle of Complementarity

According to Heisenberg (1965,155), "The solid ground of experimental proof has often been forsaken, and generalizations have been accepted uncritically, until finally contradictions between theory and experiment have become apparent."

Several efforts to classify and synthesize previous school achievement studies were surveyed in "A Practical Method of Policy Analysis by Considering Productivity-Related Research" (Phelps, 2011b). When possible, the results were converted into a consistent effect size measure, the amount of explained variance or R^2 (Phelps 2011c).

Below is a brief summary:

- A 1978 analysis of class size by Glass and Smith. Their conclusion was represented by a curve predicting increasingly higher achievement as class size decreases smaller than about 15. The review revealed errors in data preparation, application of statistics, and the application of mathematics.

After errors were corrected, a reanalysis produced results completely at odds with their conclusions and inconsistent with any notion of reality. Even so, their assumption regarding nonlinear relationships is valuable.

- <u>A 1994 study by Hedges, Laine, and Greenwald using several explanatory variables including funding.</u> Although they found no statistical evidence regarding the relationship between achievement and common explanatory variables, they found a relationship between achievement and per pupil expenditures. Their explanation was that local school officials make the appropriate decisions to produce increased achievement outcomes. No evidence was provided to support this point. However, their assumption directed interest to the notion of school effectiveness.
- <u>Walberg's 1984 study of explanatory variables in the categories of instruction, curriculum, organization, homework, and time.</u> He made several estimates of effect size; however, when taken together, the estimates were unrealistically high. Still, attention to these categories as a part of a theory is valuable.

Even with a small representation of the multitude of studies, there is substantial reason to conclude the following:

- Attention is paid mostly to the relationship between achievement measures and individual explanatory variables rather than a comprehensive consideration of multiple achievement measures and factors.
- There is no standard method of measuring effect size.
- There is no systematic inclusion of SES.
- Including a measure of individual school effectiveness is entering the research literature, but usually not as a part of a comprehensive description and explanation of school achievement.
- There is little evidence of a comprehensive theory evolving from findings of previous studies.

The Principle of Complementarity

Bohr's principle of complementarity (Born 2011, 460) is described by the following historical timeline of quantum mechanics:[19]

- 1899: Planck explained that there is a fundamental unit of energy within the atom with an integer value called "quantum."
- 1905: Einstein, building upon Planck's work, explained why electronic current is produced when light strikes metals.
- 1909: Planck summarized the knowledge gained up to that point in a lecture titled, "The Atomic Theory of Matter."
- 1911: Rutherford and Geiger concluded that the atom was comprised of electrons orbiting around a nucleus.
- 1913: Bohr concluded the orbits around the nucleus were stable, consistent with Planck's notion of quantum.
- 1927: Wilson demonstrated that atomic particles behave as particles.
- 1928: Davisson and Germer demonstrated that the atomic particles (electrons and photons) behave as waves.
- 1927. Heisenberg (2011, 164) established the uncertainty principle, stating: "It can be expressed in its simplest form as follows: One can never know with perfect accuracy both of those two important factors which determine the movement of one of the smallest particles—its position and

its velocity. It is impossible to determine accurately both the position and the direction and speed of a particle at the same instant."
- Heisenberg (1930) in "The Physical Principles of the Quantum Theory," explained the wave/particle duality of light (photon) and the electron. This sequence of building one concept on another, each making a complementary contribution, continues today—all because of the original idea of Planck's quantum. To summarize the principle, Born (2011, 460) observed: "There exist, therefore, mutually exclusive though complementary experiments which only as a whole embrace everything which can be experienced with regard to an object."

With ever-changing definitions, variables, metrics, and results in school achievement research, there is no capacity to classify and synthesize based on the principle of complementarity. Moreover, without an initial theory, there is no conceptualization against which to evaluate complementary studies. With a conceptualization—a theory—individual experiments can be conducted with the results entered back into the theory to evaluate their contribution. We return to Heisenberg (2011, 155) quote: "Difficulties arise only in the attempt to classify and synthesize the results, to establish the relation of the cause and effect between them—in short, to construct a theory."

IX. A Theory of School Achievement

Whether the focus is a planet, electron, or individual school, the purposes of research coincide to comprehensively describe and coherently explain the phenomenon via a set of laws (mathematical principles), and to accurately predict the future. The first task is to describe the initial position of the planet, electron, or the level of school achievement. The second is to explain what causes the position of the planet, electron, or the level of achievement to change. Third is to accurately predict where the planet, electron, or achievement level will be in the future. The comprehensive description, coherent explanation, laws, and accurate prediction comprise a theory.

The proposed achievement theory is *a posteriori* in nature, patterned after Galileo, Newton, Einstein, and Planck. Assumptions and conclusions are understood to be valid elements of a theory because of prior observations, experiments, and analyses, but they are confirmed only when the predictions derived from the theory are verified experimentally. Education theory, in contrast, tends to be *a priori* in nature; that is, assumptions and conclusions are evaluated via research and deductive reasoning, but no school-specific predictions are made, so verification is impossible.

The theory proposed here centers on one paramount proposition: School policies, as represented by the factors, are directed toward the educational behaviors of students, staff, families, and communities; and the combination of behavioral characteristics creates the achievement environment. In the simplest of terms, effective school policies have a positive influence on student, staff, family, and community behaviors, and these behaviors have a positive influence on student achievement. In essence, the allocation and direction of human and financial resources is the DNA of school achievement (Phelps 2011e). This theory and model are based on four propositions:

(1) It is beneficial to have a comprehensive theory and mathematical model to guide research, professional training, and practice.
(2) The theory and mathematical model apply to individual schools, the quantum microview.
(3) Complementary studies based on the average of many schools, the macroview, are vital in estimating relationships among various achievement measures and factors.
(4) A comprehensive theory and mathematical model can only be developed, solved, and verified using the quantum microview based on the percentile/probability duality and the behavioral characteristic represented by the factors.

Fundamental Laws of School Achievement

The theory and model of school achievement are based on the paramount proposition that each factor represents a behavioral trait ($f(z)$ = Behavioral Trait) and eight fundamental mathematical laws derived from the previous postulates. These laws are as follows:
(1) The sum of the weighted factors plus error equals predicted achievement: $\Sigma\ R^2 f(z) + e = PA$
(2) A factor weighting equals the product of the correlation coefficient and the standard regression coefficient: $R^2 = \Sigma\ (r * \beta)$
(3) The sum of the factor probabilities plus error equals 1: $\Sigma\ pf(z) + e = 1$
(4) Probability range (p) equals the coefficient of determination: $(R^2)p = R^2$
(5) Factors can be synthesized from individual variables with invariant weightings: $f(z) = \Sigma\ (V * \underline{W})$
(6) Individual factors are conceptually and statistically unrelated: $f(z) \neq f(z)$
(7) The difference between the averaged actual and predicted achievement is school effectiveness, an unobserved factor: $\underline{AA} - \underline{PA}$ = School Effectiveness
(8) The factor weighting of each factor divided by cost is a measure of cost effectiveness: $R^2 f(z)\ /\ \$$ = Cost-Effectiveness

From these laws evolves a comprehensive theory of school achievement whereby the status and progress of school achievement can only be described, explained, and predicted by utilizing the estimated relationships between multiple achievement measures and multiple factors (after Bohr's "reality," defined as what can be observed and measured). There is an optimal level of multiple school achievement measures that can only be predicted by identifying the optimal levels of the factor Z-values constrained by a maximum level of expenditures (principle of cost-effectiveness). The optimal factor Z-values are determined by solving simultaneous equations with parameters unique to each individual school (the quantum microview).

For the purpose of the theory and model, the following assumptions describe school operations. Schools operate:
• To achieve multiple identifiable and measurable achievement goals.
• Within a system of identifiable and measurable endogenous policy options (factors) designed to achieve the specified educational goals.
• Within a system of identifiable and measurable exogenous factors only partly responsive to school policies that influence the specified educational goals.

• Under identifiable and measurable cost constraints.
• Under practical constraints other than cost, which can be identified and measured.

Measurement requirements of the quantum microview are derived from the postulates in this article, as follows:
• All elements, achievement measures, and factors must be measured in Z-scores.
• Definitions and measures of achievement must be consistent over time.
• Definitions and measures of the endogenous and exogenous factors must be consistent across achievement measures and time.
• The relationship between achievement measures and explanatory factors must be measured by the percentile/probability duality, $R^2 f(z)$.
• Statistically-correlated and conceptually-related variables must be combined into factors, a single index representing their combined variance.
• SES must be included as exogenous factor.
• School effectiveness, the school effect, must be included as endogenous factor.
• Other endogenous factors are likely to include, staffing roles, staffing characteristics, instructional materials, methods of instruction, curriculum, time, or any measurable variables with either a distribution or a yes/no, as long as there is reasonable evidence as to the magnitude of the effect size and cost.

X. A Mathematical Model of School Achievement

Heisenberg (2011, 162) stated: "It is not surprising that our language should be incapable of describing the processes occurring within [education], for...it has been invented to describe the experiences of daily life.... Fortunately, mathematics is not subject to this limitation, and it [may be] possible to invent a mathematical scheme...adequate for the treatment of the [educational] processes."[20] The school achievement process can be mathematically modeled by a set of simultaneous equations with a separate equation for each desired achievement outcome and an equation representing the cost of each factor. There is a unique solution to these simultaneous equations representing the unique structure and circumstances of each school, the microview. As a result, alternative policy strategies can be identified and tested via simulation. A solution is possible because of the nonlinear cost-effectiveness principle explained previously. The method is to select the optimal level for each factor producing the optimal level of the multiple achievements measures, given a specified cost constraint. Other operational constraints may be included in the model. Under the macroview, no system of simultaneous equations can be constructed and solved because of the linear and unlimited returns for every explanatory variable.

The model is divided into four phases:
• Phase 1
  ➤ Determine the initial achievement level:
  ➤ Maximize the achievement predictions by identifying the best fitting factors and factor weightings.
  ➤ Factors should reflect behaviors and not just the allocation of resources.

Figure 10
Model Structure

Phase 1: Maximizing Predictions. $f(z)$ is school-specific for all achievement measure within each factor.											
Factors											
Predicted Achievement Measures	**SES**	**Effectiveness**	**Staff Quantity**	**Staff Characteristics**	**Instructional Materials**	**Instruction**	**Curriculum**	**Time**	**Other**	**Predicted Row Sum**	**Sum of Factor R²**
1 =	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	Σ	ΣR^2
2 =	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	Σ	ΣR^2
N =	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	Σ	ΣR^2
Factor Sum	Σ	Σ	Σ	Σ	Σ	Σ	Σ	Σ	Σ	ΣΣ	
Factor Cost	n/a	n/a	$	$	$	$	$	$	$	Σ $	

Phase 2: Maximizing Predictions. Optimizing Predictions. $f(z)$ is the same for all achievement measure within each optimized factor.											
Predicted Achievement with Optimized Z	Fixed Z	Fixed Z	Optimized Z	Optimized Z	Optimized Z	Optimized Z	Optimized Z	Optimized Z	Optimized Z	**Optimized Row Sum**	
1 =	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	Σ	
2 =	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	Σ	
N =	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	$R^2 f(z)$	Σ	
Optimized Factor Sum	Σ	Σ	Σ	Σ	Σ	Σ	Σ	Σ	Σ	ΣΣ	**Target Cost**
Optiized Cost	n/a	n/a	$	$	$	$	$	$	$	Σ $	**$$**
Change in Achievement	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	ΣΔ	

- Phase 2
  - ➤ Report status and progress of individual schools.
  - ➤ Once the factors are established and school data gathered and analyzed, there is great value in reporting the information to policymakers, practitioners, and the public.
- Phase 3
  - ➤ Predict new achievement levels.
  - ➤ Optimize the predictions of future achievement using the factors and weightings from phase 1.
  - ➤ Selecting the optimal Z-score for each factor for the individual school, with the Z-score levels constrained by cost, identifies the optimal achievement predictions.
- Phase 4
  - ➤ Verify the prediction of new achievement levels.
  - ➤ After the simulation model is established, the school parameters gathered and entered into the model, and the policy alternatives evaluated, it is critical to test the simulation predictions via natural experiment.
  - ➤ If the policy actions recommended by the model are implemented, assess whether they produce the predicted achievement results.

Figure 10 presents the structure of the model and the relationships between the individual elements. The structure is analogous to Mendeleev's periodic table in chemistry and the standard model in particle physics (Cox and Foreshaw 2009, 171-217).

Following the principle of complementarity, details implementing the theory and model are contained in the following two studies. Phelps (2009) described an entire reporting process based on the percentile/probability duality. The purpose was to provide policymakers, practitioners, and the public with information regarding their schools. The standing for each school on each of the factors was represented by easily understood bar graphs. The second step was to depict each school's standing on the factors in terms of the influence on achievement (effect size). Figure 8 provided an example of how this might appear. There would be a separate graph for each achievement measure with each of the constituent curves representing a factor. On each of the factor-curves, there would be a mark representing the standing (Z-score) for the individual school. Each of the graphs would provide a wealth of comprehensible information not possible in any other form.[21]

Later, Phelps (2011d) described a process of classifying and synthesizing research and placing the results into a mathematical model of individual school achievement. Individual studies are required to estimate the effect size between multiple measures of achievement and multiple factors. These effect sizes are parameters in the simulation model along with the individual school parameters. The cost of increasing (or decreasing) the level of each of the factors was

included in the model. The model selected the most cost-effective factor for improving the multiple achievement measures and increased it to a point of diminishing returns. Then the model moved to the next most cost-effective factor until the money ran out and the predicted achievement was at the optimal level.[22]

XI. Changing the Paradigm

In *The Structure of Scientific Revolutions*, Kuhn (1970, 11) recounted the importance of paradigms, like the synthesis of laws, theory, applications, and instrumentation, in the history of science, stating: A paradigm "...is what mainly prepares the student for membership in the particular scientific community with which he will later practice... Men and women whose research is based on the shared paradigms are committed to the same rules and standards for scientific practice." Kuhn (1970, 15) went on to make an observation similar to that of Heisenberg (2011) regarding research without a theory:

> In the absence of a paradigm or some candidate for paradigm, all the facts that could possibly pertain to the development of a given science are likely to seem equally relevant. As a result, early fact-gathering is a far more nearly random activity than the one that subsequent scientific development makes familiar. Furthermore, in the absence of a reason for seeking some particular form of more recondite information, early fact-gathering is usually restricted to the wealth of data that lie ready to hand.

In contrast, Iannaccone (1967, 7) described education as "of the priesthood," i.e., education is based on individual beliefs rather than a common paradigm. In this sense, education seems more akin to Aristotelian philosophy where assumptions and conclusions are identified and discussed. Aristotle (384-322 B.C.E.) was important because of his efforts to place his observations of nature into categories. Each mutually exclusive element—water, air, fire, earth, and ether (stars and planets)—had a unique place in nature and its behavior was described by observation and logic. These assumptions were actually beliefs, such as objects fall to the ground because nature has determined that is their proper place, or the stars and planets are in the heavens because nature determined that is their proper place. Having a common explanation of the elements was given no consideration. The assumptions could not be proven and were subject only to logical argument. An assumption was considered true when consistent with observation and logic. A conclusion was justified if assumptions were considered true, and the relationship between assumptions and conclusion were consistent with observation and logic. The philosophical efforts were more qualitative than quantitative because the necessary instruments of observation, measurement, and analysis were not available. There was no common practice of testing philosophical assumptions—more accurately a theory—by careful experimentation and mathematical analysis (Asimov 1966, 1-9.)

An Aristotelian-type philosophy is reflected in school achievement beliefs, such as class size makes a difference because most people believe it makes a difference, or money makes a difference because everyone believes you get what you pay for. This philosophy also finds it way into professional training and practice. Paraphrasing Heisenberg (2011, 155), "there is no classification and synthesis."

In order to make changes in research, training, and practice, new concepts must be accepted and embraced, requiring a "quantum leap"!

Implications for Research

According to Feynman (1963, 2-1, 2-2), "We try to analyze all things; to put together things which a first sight look different, with the hope that we may be able to reduce the number of different things and thereby understand them better...At first the phenomena of nature were roughly divided into classes...[23] [H]owever, the aim is to see nature as different aspects of one set of phenomena. That is the problem—to find the laws behind the amalgamation of these classes. We wish to understand the phenomenon in terms of the smallest set of principles. To express it in a simple manner, what are things made of and how few elements are there?"

Researchers who choose to further explore the quantum achievement paradigm must adhere to the laws outlined previously, at least until superior laws are established. Several research strands, which grow out of these laws with the obvious purpose of identifying and accurately identifying and measuring the relevant factors, are as follows:

(1) Factors for which there are data. Perfect factors by identifying the best constituent variables and the best invariant weightings, and determine the effect size. Relate the variables to the behavioral traits of the staff, students, families, and communities, so they may be addressed by policies.

(2) Factors for which there are only proxy variables such as SES. The more imposing task is related to the unobserved behaviors of students, families, and communities. If these factors account for the largest proportion of explained variance, these behaviors seem to warrant the largest proportion of attention.

(3) Factors for unobserved variables. After identifying effective and ineffective schools by the unobserved school effect, comparative research could be conducted to identify the observable variables representing the behaviors associated with effectiveness.

(4) Unidentified factors for which there may or may not be data. Up to this point, factors have been developed because data, proxies, or unobserved estimates are available. In this case, the goal is to identify factors that are conceptually and statistically unrelated to already identified factors.

(5) Guess the influence that unidentified factors might have. Given the many difficulties, not all is lost. It is possible to reasonably estimate the explained variance for unobserved factors from other studies because the possible range of values is relatively narrow (the sum of the explained variance plus error = 1) (Phelps 2011c). These estimates combined with cost estimates generate a cost-benefit parameter allowing reasonably good comparisons among policy options. Based on these assumptions, policy decisions can be made and tested (Phelps 2011d).

Implications for Professional Training and Practice

Kuhn (1970) also addressed the sequence necessary for a paradigm shift. A flaw must be identified in current theory, research, or practice for which there is a better theory or research scheme. Be-

fore there can be a change, there must be a change in beliefs. After an alternative is proposed, it must be rigorously tested. If shown to be better than the previous practice, it must find its way into research, textbooks, and professional classrooms. Only then does the alternative find its way into professional practice. Many disciplines rely on specialized theories and mathematical models to solve practical problems (Schrage 1991, Williams 1985). Simply put, the liner regression model with individual variable does not provide adequate opportunities to research and address school-specific achievement problems. The quantum paradigm does. Students in many other disciplines are taught to solve problems as a part of their training for use in professional life. Students in education classrooms are more likely to follow Iannaccone's "priesthood" portrayal and write papers expressing beliefs. In the final analysis, every researcher, professional trainer, policymaker, and practitioner must make epistemological choices regarding the nature of reality. Is school achievement knowledge better:

- Based on an Aristotle/Iannaccone belief system of assumptions and conclusions (philosophy), or on Bohr's notion of reality—only what can be observed, measured, and tested (science)?
- Derived from independent and unrelated research or, as Heisenberg and Bohr advocated, from the classification and synthesis of complementary research, to establish the relation of the cause and effect, i.e., a theory?
- Described, explained, and predicted by the macroview (the average of a large number of schools), or by the quantum, school-specific microview?

Evidence and logical support have been presented for a substantial number of concepts, in the form of postulates, propositions, and mathematical principles, culminating in a theory and mathematical model of school achievement. To close, I again quote Heisenberg (2011, 155): "It is advisable to introduce a great wealth of concepts into a theory...and then to allow experiment to decide at what points a revision is necessary."

Endnotes

[1] "Quantum" comes from the Latin *quantus*, for "how much." A new branch of physics began when Max Planck discovered "...the energy radiated from a particle such as a photon or electron must be an integer multiple of a fundamental quantum" (Hawking 2011, ix).

[2] Brackets indicate my substitution of education language and examples.

[3] In this article, a postulate is defined as a claim of truth for the purpose of sequential reasoning leading to a final theory.

[4] This section is drawn from *Fundamentals Statistics in Psychology and Education*, by Joy Paul Guilford (New York: McGraw-Hill, 1965), 113-133.

[5] According to the *Stanford Encyclopedia of Philosophy*, "In the summer of 1654, Pascal returned briefly to mathematics in correspondence with Pierre Fermat (1601–65) about calculating probabilities associated with gambling. He summarized his findings in the *Traité du triangle arithmétique* which, like much of his other work,

remained unpublished until after his death." See "Blaise Pascal," http://plato.stanford.edu/entries/pascal.

[6] The information in this section is taken from Taylor (1982, 99-127).

[7] In calculus, dx means the change in x, and dy, the change in y.

[8] $q = 1 - p$, the chance of being incorrect.

[9] Euler's "e" is commonly used when rates of change are involved. Z is negative to make the curve path up then down (rather than the reverse), and it is squared to make it symmetrical around the mean—a Z-score of zero. The value of $e^{\wedge 0}$ is 1 (Barnett and Ziegler 1984, 775).

[10] Cumulative is the sum of the preceding values. In calculus it is integration. Therefore, the slope of the cumulative curve is the value of the normal curve at the same Z-score.

[11] See, James L. Phelps, "A Practical Method of Policy Analysis by Simulating Policy Options," *Educational Considerations* 39(1): 50, 62.

[12] For a review of cost-effectiveness, see *Cost-Effectiveness and Educational Policy*, edited by Henry M. Levin and Patrick J. McEwan, 2002 *Yearbook of the American Education Finance Association* (Larchmont, NY: Eye On Education, 2002).

[13] In quantum physics, the position and momentum of a particle (photon or electron) cannot be measured simultaneously. This phenomena is called the Heisenberg Principle of Uncertainty (Hawking 2011, 148-149). Separate but complementary analysis of position and momentous are required. Bohr (2011, 417) refers to this as the principle of complementarity.

[14] A principle of chaos theory is small changes in inputs produce huge changes in outcomes (Gleick 1987, 9-33).

[15] Taylor (1982) described this as separating the systematic error from the random error.

[16] Interestingly, the "school effect" technique, averaging over time, is the same technique as determining factors, so the same conditions must apply; that is, the definitions and measure of the predicting variables must be consistent.

[17] Later it was discovered that the nucleus could be divided into smaller particles.

[18] See Guilford (1965, 403-404) for a vivid example.

[19] The summary that follows is drawn from Heisenberg (2011).

[20] Language in brackets was added by the author.

[21] While the purpose here is to describe a theory and model of school achievement rather than to present research results, the estimates of the explained variance in the above study are instructive. SES accounted for the largest percent of explained variance (in the range of $\pm$ 60%) and the unobserved effectiveness was second (in a range of $\pm$ 25%). The factors identified by Hedges, Laine, and Greenwald (1994), such as staff roles and staff characteristics, were small ($\pm$ 7%). No data were available for the factors identified by Walberg (1984).

[22] Computer software is readily available for this purpose. Microsoft Excel and Solver, a function within Microsoft Excel, were used in the studies cited.

[23] This is what Aristotle called elements and this article refers to as factors.

References

Asimov, Isaac. 1966. *Understanding Physics*. New York: Dorset Press.

Barnett, Raymond A., and Michael R. Ziegler. 1984. *College Mathematics*. San Francisco, CA: Dellen Publishing.

Bohr, Niels. 2011. "Can Quantum Mechanical Descriptions of Physical Reality Be Considered Complete?" In *The Dreams that Stuff Is Made of: The Most Astounding Papers on Quantum Physics and How They Shook the Scientific World*, edited by Stephen Hawking, 471-483. Philadelphia, PA: Running Press.

Born, Max. 2011. "Statistical Interpretation of Quantum Mechanics." In Hawking, *The Dreams that Stuff Is Made of*, 448-461.

Cox, Brian, and Jeff Forshaw. 2009. *Why Does E = MC²?* Cambridge, MA: Da Capo Press.

_____. *The Quantum Universe*. 2011. Cambridge, MA: Da Capo Press.

Einstein, Albert, Boris Podolsky, and Nathan Rosen. 2011. "Can Quantum Mechanical Descriptions of Physical Reality Be Considered Complete?" In Hawking, *The Dreams that Stuff Is Made of*, 463-470.

Feynman, Richard P. 1963. *The Feynman Lectures on Physics*. Reading, MA: Addison-Wesley.

Glass, Gene V., and Mary Lee Smith. 1978. *Meta-Analysis of Research on the Relationship of Class-size and Achievement*. San Francisco, CA: Far West Laboratory for Educational Research and Development.

Gleick, James. 1987. *Chaos: Making a New Science*. New York: Penguin Books.

Guilford, Joy Paul. 1965. *Fundamentals Statistics in Psychology and Education*. New York: McGraw-Hill.

Hawking, Stephen, ed. 2011. *The Dreams that Stuff Is Made of: The Most Astounding Papers on Quantum Physics and How They Shook the Scientific World*. Philadelphia, PA: Running Press.

Hedges, Larry V., Richard D. Laine, and Rob Greenwald. 1994. "Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes." *Educational Researcher* 23 (3): 5-14.

Heisenberg, Werner. 2011. "Excerpts from 'The Physical Principles of the Quantum Theory'." In Hawking, *The Dreams that Stuff Is Made of*, 151-236.

Iannaccone, Laurence. 1967. *Politics in Education*. New York: Center for Applied Research in Educational.

Kuhn, Thomas S. 1970. *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.

Levin, Henry M. 1997. "Raising School Productivity: An X-Efficiency Approach." *Economics of Education Review* 16 (3): 303-311.

Levin, Henry M., and Patrick J. McEwan, eds. 2002. *Cost-Effectiveness and Educational Policy*. 2002 Yearbook of the American Education Finance Association. Larchmont, NY: Eye On Education.

Phelps, James L. 2009. "Measuring and Reporting School and District Effectiveness." *Educational Considerations* 36(2): 40-52.

_____. "Another Look at the Glass and Smith Study on Class Size." 2011a. *Educational Considerations* 39(1): 3-17.

_____. "A Practical Method of Policy Analysis by Considering Productivity-Related Research" 2011b. *Educational Considerations* 39(1): 18-32.

_____. 2011c. "A Practical Method of Policy Analysis by Estimating Effect Size." *Educational Considerations* 39(1): 33-48.

_____. 2011d. "A Practical Method of Policy Analysis by Simulating Policy Options." *Educational Considerations* 39(1): 49-62.

_____. 2011e. "Factor Analysis of Explanatory Variables in a Achievement Production Function." *Educational Considerations* 39(1): 71-74.

Phelps, James L., and Michael E. Addonizio. 2006. "How Much Do School and Districts Matter? A Production Function Approach to School Accountability." *Educational Considerations* 33(2): 51-62.

Planck, Max. "The Atomic Theory of Matter." 2011. In Hawking, *The Dreams that Stuff Is Made of*, 32-48.

Schrage, Linus. 1991. *Lindo: An Optimization Modeling System*. 1991. San Francisco, CA: Scientific Press.

Stanford Encyclopedia of Philosophy. "Blaise Pascal." http://plato.stanford.edu/entries/pascal.

Taylor, John. 1982. *An Introduction to Error Analysis*. Sausalito, CA: University Science Books.

Walberg, Herbert. 1984 "Improving the Productivity of America's Schools." *Educational Leadership* 41(8): 19-27.

Williams, Hilary P. 1985. *Model Building in Mathematical Programming*. Chichester, U.K.: John Wiley and Sons.

Wooldridge, Jeffrey. 2000. *Introductory Econometrics: A Modern Approach*. Cincinnati, OH: South-Western College Publishing.

Doctoral Programs in Educational Leadership: A Duality Framework of Commonality and Differences

Perry A. Zirkel

Supreme Court Justice Oliver Wendell Holmes reportedly characterized President Franklin Delano Roosevelt as having "a second-class intellect but a first-class temperament" (Ward 1989). The present state of, and the proposals to date for, doctoral programs in educational leadership do not sufficiently reflect this implicit recognition of a common core of competencies and this explicit differentiation for what Sergiovanni (1986, 17) and other leadership scholars (e.g., Sternberg and Wagner 1986) have termed "practical intelligence."

In recent years, doctoral programs in education leadership have been subject to notable criticism and proposals for reform. Starting with a synthesis of this criticism, this article focuses on the two primary constituencies—university faculty members who teach in such programs and school superintendents, who are the

---

Perry A. Zirkel is University Professor of Education and Law at Lehigh University, where he formerly served as Dean of the College of Education and more recently held the Iacocca Chair in Education for its five-year term. He has a Ph.D. in Educational Administration and a J.D. from the University of Connecticut, and a Master of Laws degree from Yale University. He has written more than 1,300 publications on various aspects of school law, with an emphasis on legal issues in special education. He writes a regular column for "Principal" magazine and did so previously for "Phi Delta Kappan" and "Teaching Exceptional Children." Past president of the Education Law Association and cochair of the Pennsylvania special education appeals panel from 1990 to 2007, he is the author of the CEC monograph "The Legal Meaning of Specific Learning Disability"; the recently published books, "A Digest of Supreme Court Decisions Affecting Education" and "Student Teaching and the Law"; and the two-volume reference "Section 504, the ADA and the Schools," now in its third edition. He recently received the Research into Practice Award from the American Educational Research Association (AERA) and the Excellence in Research Award from AERA's Division A (Administration, Organization, and Leadership).

leading practitioners such programs serve. Literature concerning other constituencies, e.g., school principals and certification programs in educational leadership, are included to a limited degree to help inform or sharpen this focus. The thematic lens for the foundational literature review is to determine the extent that education leadership faculty and school superintendents share a community of interest and, conversely, the scope of the remaining divide between these two groups in terms of shaping the appropriate approach at the doctoral level. The culminating vision is for doctoral study in education leadership that reflects both this commonality and differentiation.

More specifically, this article consists of three parts. The first part reviews the literature that contains the criticism, along with proposals and responses for reform. The second part canvasses the competencies jointly developed and separately assessed by faculty and school superintendents. The third part examines other relevant evidence as to the extent of common vs. divided interests between these two constituencies. The purpose is to provide a foundational framework for re-examining doctoral programs in educational leadership. As with other analyses (e.g., Murphy 1991), the focus on the pinnacles of the doctorate and the superintendent may incidentally but not necessarily result in more general lessons for practitioners and the professoriate in educational leadership.

Criticism

The recent criticism, centering on the national movement for school reform and blanketing schools of education generally, has extended to education leadership programs in particular. For example, despite extensive redesign efforts in educational leadership programs dating back more than a decade, the U.S. Department of Education (2005) has criticized these programs as lacking programmatic vision and coherence. At the same time, Levine's (2005, 23) well-publicized study of educational leadership programs characterized their trajectory during the most recent decades as "a race to the bottom," with the weaknesses including low admissions standards, inadequate clinical components, and "curricula ... disconnected from the needs of leaders and their schools." For example, he reported 2004 data from the Educational Testing Service showing that the mean Graduate Record Exam scores in education leadership were the second lowest for 16 reported fields, including elementary and secondary education. Echoing previous recommendations within the profession itself, specifically the National Commission on Excellence in Educational Administration in 1987 and the National Policy Board for Educational Administration in 1989 (McCarthy 1999), Levine called for drastic elimination of the many programs in educational leadership.

Most recently, the current head of the U.S. Department of Education, Arne Duncan, who came directly to this position from a school district superintendency, criticized schools of education for lack of rigor (Duncan 2009). Although his particular focus was teacher preparation, his criticism of schools of education was generic. Similarly, using the M.B.A. reform movement as the analogy, Maranto, Ritter, and Levine (2010, 25) criticized schools of education for "lack of sufficient academic rigor and applied acuity," recommending reorganization "around highly rigorous academic disciplines with well-established academic quality, and which seem likely to offer the skills and content teachers and administrators need."

The criticism specific to the doctoral level, like that of educational leadership preparations programs generally, is nothing new (McCarthy 2002). For example, Brown (1990) traced concerns about the quantity and quality of Ph.D. and Ed.D. programs in education back to the 1930s and 1940s. Focusing on prestigious universities, Clifford and Guthrie (1988) tracked back the lack of distinction of, and distinctiveness between, such programs even earlier.

Research supported such criticism. For example, Osguthorpe and Wong (1993) found—consistent with a string of earlier studies for education generally (Anderson 1983; Deering and Whitworth 1982; Dill and Morrison 1985; Moore, Russel, and Ferguson 1960; Robertson and Sistler 1971; Schneider et al. 1984) and educational leadership specifically (Davis and Spuck 1978; Norton 1992; Norton and Levan 1987)—that Ed.D. and Ph.D. programs in education were remarkably similar, including their research and statistics requirements. As a framework for the resulting proposals, Osguthorpe and Wong (1993, 60) outlined the following four basic options for schools of education:

(a) continue to offer both the Ed.D. and Ph.D. in their current undifferentiated state …; (b) continue to offer both degrees, but differentiate between program requirements for each; (c) offer only one degree and define more clearly the expectations for the degree, specifically the role of the dissertation; or, (d) offer a degree with a title other than Ed.D. or Ph.D.

They characterized the first option as dangerous.

The Critics' Proposals

Predating the recent wave of criticism, the National Policy Board in Educational Administration (NPBEA 1989) advocated the second option, recommending that the preparation of educational leaders be limited to the doctoral level altogether. At about the same time, Courtenay (1988, 18) argued for the third option, more specifically suggesting "the various fields of education use the Ph.D. only, but with two tracks, one for scholars of practice and one for scholarly practitioners." Instead, Goodlad (1990) proposed the fourth option in the form of a Doctor of Pedagogy (D. Paed.) as the only terminal degree in education. Similarly, the education leadership faculty at Texas A&M University not only proposed but also implemented a Professional Studies Doctorate (PSD), including a cohort of mid-level school administrators, local superintendents as clinical professors, and a formal field component for reflective practice, as an alternative to the Ph.D. or Ed.D. (Bratlien et al. 1992). The more recent proposals have varied considerably.

Initially advocating the second option, Shulman (2004), the then president of the Carnegie Foundation for the Advancement of Teaching, recommended differentiation between the Ed.D. for practitioners and the Ph.D. for scholars. Subsequently, the Carnegie Foundation and the Council of Academic Deans from Research Education Institutions launched an initiative among 21 universities nationwide to "reclaim" the Ed.D. by distinguishing it from the Ph.D. as specifically oriented to preparing practitioners rather than professors, including applied rather than academic research (Redden 2007).

In the meanwhile, however, Shulman and his Carnegie colleagues proffered their prescription for reclaiming and distinguishing the education doctorate under the rubric of the fourth option. More specifically, based on a Carnegie study of doctoral programs in six disciplines, Shulman and his colleagues characterized the problems of the education doctorate as "chronic and crippling" (Shulman et al. 2006, 25, 27) and proposed—instead of designing the prevailing Ed.D. by subtraction as a "Ph.D.-Lite"—development, on a "zero-base" approach, of a separate new Professional Practice Doctorate (P.P.D.) akin to the differentiation between the M.D. and the Ph.D. in medical sciences. Like the M.D., the P.P.D. would have a "rigorous" (29) substantive professional assessment but no dissertation requirement at the end. Although acknowledging that the name was not the primary issue and that "[t]here is real danger in taking to extremes the distinction between a professional practice degree and a research degree" (30), Shulman and colleagues did not explore the scope of the overlap.

More generally, Lagemann's (2008) advocacy of a distributed model of educational research provides indirect support for a separable doctoral program in education. She argued that universities should reserve clinical research, more specifically problem-finding and translational research, for schools of education whereas problem-solving research in education should be centered in the disciplines of arts and sciences.

Specifically in educational leadership, Levine (2005) recommended a combination of the third and fourth options—eliminating the Ed.D. degree as being academically inadequate for practitioners and retooling the master's curriculum into a new terminal Master's in Educational Administration (M.E.A.) analogous to the M.B.A. At the same time, he recommended reserving the Ph.D. in educational leadership for the nation's most research-oriented universities and exclusively for academic careers as scholars of school leadership, resulting in reduction to one-quarter of the present expansive doctoral enrollments in educational leadership.

The Reactions and Counterproposals

Assessing the response to this criticism, Levine and Dean (2007) noted major differences among the stakeholders, with the American Association of School Administrators (AASA) being partially supportive and the University Council on Educational Administration (UCEA) providing a negative response. More specifically, the excerpted AASA response, issued jointly with the two national principals' organizations, affirmed the disconnect between the scholarly preparation and practical needs; however, they did not support replacing the Ed.D. with a M.E.A., reasoning as follows: "Changing a label will not solve a problem; changing the rigor the programs will" (67). The UCEA similarly supported Levine's recommendation for rigorous standards but criticized the quality of his research. Moreover, with regard to the Ed.D., the UCEA representatives endorsed distinctively redefining the Ed.D. but along the lines of the Carnegie initiative rather than Levine's proposed reduction to a Master's level professional degree (Young et al. 2005).

The other views within academia have been diverse with regard to the doctoral level. For example, agreeing with Levine's recommendation for elimination of the Ed.D. and specifically targeting "general managers" (e.g., superintendents), Murphy (2006b, 533) acknowledged that "one could imagine a renamed doctoral degree, as suggested by Lee Shulman, that addresses the muddled distinction between the Ph.D. and the Ed.D.," but he concluded that "[c]reating an entirely new master's degree such as the MEL [Master of Educational Leadership] would make the most sense because it would have the cachet of something special."

Agreeing with the indistinctiveness problem but not the pro-grammatic solution, Evans (2007, 555) argued for the opposite of Shulman et al.'s proposal for a P.P.D—namely, a single Ph.D. program in educational leadership based on a "unitary scholar-educator class or set of activities to which people make differential contributions according to time, talents, interest, and abilities." In his view, a separate practitioner's degree, whether the traditional Ed.D. or the proposed M.E.A. or P.P.D., "institutionalized a philosophical and practical separation that would contribute to a flawed conception of both." Counterarguing that the disagreement was largely a matter of semantics, Shulman (2007, 561) responded that the P.P.D. has a broad basis composed of a "wisdom of practice," which is "deeply theoretical," and other sources, such as "values, visions of the pos-sible, … and equity." Thus, while agreeing that the worlds of the scholar and the practitioner overlap, each of them fused the two into their respective program polarities.

Similar to Evans, Bredeson (2006, 21) argued for "integrated Ph.D. programs" in educational leadership, characterized by "flexibility to address individual specialization needs while not sacrificing the substantive dialog between scholar/researchers and educational practitioners that comes in commonly shared seminars and learning activities where there is substantial overlap in professional knowl-edge." Reaching the same conclusion via advocating the elimination of the Ed.D., Deering (1998, 247) argued: "By offering two terminal degrees that are more similar than different, colleges and depart-ments of education unwittingly cause confusion among students and faculties, undermining the standing of all terminal degrees in education." Using the nursing profession as the analogy, he recommended strengthening the Ed.S. to replace the Ed.D.

In contrast, pointing out the lack of distinction both between and for the Ph.D. and the Ed.D. and reiterating the conclusion of his earlier coauthored book (Clifford and Guthrie 1988), Guthrie (2006) argued for entirely separate tracks with respective rigor for practitio-ners and researchers. Selecting the health and engineering profes-sions as the appropriate analogy, he argued that a "dual purpose single track program" (24) woefully compromised research prepara-tion and practitioner training. Similarly agreeing with Levine's "mission muddle" criticism, Shepard, as the president of the National Academy of Education, was quoted by *Education Week* as follows: "By blending both programs, you serve neither purpose well" (Viadero 2008, 6). Taking the matter a step further, Young (2006) outlined, as a working model, the potential differentiation between the Ed.D. and Ph.D. in educational leadership. More spe-cifically, she proposed the following differential features for the Ed.D.: the use or portfolios (rather than exams) for comprehensive assessment; a field (rather than teaching/research) internship, which includes program evalua-tion experience/proficiency (rather than, for example, a professional conference presentation); and applied (in contrast to original) research for the dissertation with at least one practicing profes-sional (in contrast to a faculty scholar from a related discipline or another institution) as a member of the dissertation committee. The proposed coursework differed in both titles and amount for the leadership and research cores, with the Ph.D. having the additions of a specialized concentration and a cognate area. However, she did not address any purposeful commonality in the design or in the specific competencies at the entry and exit levels.

Similarly, the debate concerning the Ed.D. has gone in diverse directions more specifically in terms of the doctoral dissertation. Representing the integrative view with respect to the dissertation, Malen and Prestine (2005, 7) advocated "blurring the distinction between scholars and practitioners, 'producers' and 'consumers' of research, and professional (Ed.D.) and research (Ph.D.) degrees" by retaining yet revitalizing inclusive but rigorous dissertation require-ments. Representing a moderate step in the opposite, new direc-tion, Duke and Beck (1999) advocated expansion, not replacement, of the traditional dissertation in education via alternative formats, such as a series of publishable articles, based on precedents in vari-ous fields in arts and sciences. As another variation in the differenti-ated direction, Andrews and Grogan (2005)—using the analogy of other professional doctoral degrees, such as the J.D. and the M.D.—argued for a differentiated Ed.D. dissertation, replacing the traditional arts and science scholarly study with a portfolio that included not only reflection papers but also a capstone action research project. Implementation of these proposals has been uneven. Describing the dissertation as "an artifact of the arts and science model that is conspicuous by its absence in nearly every other professional school (e.g., law schools, college of veterinary medicine)," Murphy and Vriesenga (2005, 33) traced the contours of the rare—i.e., four of 161—Ed.D. programs that appeared to have truly alternative, professionally-anchored dissertations. The key characteristics included a practice, rather than theory, orientation; integrated activities; collaborative grounding; and a client, rather than faculty, focus. However, they admitted that these programs were only "inchoate initiatives" that thus far lacked "evaluation information" (50). Reporting more recent developments in this differentiated direction, Imig (2011), as director of the Carnegie Project on the Education Doctorate (CPED), recounted movement toward a capstone project to replace the traditional dissertation among Ed.D. programs. Exemplifying their efforts, the various mem-ber institutions of the CPED are considering alternatives to a written product, and, according to Imig (2011, 12), "there is preliminary agreement … that more than one candidate may work on a single capstone." Imig predicted "we will continue to have multiple forms of the capstone or culminating project for the foreseeable future, but through studying these variations, a collective understanding of effective outcomes will emerge."

Explaining that the redesign of a differential Ed.D. will require changes in the organizational structures of and faculty roles in schools of education, Perry (2011) reported that the second phase of the CPED consortium will facilitate this process. More specifically, armed with a $700,000 FIPSE grant for 2010-2013, the focus is to document, evaluate, and disseminate the implementation of these "professional practice doctorates" (Perry 2011, 4). She cautioned, however, that this period is not sufficient to reverse and resolve the "century of confusion" concerning the Ed.D.

Finally and most broadly, various respected sources within the education leadership professoriate have recommended improve-ments in educational leadership preparation programs generally, ranging, for example, from Bredeson's (1991) call for reflective incorporation of personal experience to more recent emphases on adopting the transformative model of leadership (Brown 2006a, 2006b; Leithwood et al. 2005), focusing this transformation on social justice (Brown 2008; Cambron-McCabe and McCarthy 2005),

or focusing it more narrowly on student achievement (Hale and Moorman 2003).

The Recent Trends

During earlier decades, doctoral degrees in educational leadership proliferated, with the rate of growth higher for the more prestigious Ph.D. than for the Ed.D., as universities reduced or waived the foreign language requirement and the two programs became more similar to each other. For example, Nelson and Coorough (1994) reported that the field of educational administration accounted for 40 Ph.D. dissertations and 221 Ed.D. dissertations in 1960 as compared to 494 Ph.D. dissertations and 802 Ed.D. dissertations in 1990.

Serving in effect as a baseline for the more recent period, Hackmann and Price's (1995) national survey found rather wide variety within a common template for doctoral programs in educational leadership. For example, entry requirements for almost all of the responding 127 institutions (representing a 68% response rate) used grade point average (GPA) as an admissions criterion, but they varied notably in terms of whether the GPA was at the undergraduate and/or graduate level and what the minimum was for either one. Similarly, the number of credit hours varied widely from 28 to 67 for coursework and from zero to 30 for the dissertation. At the exit end, only three institutions reported no comprehensive examination, and three programs reported having the following respective replacements for a dissertation: a field research project, an executive position paper, or a portfolio that includes a synthesis exercise. As for the clinical side, the majority of the programs did not require prior teaching (52%) or administrative experience (73%), but half of the programs reported requiring completion of an administrative internship. However, none of these analyses differentiated Ed.D. from Ph.D. programs.

Since then, as Orr (2006) observed, of the approximately 200 institutions offering doctoral programs in educational leadership, a handful has redesigned the Ed.D. in educational leadership as distinguishably practitioner-oriented compared with the more scholarly Ph.D. Baker, Orr, and Young (2007) determined that the number of doctorates granted in educational leadership increased 31% from 1993 to 2003, with most of the growth attributable to less selective institutions newer to the field that had far more limited graduate resources and yet no more likelihood for innovation. In addition, Orr (2007) also noted a movement at a few institutions away from the traditional dissertation to a project-based study by an individual doctoral student or a team of them.

Other efforts at reform have surfaced as well. For example, Hoyle, English, and Steffy (1998, 181) advocated a "professional studies model" that starts with mapping the various sets of standards, such as those of AASA and ISLLC. However, while parenthetically noting that "[a] review of current standards reveals an eighty to ninety percent overlap between indicators," they did not present the particulars of this review. Moreover, their model is not specific to the doctoral level, much less the distinction between an Ed.D. and a Ph.D. The program that they cite as illustrative of the doctoral version of their model is the Ed.D. program in educational leadership at Duquesne University, which had the reported features of a cohort of practicing administrators, concentrated monthly and summer classes, university-district learning communities, problem-based learning, and portfolios. Separately and without any specified evaluative framework, Hoyle and Torres (2010) recommended

model status for Seton Hall University's executive "fast track" Ed.D. program along with the contrasting University of Wisconsin's Ph.D. program in educational leadership and policy analysis,

Similarly, Everson (2006, 7) promoted the Ed.D. program at St. Louis University as including cohort-based teams of three to four doctoral students who are mid-level school leaders who conduct "field-based or field focused" projects as their culminating activity. She reported positive perceptions among the participants as preliminary evidence of successful progress. In a follow-up article, Everson (2009) reported that the program currently had 242 participants, compared to 28 in the Ph.D. program in educational leadership, and further explained the emphasis on problem-based learning and team-based culminating projects, including individual analysis reports and oral examinations. However, the only additional assessment information was reported enhancement of the evaluation design "to address specific areas of interest to the faculty regarding the practices of program graduates" (97).

A separate, although overlapping, example in the literature is the Ed.D. at Arizona State University. In accordance with the Carnegie recommendation (Golde 2007; Shulman 2005) for developing "signature pedagogies" akin to those in medicine, law, and neuroscience, Olson and Clark (2009) described the invention and refinement of a "leaders-scholar community" approach in the Ed.D. program in educational leadership at Arizona State University. This signature pedagogy includes cohort subgroups of five to seven students assigned to one faculty member as their collective doctoral adviser and "culminating in action research dissertation defenses and degree completion by all student members" (217). Although the effectiveness of such a program is not settled, Olson and Clark (2009, 218) presented the preliminary results of their ongoing qualitative research evaluation in terms of the "testimony" of the participating faculty and students.

Thus, similar to the common characteristics of "promising" principal preparation programs (Jackson and Kelley 2002, 197), these innovative doctoral programs in educational leadership tend to include problem-based learning, cohorts, and collaborative partnerships, and "a clear, well-defined curriculum focus reflecting agreement on the relevant knowledge base" (208). Also similar to the research concerning educational leadership preparation programs more generally, the studies of the combined effect of these best-practice doctoral components is scant. As McCarthy and Forsyth (2009, 117) have pointed out, the prevalent "perception studies" are not sufficient to establish effectiveness. Hoyle and Torres' (2008) interview study of current program faculty and their selected graduates of six top-ranked education leadership doctoral programs serves as an example. Instead of limiting the study to participant perceptions, the ultimate dependent variables would appear to include, for example, superintendent renewal and student achievement. However, as Hoyle's (2007) case study of the first of these two variables showed, the research to date has been largely limited to initial explorations. Similarly, the research to date that uses student achievement as the dependent variable is either based on varying broad conceptions of leadership (e.g., Leithwood, Patten, and Jantzi 2010; Robinson, Lloyd, and Rowe 2008) or an insufficiently clear conception of superintendent effectiveness (e.g., Waters and Marzano 2006). More promising would be a mediated model—akin to Kottkamp's (2010) longitudinal evaluation model

that included, along with mediating variables, doctoral program characteristics, graduates' leadership effectiveness, and student learning.

Overall, in the absence of more objective data and in light of the institutional drift to less selective colleges and universities, the innovations do not seem to have provided a net elevation of the doctoral programs in education leadership. Murphy's (2006a, 490) assessment would appear to be on target: "While a number of programs are better than [Levine] suggests, far too many are inadequate and, with the heightened pressures [among administrators] for high-status credentials and fast-track programs, may be getting worse."

Competencies

The reconfiguration of the terminal degree structure in educational leadership ultimately depends on the "competencies"—used here as a generic rubric for the various content areas of the standards, including knowledge, dispositions and performances—that programs seek to target and nurture. During the past three decades, superintendents and professors have led collaborative groups in developing successive sets of standards for educational leaders. Although other organizations led the parallel development of competency inventories for principals (Jackson and Kelley 2002), the two major sets specific to the focus here are those developed under the rubric of AASA and the Interstate School Leaders Licensure Consortium (ISLLC).

The Development of Standards

In 1982, the AASA, which is the national organization that represents superintendents and other central office school administrators, published *Guidelines for the Preparation of School Administrators*. One of the purposes for the guidelines was to "assist … training institutions in refining … doctoral programs in educational administration" (AASA 1982, 2). The development included the input of education leadership professors (Hoyle 1985; Hoyle 1987). The 1982 guidelines consisted of seven performance goal areas—each with identified competency and related skills, for a total of 43 skills—concerning the learning climate, governmental support, curriculum, instructional management, evaluation/improvement, resource allocation, and research (AASA 1982; Hoyle 1985). Subsequently, the AASA published successive texts based on these standards (Hoyle, English, and Steffy 1985, 1998). Further, in 1993 the AASA published more specialized guidelines specific to the preparation of superintendents, *Professional Standards for the Superintendency*, which were the basis for a textbook that the UCEA Center for the Study of the Superintendency developed in 2005 (Hoyle et al. 2005).

In 1996, the National Policy Board for Educational Administration (NPBEA), which represents ten major organizations including the AASA and UCEA, developed the ISLLC standards for educational leaders. Designed as a new foundation from both the academic and practice domains and deliberately left as broad, evolving conceptions (Murphy 2005), these six standards, which each have from three to nine more specific functions, concern a shared vision; effective school culture/curriculum; efficient management; school/community relations; ethical conduct; and advocacy/responsiveness (CCSSO 2008). More than 40 states use the ISLLC standards as the platform for their certification programs for educational leaders (Roach, Smith, and Boutin 2010; Toye et al. 2007).

In 2002, the National Council for Accreditation of Teacher Education (NCATE) adopted the Educational Leadership Constituent Council (ELCC) program standards, which are an adapted version of the ISSLC standards that includes a seventh standard for a culminating internship, for its review of educational leadership programs (NPBEA 2002). NCATE's ELCC implemented these standards for accreditation reviews (Jackson and Kelley 2002). Recently, Young (2011, 7) reported, "over half of the 500 programs nationwide have revised their leadership programs to align with ELCC standards and have been reviewed by the ELCC on behalf of NCATE."

At about the same time as NCATE's adoption of ELCC standards, the Educational Testing Service developed the School Leaders Licensure Assessment based on the ISLLC standards (Murphy et al. 2009). As of 2006, despite Anderson's (2001) criticism of this examination from a social justice perspective, approximately 25 states required its use for initial certification (Toye et al. 2007).

In a two-year project starting in 2006, a national panel revised the original, 1996 version to tie each function to "research-based pedagogical practices as well as empirical knowledge" (Young 2008, 1). In 2008, the NPBEA issued the resulting revision, renamed the Educational Leadership Policy Standards (NPBEA 2008). NCATE adopted these standards as the benchmarks for evaluating educational leadership program and licensure exams for aspiring school administrators (Hoyle and Torres 2010). As the latest phase in the updating process, NPBEA (2010) issued draft ELCC standards for building-level leaders, including principals, and district-level leaders, including superintendents. The two sets both consisted of eight standards and subparts, called "elements," that are in parallel but customized to their respective organizational level in both the wording and supporting, updated research and commentary. After a consultation process for review, comment, and revision, the ELCC Standards Revision Steering Committee submitted the final versions to NCATE (Mawhinney and Young 2010).

The Perspectives of the Constituencies

Although the various surveys from the single perspective of professors or superintendents at the state or national level seem to show general endorsement of these overlapping sets of standards, the surveys that measure multiple perspectives reveal that these two constituencies also differ in significant respects in their assessments of the relevance and importance of the standards.

Single perspective. Two successive clusters of educational leadership dissertations provided single constituency perspectives of the 1982 AASA guidelines. First, a cluster of dissertations at Texas A&M University in the mid-1980s assessed the extent of support within separate constituencies for this set of competencies. More specifically, these successive surveys found general endorsement of the AASA list among the representatives of the UCEA and the National Council of Professors of Educational Administration (Edgell 1983); a national sample of school superintendents (McClellan 1985); a random sample of members of the National Association of Secondary School Principals (Fluth 1986); and, more peripherally in terms of constituencies, Texas junior/community college administrators (Voelter 1985).

However, despite the relatively defensible sampling design and response rates of these studies, a final study revealed that the results from the professoriate could be merely politically correct "lip service" to this significant practitioner organization's document.

More specifically, in a national survey of educational leadership department heads conducted by Piper (as cited in Hoyle 1985), 69% endorsed the 1982 AASA guidelines, but 54% opposed NCATE's adopting them for use as criteria in accrediting educational leadership programs.

Second and less relevant here, a cluster of dissertations toward the end of the same decade focused on prioritization of the 1982 AASA goals and skills by national and state samples of superintendents. More specifically, Sclafani's (1987) national sampling—which consisted not only of a representative sample but also a separate sample of superintendents whom peers in their state had nominated as highly effective—and the follow-up state samples of superintendents in Texas (Collier 1987) and Tennessee (Douglas 1990) found various significant differences in priorities within and among these groups of superintendents. However, the instrument in these three dissertations consisted of a revised version of the AASA list; for example, based on pilot testing with small groups of superintendents in three states, school finance became an additional performance goal area for management, and an additional 13 skills replaced five of the original total of 43.

In a follow-up to the Sclafani study, Sass's 1989 dissertation revealed limited significant differences for various demographic variables, including prior superintendency experience, among a national sample of educational leadership professors with regard to their rankings of the AASA goals and skills. On the limitations side, his response rate was 42.5%, and he performed an excessive number of analyses of statistical significance.

A pair of peripherally pertinent studies focused on single perspectives related to the ISLLC standards. First, in a study intended to determine to the extent to which superintendent search announcements reflected the perspective of school boards, Ramirez, Carpenter, and Guzman (2007) found general but not completely consistent alignment between the ISLLC standards and the selection criteria of these announcements. However, the sample was not random, and the authors acknowledged that such criteria result from a broad-based, multiple-constituency process rather than a single board perspective. Second, in a survey of 500 principals who worked in specially designated urban districts in New Jersey, the respondents identified topics that fit within standards two and three, but their response rate was limited to 16% of this relatively restricted population, and the congruence between the responses to their open-ended survey item and these broad categories was unclear (Friedland, Fleres, and Hill 2007).

Multiple perspectives. The corresponding studies that compared the assessments of more than one constituency, however, found not only commonalities but also significant differences. Although the focus here is on superintendents' and professors' perspectives of these successive sets of standards, findings are also included for other constituencies.

Although the Ed.D. dissertation of Sass (1989) collected rankings of AASA standards from educational leadership professors, he cautiously compared his results with those Sclafani had obtained two years earlier for superintendents. Upon doing so, he observed that both groups ranked climate first and research last, but they appeared to differ in terms of some of the other goals and skills.

In another Ed.D. dissertation the same year, which was based on the eight competency domains of California's principal licensure, education leadership professors gave significantly higher ratings than did school principals with regard to the relevance (two of the eight domains) and effectiveness (six of the eight domains) of their preparation programs; however, the limited size and scope of the sample and the unsophisticated statistical analysis left the generalizability of these findings in question (Lem 1989).

Similarly, the conference paper of Gousha and Mannon (1991) reported no significant difference among large-city superintendents, administrator preparation faculty, and state education agency personnel with regard to their perceived importance of eleven of thirteen competencies, but their report had several serious limitations. First, their paper provided only cryptic information about the subjects and instrument of the study. Second, the authors reported using the entire population of these three groups, which did not square with their use of inferential statistics. Third, the superintendent group was limited to large city superintendents, and only eleven members of this group responded to the survey. Fourth, some of the competency items were vague and without elaboration or example, such as "foundational knowledge" and "specific knowledge," and their relationship to the established sets of standards was unclear.

Subsequently, a pair of doctoral dissertations examined multiple constituencies' prioritization of the ISLLC standards. First, in a study of four stakeholder groups in Alabama—teachers, parents, administrators, and professors—administrators differed significantly from professors with regard to the perceived importance of one of the six ISLLC standards; specifically, administrators perceived management as more important than the professors did (Marshall and Spencer 1995). Yet, the limitation of the study to one state, the difference in sampling procedure for the education leadership professors from that for the other three constituencies, and the brief presentation of the data analysis warn against overreliance on the results.

Second and less relevant in the absence of a sample of professors, a study of three stakeholder groups in Missouri—superintendents, principals, and school board presidents—determined that superintendents significantly differed from the principals with regard to the perceived importance of five of the six ISLLC standards, although their ratings did not significantly differ from board presidents (Ray 2003). The response rates, especially the 34% for school board presidents, and the failure to reach the threshold sample size for representativeness for each of these three populations limited the generalizability of the results even for a single state.

In sum, the evolving standards represented most recently by the revised ELCC standards provide a common core developed by both practitioners and professors and largely accepted by both constituencies. Despite limitations in the various research studies to date, their cumulative and rather comprehensive extent suggests a common foundation for parallel but differentiated extensions.

Complementarity

Other sources of evidence of the extent of the commonality of, yet differences between, superintendents and educational leadership faculty include research findings regarding their respective demographics and their interests or values. The rather consistent theme that emerges from these various sources is the substantial overlap, or shared foundation with distinguishable orientation and applications.

Demographics of the Superintendency and the Professoriate

A series of 10-year studies has provided successive snapshots of the characteristics of school superintendents. For example, Bjork, Glass, and Brunner (2005) synthesized the results of the survey for the year 2000 along with that of various other studies of the superintendency, reporting that, despite variation in relation to district size and decade, superintendents continued to perceive management and instructional leadership as key competency areas. They also concluded that, on average, superintendents in the 2000 study had spent more time moving through "the chairs" than those in the 1992 study. According to the accompanying synthesis, superintendents reported general satisfaction with their preparation programs, with the primary perceived weakness being insufficient connections and applications to practice, leading to the recommendation of Bjork, Kowalski, and Browne-Ferrigno's (2005, 87) for more emphasis on "tacit knowledge (practical intelligence)." Various other sources have also concluded that communication is increasingly a core competency for successful superintendents (e.g., Kowalski and Keedy 2005).

The more recent study (Glass and Franceschini 2007) revealed the increased importance of the instructional leader competency area in terms of the school boards' hiring expectations. Other notable findings were that the proportion of females and minorities had increased to 21.7% and 6.2% respectively while white males continued to be the dominant demographic group of superintendents; and the proportion of superintendents with doctorates increased from 46% to 51% in the six years since the previous survey, with the majority being particularly predominant (i.e., more than 75%) in districts with more than 5,000 students. The responding superintendents, like those in the 2000 survey, continued to rate their preparation programs as effective or very effective, although the total percentage for these two categories together was lower for doctoral than master's level programs.

In the findings of the most recent study in this series (Kowalski et al. 2011), respondents expressed a generally high level of job satisfaction, but that only half of them expected to be in a superintendency in the year 2015. Additionally, the proportion of female superintendents had reached 24.1%. Consistent with earlier AASA studies, a substantial majority of the respondents rated their academic preparation as good (53.9 %) or excellent (24.8%). The proportion of respondents who reported having a doctoral degree (45.3%) was identical to that found in the Glass, Bjork, and Brunner (2000) study; yet, the ratings of their former professors as good or excellent was 80% compared to 65.9% in the 2000 study.

For the education leadership professoriate, following an early survey (Campbell and Newell 1973), McCarthy and her colleagues provided a corresponding series of snapshots that reveals both commonality with, and differences from, superintendents. First, for the intervening period of the later 1970s and early 1980s, McCarthy (1999) noted the development of subspecialties in education law, finance, and politics, as evidenced by the growth of specialized organizations for each of these fields. More specifically, from the survey in 1986 (McCarthy et al. 1988) to the one in 1994 (McCarthy and Kuh 1997), significant turnover in the educational leadership professoriate was found, but most of the "new breed" of faculty members were not at the research and doctoral universities (McCarthy and Kuh 1998, 361). Additionally, as McCarthy and Kuh (1998) noted, the 1994 new faculty members were far less likely than their 1986 counterparts to list research as their primary strength. Similarly, the proportion with significant experience as K-12 administrators had increased from 28% to 45%, but this priority was much less pronounced at research and doctoral universities. As a result, they observed that the most critical need cited by the largest percentage of faculty had evolved from "a more extensive knowledge base" in 1972 to "curricular reform" in 1986 to "more attention to problems of practice" in 1994. Viewing this shift to a "field sensitive" orientation as part of a historical "pendulum-like propensity in responding to criticism" (McCarthy and Kuh 1998, 368), they warned against "an unintended over-correction toward praxis" (469).

The preliminary results from the most recent survey, conducted in 2008, revealed a dramatic overall shift in the proportion of females—45% compared to 2% in 1972—and minorities—17% as compared with 3% in 1972—in the education leadership professoriate, which largely parallels the overall composition of the faculty in higher education nationally (McCarthy and Hackmann 2009). They also reported an increase from 1% to 3% in 1972 to 17% of nontenure-line faculty in educational leadership, presumably not only visiting or part-time lines but also clinical faculty increasingly referred to as "professors of practice." In terms of the faculty's listings of their primary strengths, they found a pendulum-like reverse cycle for research. (See Table.) Thus, only a minority of education leadership faculty self-reported research as a primary strength during this 36-year period, with the initial stronger emphasis in UCEA institutions re-emerging even more strongly in 2008 after a merging movement with non-UCEA member institutions at the half-way point. In contrast, there was a general decline in the faculty-respondents' listing of service/outreach as the primary strength for the faculty in both UCEA and non-UCEA programs, a trend that was even more pronounced among tenure-line faculty. One may speculate that a two-track system similar to that of clinical faculty at law schools may be developing.

Interests and Values in Professional Reading

The overlapping interests and values of superintendents and educational leadership faculty are also evident in terms of their choices of professional periodicals. More specifically, in Zirkel's (2007) comparison of the respective ratings and usage of super-intendents (Mayo and Zirkel, 2002) and educational leadership faculty (Mayo, Zirkel, and Finger 2006), both constituencies highly ranked and regularly read *Educational Leadership* and *Phi Delta Kappan*. Yet, the two groups notably differed in their other choices, with professors choosing refereed journals, such as *Educational Administration Quarterly* and the *American Education Research Journal*, and superintendents selecting practitioner magazines, such as *School*

Table
Percentage of Faculty Reporting Research
as a Primary Strength

Institutions	Year		
	1986	*1994*	*2008*
UCEA Member	24%	16%	33%
Non-UCEA Member	11%	15%	11%

Administrator and the *American School Board Journal*. Observing that "[s]uperintendents and their counterparts in academe work in different contexts, but the connections need to be strong and interactive," Zirkel (2007, 589) concluded that "if educational leadership is to become a fully realized and preeminent profession, then *Educational Leadership* or some other journal will ultimately have to become the effective equivalent of the *New England Journal of Medicine*." More recently, Goodyear et al. (2009) found that various scholars in the broad field of education perceived that only two of the eleven core journals—again, *Educational Leadership* and *Phi Delta Kappan*—had a greater effect on policy and practice than on scholarship.

Other Differences beyond the Common Core

More generally, a recent review noted the gap and tension between the perceptions of education leadership faculty and practitioners in terms of the content and delivery of preparation programs (Hackmann et al. 2009). Similarly, Murphy (1999) reported a separation and mutual suspicion between AASA and UCEA that reflected the different values and orientations of their respective constituencies. In a personal account of a professor at a nationally acclaimed school of education, who was the only former superintendent on the faculty, Davis (2007, 570-571) noted "a growing sense of disconnection" between the research and practice that he attributed to the "arrogance of academe," the careless consumerism of practitioners, and the gap in journals and language between these two groups. In an accompanying analysis, Murphy (2007, 582) suggested that "the cottage industry of criticism of administrator preparation" missed the fatal flaw of education leadership programs—the marginalization of practice. Reporting his sense of a "palpable, though quite civilized, presumption of superiority embedded in the culture of university preparation programs" (583), he urged making administrators' practice, rather than overintellectualized theory, the organizing force for such programs.

On a more abstract and indirect level, a set of position papers in the October 2008 issue of the *Educational Researcher* recognized and responded to "the Divide" (Noffke 2008, 430) between practitioner and scholar. In his paper, Labaree (2008, 421) viewed the separation as inevitable based on "the division of educational labor structured by the institutional settings, occupational constrains, daily work demands, and provisional incentives" of these two role realms. At the opposite pole, Bulterman-Bos (2008) called for a unifying approach, based on the medical model, of clinical research, which would require extensive and continuing experience in the world of practice for all research in education. Both sides recognized that the two worlds overlap rather than being mutually exclusive or coterminous. However, their polar positions have two limitations as applied to the focus here. First, each perspective was at the respective extremes of separation or integration without tailoring to the extent of commonality and difference. Second, the worldview that they both identified on the practice side is the role of classroom teacher, which is significantly different from the position of school district superintendent.

The root duality is between "academic knowledge" and "practice knowledge" (Murphy 2002, 184). As an advocate for "reculturing" the educational leadership profession, Murphy suggested alternative metaphors of moral steward (i.e., social justice), educator (i.e., school improvement), and community builder (i.e., democratic community) as providing the synthesizing paradigm. In doing so, he suggested the futility of the traditional metaphor of bridge-building as follows: "Trying to link theory and practice in school administration has been for the past 30 years a little like attempting to start a car with a dead battery. The odds are fairly long that the engine will ever turn over" (Murphy 2002,181). More comprehensively, McCarthy and Forsyth (2009, 88) elaborated the poles as "technical-rational knowledge" and "practice knowledge/artistry" while adding the mediating constructs, such as context and valuation, as a model for analyzing educational leadership preparation. These successive conceptions further reveal the commonality and differences between the professoriate and the superintendency.

Conclusion

At first glance, the current quality standards for preparation of educational leaders (e.g., Young 2011) make sense in terms of the superintendency as the chief educational leader at the local level, but stand in stark contrast to the enduring conception of the Ph.D., as "the monarch of the academic community" and as "the academy's own means of reproduction" (Shulman 2008, x-xi). For example, the common elements of intensive internships and cohort structures are obviously intended for practitioners whereas for professors the missing components are subject specializations and sophisticated research skills.

Yet, a unifying vision provides a way of harmonizing the commonalities and the differences between the practitioners, as led by the superintendents, and the professoriate, as marked by academia's doctoral degree, in education leadership. This three-part review will help inform the design debate and decisions for providing more effective doctoral programs that align more closely with overlapping but differentiating duality of these primary groups of leadership practitioners and scholars.

References

American Association of School Administrators (AASA). 1982. *Guidelines for the Preparation of School Administrators*. Arlington, VA: AASA. ED 229825.

_____. *Professional Standards for the Superintendency*. 1993. Arlington, VA: AASA.

Anderson, Dale G. 1983. "Differentiation of the Ed.D. and Ph.D. in Education." *Journal of Teacher Education* 34 (3): 55-58.

Anderson, Gary L. "Disciplining Leaders: A Critical Discourse Analysis of the ISLLC National Examination and Performance Standards in Educational Administration." 2001. *International Journal of Leadership in Education* 4 (3): 199-216.

Andrews, Richard, and Margaret Grogan. "Form Should Follow Function: Removing the Ed.D. Dissertation from the Ph.D. Straight Jacket." 2005. *UCEA Review* 46 (2): 10-13.

Baker, Bruce D., Margaret Terry Orr, and Michelle D. Young. 2007. "Academic Drift, Institutional Production, and Professional Distribution of Graduate Degrees in Educational Leadership." *Educational Administration Quarterly* 43 (3): 279-318.

Bjork, Lars G., Thomas E. Glass, and C. Cryss Brunner. 2005. "Characteristics of American School Superintendents." In *The Contemporary Superintendent: Preparation, Practice, and Development*, edited by Lars G. Bjork and Theodore J. Kowalski, 19-44. Thousand Oaks, CA: Corwin.

Bjork, Lars G., Theodore J. Kowalski, and Tricia Browne-Ferrigno. 2005. "Learning Theory and Research." In *The Contemporary Superintendent: Preparation, Practice, and Development*, edited by Lars G. Bjork and Theodore J. Kowalski, 71-106. Thousand Oaks, CA: Corwin.

Bratlien, Maynard J., Susan M. Genzer, John R. Hoyle, and Arnold D. Oates. 1992. "The Professional Studies Doctorate: Leaders for Learning." *Journal of School Leadership* 2 (1): 75-89.

Bredeson, Paul V. 2006. "Integrated Doctoral Programs in Educational Leadership: The Case for Preparing Practitioners and Researchers Together." *UCEA Review* 45 (2): 20-23.

_____. 1991. "Reappraising Personal Experience in the Preparation of School Administrators." *Journal of School Leadership* 1 (2): 176-189.

Brown, Kathleen M. 2008. *Preparing Future Leaders for Social Justice, Equity, and Excellence: Bridging Theory and Practice Through a Transformative Andragogy*. Norwood, MA: Christopher-Gordon.

_____. 2006a. "A Transformative Andragogy for Principal Preparation Programs." *UCEA Review* 45 (2): 1-5.

_____. 2006b. "Leadership for Social Justice and Equity: Evaluating a Transformative Framework and Andragogy." *Educational Administration Quarterly* 42 (5): 700-745.

Brown, Laurence D. *A Perspective on the Ph.D.-Ed.D. Discussion in Schools of Education*. 1990. Paper presented at the annual meeting of the American Educational Research Association, Boston, April. ED 320882.

Bulterman-Bos, Jacquelien A. 2008. "Will a Clinical Approach Make Education Research More Relevant for Practice?" *Educational Researcher* 37 (7): 412-420.

Cambron-McCabe, Nelda, and Martha McCarthy. 2005. "Educating School Leaders for Social Justice." *Educational Policy* 19 (1): 201-222.

Campbell, Roald F., and L. Jackson Newell. 1973. "A Study of Professors of Educational Administration." *Educational Administration Quarterly* 9 (3): 3-27.

Clifford, Geraldine J., and James W. Guthrie. 1988. *Ed School*. Chicago: University of Chicago Press.

Collier, Virginia Sue Lester. 1987. "Identification of Skills Perceived by Texas Superintendents as Necessary for Successful Job Performance." PhD diss., University of Texas at Austin. ProQuest (AAT 8717390).

CCSSO (Council of Chief State School Officers). *Educational Leadership Policy Standards: ISLLC 2008*. 2008. As Adopted by the National Policy Board for Educational Administration. Washington DC. http://www.ccsso.org/Documents/2008/Educational_Leadership_Policy_Standards_2008.pdf.

Courtenay, Bradley C. 1998. "Eliminating the Confusion over the Ed.D. and Ph.D. in Colleges and Schools of Education." *Innovative Higher Education* 13 (1): 11-20.

Davis, Stephen H. 2007. "Bridging the Gap Between Research and Practice: What's Good, What's Bad, and How Can One Be Sure?" *Phi Delta Kappa* 88 (8): 569-578.

Davis, W. Jeremy, and Dennis W. Spuck. 1978. "A Comparative Analysis of Masters, Certification, Specialist and Doctoral Programs." In *Preparatory Programs for Educational Administrators in the United States*, edited by Paula F. Silver and Dennis W. Spuck, 151-177. Columbus, OH: UCEA.

Deering, Thomas E. 1998. "Eliminating the Doctor of Education Degree: It's the Right Thing to Do." *Educational Forum* 62 (3): 243-248.

Deering, Thomas E., and Jerry Whitworth. 1982. "The Doctoral Degree in Education: The Ed.D. versus the Ph.D." *Eastern Education Journal* 15 (2): 21-24.

Dill, David D., and James L. Morrison. 1985. "Ed.D. and Ph.D. Research Training in the Field of Education: A Survey and a Proposal." *Review of Higher Education* 8 (2): 169-182.

Douglas, Betty Jo. 1990. "Skills Perceived by Tennessee Superintendents as Necessary for Successful Job Performance." EdD diss., Memphis State University. ProQuest. (AAT 9102954).

Duke, Nell K., and Sarah W. Beck. 1999. "Education Should Consider Alternative Formats for the Dissertation." *Educational Researcher* 28 (3): 31-36.

Duncan, Arne. 2009. "Teacher Preparation: Reforming the Uncertain Profession." Remarks of [U.S. Department of Education] Secretary Arne Duncan at Teachers College, Columbia University." New York, October 22. http://www.ed.gov/news/speeches/teacher-preparation-reforming-uncertain-profession.

Edgell, Wallace Alan. 1983. "Educational Administration Professors' Perceptions of the Importance of the Competencies and Related Skills in the 'Guidelines for the Preparation of School Administrators.'" PhD diss., Texas A&M University. ProQuest (AAT 8408426).

Evans, Rodney. 2007. "Comments on Shulman, Golde, Bueschel, and Garabedian: Existing Practice is Not the Template." *Educational Researcher* 36 (9): 553-559.

Everson, Susan T. 2006. "The Role of Partnerships in the Professional Doctorate in Education: A Program Application in Educational Leadership." *Educational Considerations* 33 (2): 5-9.

_____. 2009. "A Professional Doctorate in Educational Leadership: Saint Louis University's Ed.D. Program." *Peabody Journal of Education* 84 (1): 86-99.

Fluth, John Adam. 1986. "Senior High School Principals' Perception of the Importance of the Competencies and Related Skills in the 'Guidelines for the Preparation of School Administrators.'" PhD diss., Texas A&M University. ProQuest (AAT 8614952)

Friedland, Helen A., Carol Fleres, and Frederick Hill. 2007. "The Efficacy of School Leadership Programs." In *At the Tipping Point: Navigating the Course for the Preparation of Educational Administrators*. 2007 Yearbook of the National Council of Professors of Educational Administration, edited by Linda K. LeMasters and Rosemary Papa, 319-328. Lancaster, PA: Pro-Active.

Glass, Thomas E., and Louis A. Franceschini. 2007. *The State of the American Superintendency: A Mid-Decade Study*. Lanham, MD: Rowman & Littlefield.

Golde, Chris M. 2007. "Signature Pedagogies in Doctoral Education." *Educational Researcher* 36 (6): 344-351.

Goodlad, John I. 1990. *Teachers for Our Nation's Schools*. San Francisco: Jossey-Bass.

Goodyear, Rodney K., Dominic J. Brewer, Karen Symms Gallagher, Terence J. G. Tracey, Charles D. Claiborn, James W. Lichtenberg, and Bruce E. Wampold. 2009. "The Intellectual Foundations of Education: Core Journals and Their Impacts on Scholarship and Practice." *Educational Researcher* 38 (9): 700-706.

Gousha, Richard P., and Golam Mannan. 1991. *Analysis of Selected Competencies: Components, Acquisition and Measurement Perceptions of Three Groups of Stakeholders in Education*. Paper presented at the annual meeting of the National Council of Professors of Educational Administration, Fargo, ND, August. ED336850.

Guthrie, James W. 2006. "Multi-Purpose Education Doctorate No Longer Palatable." *UCEA Review* 45 (2): 24-26.

Hackmann, Donald G., Scott C. Bauer, Nelda H. Cambron-McCabe, and David M. Quinn. 2009. "Characteristics, Preparation, and Professional Development of Educational Leadership Faculty." In *Handbook of Research on the Education of School Leaders*, edited by Michelle D. Young, Gary M. Crow, Joseph Murphy, and Rodney T. Ogawa, 225-268. New York: Routledge.

Hackmann, Donald G., and William J. Price. 1995. *Preparing School Leaders for the 21st Century: Results of a National Survey of Educational Leadership Doctoral Programs*. Paper presented at the annual meeting of the National Council of Professors of Educational Administration, New Orleans. February. ERIC, ED 378863.

Hale, Elizabeth L., and Hunter N. Moorman. 2003. *Preparing School Principals: A National Perspective on Policy and Program Innovations*. Washington, DC, and Edwardsville, IL: Institute for Educational Leadership and Illinois Education Research Council.

Hoyle, John R. 2007. "A Preparation Mystery: Why Some Succeed and Others Fail." *Planning and Changing* 38 (3-4): 148-163.

_____. 1987. "The AASA Model for Preparing School Leaders." In *Approaches to Administrative Training in Education*, edited by Joseph Murphy and Philip Hallinger, 83-95. Albany, NY: State University of New York Press.

_____. 1985. "Programs in Educational Administration and the AASA Preparation Guidelines." *Educational Administration Quarterly* 21 (1): 71-93.

Hoyle, John R., Lars G. Bjork, Virginia Collier, and Thomas Glass. 2005. *The Superintendent as CEO*. Thousand Oaks: Corwin.

Hoyle, John R., Fenwick English, and Betty Steffy. 1998. *Skills for Successful 21st Century School Leaders: Standards for Peak Performance*. Arlington, VA: AASA. ERIC, ED 420911.

_____. 1985. *Skills for Successful School Leaders*. Arlington, VA: AASA. ERIC, ED 264654.

Hoyle, John R., and Mario S. Torres. 2010. *Six Steps to Preparing Exemplary Principals and Superintendents: Leadership Education at Its Best*. Lanham, MD: Rowman & Littlefield.

_____. 2008. "Exploring the Link between School Leadership Preparation and Practice." *Planning and Changing* 39 (3/4): 213-239.

Imig, David. 2011. "Dissertations or Capstones: The Conundrum Facing Ed.D. Programs." *UCEA Review* 52 (3): 12-13.

Jackson, Barbara L., and Carolyn Kelley. 2002. "Exceptional and Innovative Programs in Educational Leadership." *Educational Administration Quarterly* 38 (2): 192-212.

Kottkamp, Robert B. 2010. "Introduction: Leadership Preparation in Education." *Educational Administration Quarterly* 47 (1): 3-17.

Kowalski, Theodore J., and John L. Keedy. 2005 "Preparing Superintendents to Be Effective Communicators." In *The Contemporary Superintendent: Preparation, Practice, and Development*, edited by Lars G. Bjork and Theodore J. Kowalski, 207-226. Thousand Oaks, CA: Corwin.

Kowalski, Theodore J., Robert S. McCord, George J. Petersen, I. Phillip Young, and Noelle M. Ellerson. 2011. *The American School Superintendent: 2010 Decennial Study*. New York: Rowman & Littlefield.

Labaree, David F. 2008. "Comments on Bulterman-Bos: The Dysfunctional Pursuit of Relevance in Education Research." *Educational Researcher* 37 (7): 421-423.

Lagemann, Ellen C. 2008. "Education Research as Distributed Activity across Universities." *Educational Researcher* 37 (7): 424-428.

Leithwood, Kenneth, Karen Seashore Louis, Stephen Anderson, and Kyla Wahlstrom. 2005. *How Leadership Influences Students Learning*. New York: Wallace Foundation.

Leithwood, Kenneth, Sara Patten, and Doris Jantzi. 2010. "Testing a Conception of How School Leadership Influences Student Learning." *Educational Administration Quarterly* 46 (5): 671-706.

Lem, Robert W. 1989. "Perceptions of School Administrators and Professors of Educational Administration Regarding the Relevance and Effectiveness of Administrator Preparation Programs." EdD diss., University of San Francisco. ProQuest (AAT 9023841)

Levine, Arthur. 2005. *Educating School Leaders*. Washington, DC: Education Schools Project.

Levine, Arthur, and Diane R. Dean. 2007. "Deleting the Doctorate (and Other Vestiges of Outmoded Preparation)." *School Administrator* 7 (64): 10-13.

Malen, Betty, and Nona A. Prestine. 2005. "The Case for Revitalizing the Dissertation." *UCEA Review* 44 (2): 7-9.

Maranto, Robert, Gary Ritter, and Arthur E. Levine. 2010. "The Future of Ed. Schools." *Education Week*, January 6. http://www.edweek.org/ew/articles/2010/01/06/16maranto_ep.h29.html.

Marshall, Mark E., and William A. Spencer. 1999. *Public School Administrator Competencies: A Comparison of the Perceptions of Stakeholders in Alabama*. Paper presented at the annual meeting of the Mid-South Educational Research Association, Boston, November. ERIC, ED 439516.

Mawhinney, Hanne, and Michelle D. Young. 2010. "Anchoring the ELCC Standards to the Existing Research Base in Educational Leadership." *UCEA Review* 51 (3): 9.

Mayo, C. Russell, Perry A. Zirkel, and Brian A. Finger. 2006. "Which Journals Are Education Leadership Professors Choosing?" *Educational Administration Quarterly* 42 (5): 806-811.

Mayo, C. Russell, and Perry A. Zirkel. 2002. "School Superintendents' Choices of Professional Periodicals." *Journal of School Leadership* 12 (4): 397-410.

McCarthy, Martha M. 2002. "Educational Leadership Preparation Programs: A Glance at the Past with an Eye to the Future." *Leadership and Policy in the Schools* 1 (3): 201-221.

_____. 1999. "The Evolution of Educational Leadership Preparation Programs." In *Handbook of Research on Educational Administration*, edited by Joseph Murphy and Karen Seashore Louis, 119-139. Thousand Oaks, CA: Corwin.

McCarthy, Martha M., and Patrick B. Forsyth. 2009. "An Historical Review of Research and Development Activities Pertaining to the Preparation of School Leaders." In *Handbook of Research on the Education of School Leaders*, edited by Michelle D. Young, Gary M. Crow, Joseph Murphy and Rodney T. Ogawa, 1-22. New York: Routledge,.

McCarthy, Martha M., and Donald G. Hackmann. 2009. "Demographic Trends in the Educational Leadership Professoriate, 1972-2008." *UCEA Review* 50 (2): 24-26.

McCarthy, Martha M., and George D. Kuh. 1997. Continuity and Change: The Educational Administration Professoriate. Columbia, MO: University Council of Educational Administration.

McCarthy, Martha M., and George D. Kuh. 1998. "A New Breed of Educational Leadership Faculty Members." *Journal of School Leadership*, 8 (4): 360-373.

McCarthy, Martha M., George D. Kuh, Jackson Newell, and Carla M. Iaconi. 1998. *Under Scrutiny: The Educational Administration Professoriate*. Tempe, AZ: University Council of Educational Administration.

McClellan, Michael Joseph. 1984. "A National Study of Public School Superintendents' Perceptions of the Relevancy of the Competencies and Related Skills Outlined in the 'Guidelines for the Preparation of School Administrators.'" PhD diss., Texas A&M University. ProQuest (AAT 8428789)

Moore, Harold Emerson, John H. Russel, and Donald G. Ferguson. 1990. *The Doctorate in Education: An Inquiry into the Conditions Affecting Pursuit of the Doctoral Degree in the Field of Education*, vol. 2. Washington, DC: American Association of Colleges for Teacher Education.

Murphy, Jerome T. 2006a. "Dancing Lessons for Elephants: Reforming Ed School Leadership Programs." *Phi Delta Kappan* 87 (7): 489-491.

_____. 2006b. "Elephants or Dinosaurs? A Call to Action for Ed Schools." *Phi Delta Kappan* 87 (7): 529-535.

_____. 1991. "Superintendents as Saviors: From the Terminator to Pogo." *Phi Delta Kappan* 73 (7): 507-513.

Murphy, Joseph. 2007. "Questioning the Core of University-Based Programs for Preparing School Leaders." *Phi Delta Kappan* 88 (8): 582-584.

_____. 2005. "Unpacking the Foundations of ISLLC Standards and Addressing Concerns in the Academic Community." *Educational Administration Quarterly* 41 (1): 154-191.

_____. 2002. "Reculturing the Profession of Educational Leadership: New Blueprints." *Educational Administration Quarterly* 38 (2): 176-191.

_____. 1999. "The Reform of the Profession: A Self Portrait." In *Educational Administration: A Decade of Reform*, edited by Joseph Murphy and Patrick B. Forsyth, 39-68. Thousand Oaks, CA: Corwin.

Murphy, Joseph, and Michael Vriesenga. 2005. "Developing Professionally Anchored Dissertations: Lessons from Innovative Programs." *School Leadership Review* 1 (1): 33-57.

Murphy, Joseph, Michelle D. Young, Gary M. Crow, and Rodney T. Ogawa. 2009. "Introduction: Exploring the Broad Terrain of Leadership Preparation Programs in Education." In *Handbook of Research on the Education of School Leaders*, edited by Michelle D. Young, Gary M. Crow, Joseph Murphy, and Rodney T. Ogawa, 1-22. New York: Routledge.

National Policy Board for Educational Administration (NBPEA). 2010. *Educational Leadership Program Standards: ELCC District-Level Standards*. Charlottesville, VA. http://www.npbea.org/projects.php.

_____. 2008. *National Leadership Standards*. Charlottesville, VA. http://www.npbea.org/projects.php.

_____. 2002. *Standards for Advanced Programs in Educational Leadership*. Charlottesville, VA. http://www.ncate.org/public/programStandards.asp?ch=4#ELCC.

_____. 1989. *Improving the Preparation of School Administrators: An Agenda for Reform*. Charlottesville, VA.

Nelson, Jack K., and Callen Coorough. 1994. "Content Analysis of the PhD Versus Ed.D. Dissertation." *Journal of Experimental Education* 62 (2): 158-168.

Noffke, Susan E. 2008. "Comments on Bulterman-Bos: Research Relevancy or Research for Change?" *Educational Researcher* 37 (7): 429-431.

Norton, M. Scott. 1992. "Doctoral Studies of Students in Educational Administration Programs in Non-Member UCEA Institutions." *Educational Considerations* 20 (1): 37-41.

Norton, M. Scott, and Frederick D. Levan. 1987. "Doctoral Studies of Students in Educational Administration Programs in UCEA Member Institutions." *Educational Considerations* 14 (1): 21-24.

Olson, Kate, and Christopher M. Clark. 2009. "A Signature Pedagogy in Doctoral Education: The Leader-Scholar Community." *Educational Researcher* 38 (3): 216-221.

Orr, Margaret Terry. 2007. "The Doctoral Debate." *School Administrator* 64 (7): 16-19.

_____. 2006. "Mapping Innovation in Leadership Preparation in Our Nation's Schools of Education." *Phi Delta Kappan* 87 (7): 492-500.

Osguthorpe, Russell T., and Mei Jiuan Wong. 1993. "The Ph.D. Versus the Ed.D.: Time for a Decision." *Innovative Higher Education* 18 (1): 47-63.

Perry, Jill Alex. 2011. "The Carnegie Project on the Education Doctorate: Phase II—A Quest for Change." *UCEA Review* 52 (3): 1-4.

Ramirez, Al, Dick Carpenter, and Nadyne Guzman. 2007. "Analysis of Superintendent Search Criteria and the Standards for Administration Preparation Programs." In LeMasters and Papa, *At the Tipping Point*, 266-278.

Ray, Huba Ann. 2003. "ISLLC Administrator Competencies: A Comparison of Perceptions among Superintendents, School Board Presidents, and Principals." EdD diss., St. Louis University. ProQuest (AAT 3102922)

Redden, Elizabeth. 2007. "Envisioning a New Ed.D." *Inside Higher Education*, April 10. http://www.insidehighered.com/news/2007/04/10/education.

Roach, Virginia, L. Wes Smith, and James Boutin. 2010. "School Leadership Policy Trends and Developments: Policy Expediency or Policy Excellence." *Educational Administration Quarterly* 47 (1): 71-113.

Robertson, Neville, and Jack K. Sistler. 1971. *The Doctorate in Education*. Bloomington, IN: Phi Delta Kappa.

Robinson, Viviane M. J., Claire A. Lloyd, and Kenneth J. Rowe. 2008. "The Impact of Leadership on Student Outcomes: An Analysis of the Differential Effects of Leadership Types." *Educational Administration Quarterly* 44 (5): 635-674.

Sass, Michael William. 1989. "The AASA Performance Goal and Skill Areas' Importance to Effective Superintendency Performance as Viewed by Professors of Educational Administration and Practicing Superintendents." EdD diss., Northern Illinois University. ProQuest (AAT 9006076)

Schneider, Barbara L., Laurence Brown, Terry Denny, B. Claude Mathis, and William Schmidt. 1984. "The Deans' Perspective on the Status of Doctoral Programs in Schools of Education." *Phi Delta Kappan* 65 (9): 617-620.

Sclafani, Susan B. Krouner. 1987. "AASA Guidelines for the Preparation of School Administrators: Do They Represent the Important Job Behaviors of Superintendents?" PhD diss., the University of Texas at Austin. ProQuest (AAT 8717530)

Sergiovanni, Thomas. 1989. "Mystics, Neats, and Scruffies: Informing Professional Practice in Educational Administration." *Journal of Educational Administration* 27 (2): 7-21.

Shulman, Lee S. 2008. Foreword to *The Formation of Scholars: Rethinking Doctoral Education for the Twenty-First Century*, by George E. Walker, Chris M. Golde, Laura Jones, Andrea Conklin Bueschel, and Pat Hutchings. San Francisco: Jossey-Bass.

_____. 2007. "Response to Comments: Practical Wisdom in the Service of Professional Practice." *Educational Researcher* 36 (9): 560-563.

_____. 2005. "Signature Pedagogies in the Professions." *Daedalus* 134 (3): 52-59.

_____. 2004. *A New Vision for the Doctorate in Education*. Symposium conducted at the annual meeting of the American Educational Research Association, San Diego, April.

Shulman, Lee S., Chris M. Golde, Andrea Conklin Bueschel, and Kriten J. Garabedian. 2006. "Reclaiming Education's Doctorates: A Critique and a Proposal." *Educational Researcher* 35 (3): 25-32.

Sternberg, Robert J., and Richard K. Wagner, eds. 1986. *Practical Intelligence: Nature and Origins of Competence in the Everyday World*. Cambridge, UK: Cambridge University Press.

Toye, Carla, Rolf Blank, Nancy Sanders, and Andra Williams. 2007. *Key State Education Policies on P-12 Education, 2006: Results of a 50-State Survey*. Washington, DC: Council of Chief State School Officers.

U.S. Department of Education. 2005. *Innovative Pathways to School Leadership*. Washington, DC.

Viadero, Debra. 2008. "Study Will Size Up Doctoral Programs for Education Researchers." *Education Week*, October 24. http://www.edweek.org/ew/articles/2008/10/24/10aera.h28.html.

Voelter, Donald Edward. 1985. "Perceptions of Selected Junior/Community College Administrators Regarding the Relevancy of the Competencies and Related Skills of the 'Guidelines for the Preparation of School Administrators.'" PhD diss., Texas A&M University. ProQuest (AAT 8514304)

Ward, Geoffrey C. 1989. *A First-Class Temperament: The Emergence of Franklin Delano Roosevelt*. New York: Harper & Row.

Waters, J. Timothy, and Robert J. Marzano. 2006. "School District Leadership that Works: The Effect of Superintendent Leadership on Student Achievement." Working Paper. Denver, CO: Mid-continent Research for Education and Learning, September. https://www.cosa.k12.or.us/downloads/profdev/SuperintendentLeadership.pdf.

Young, Michelle D. 2011. "From the Director: Quality Assurance in Educational Leadership Preparation." *UCEA Review* 52 (2): 6-9.

_____. 2008. "Programs Are Making Progress in Providing Research-Based Preparation that Supports Student Learning." *UCEA Review* 50 (2): 9-10.

_____. 2006. "From the Director: The M.Ed., Ed.D. and Ph.D. in Educational Leadership." *UCEA Review* 45 (2): 6-9.

Young, Michelle D., Gary M. Crow, Margaret Terry Orr, Rodney T. Ogawa, and Ted Creighton. 2005. "An Education Look at 'Education School Leaders.'" *UCEA Review* 47 (2): 1-4.

Zirkel, Perry A. 2007. "The Professoriate, the Practitioners, and 'Their' Periodicals." *Phi Delta Kappan* 88 (8): 586-588.

ISSUES 1990-2012

Educational Considerations is a leading peer-reviewed journal in the field of educational leadership. Since 1990, *Educational Considerations* has featured outstanding themes and authors related to leadership:

SPRING 1990: a theme issue devoted to public school funding.
Edited by David C. Thompson, Codirector of the UCEA Center for Education Finance at Kansas State University and Board of Editors of *Educational Considerations.*

FALL 1990: a theme issue devoted to academic success of African-American students.
Guest-edited by Robbie Steward, University of Kansas.

SPRING 1991: a theme issue devoted to school improvement.
Guest-edited by Thomas Wicks & Gerald Bailey, Kansas State University.

FALL 1991: a theme issue devoted to school choice.
Guest-edited by Julie Underwood, University of Wisconsin-Madison and member of the Editorial Advisory Board of *Educational Considerations.*

SPRING 1992: a general issue devoted to philosophers on the foundations of education.

FALL 1992: a general issue devoted to administration.

SPRING 1993: a general issue devoted to administration.

FALL 1993: a theme issue devoted to special education funding.
Guest-edited by Patricia Anthony, University of Massachusetts-Amherst and member of the Editorial Advisory Board of *Educational Considerations*

SPRING 1994: a theme issue devoted to analysis of funding education.
Guest-edited by R. Craig Wood, Codirector of the UCEA Center for Education Finance at the University of Florida and member of the Editorial Advisory Board of *Educational Considerations.*

FALL 1994: a theme issue devoted to analysis of the federal role in education funding.
Guest-edited by Deborah Verstegen, University of Virginia and member Editorial Advisory Board of *Educational Considerations.*

SPRING 1995: a theme issue devoted to topics affecting women as educational leaders.
Guest-edited by Trudy Campbell, Kansas State University.

FALL 1995: a general issue devoted to administration.

SPRING 1996: a theme issue devoted to topics of technology innovation.
Guest-edited by Gerald D. Bailey and Tweed Ross, Kansas State University.

FALL 1996: a general issue of submitted and invited manuscripts on education topics.

SPRING 1997: a theme issue devoted to foundations and philosophy of education.

FALL 1997: first issue of a companion theme set (Fall/Spring) on the state-of-the-states reports on public school funding.
Guest-edited by R. Craig Wood, University of Florida, and David C. Thompson, Kansas State University.

SPRING 1998: second issue of a companion theme set (Fall/Spring) on the state-of-the-states reports on public school funding.
Guest-edited by R. Craig Wood, University of Florida, and David C. Thompson, Kansas State University.

FALL 1998: a general issue on education-related topics.

SPRING 1999: a theme issue devoted to ESL and Culturally and Linguistically Diverse populations.
Guest edited by Kevin Murry and Socorro Herrera, Kansas State University.

FALL 1999: a theme issue devoted to technology.
Guest-edited by Tweed Ross, Kansas State University.

SPRING 2000: a general issue on education-related topics.

FALL 2000: a theme issue on 21st century topics in school funding.
Guest edited by Faith Crampton, Senior Research Associate, NEA, Washington, D.C.

SPRING 2001: a general issue on education topics.

FALL 2001: a general issue on education funding.

SPRING 2002: a general issue on education-related topics.

FALL 2002: a theme issue on critical issues in higher education finance and policy.
Guest edited by Marilyn A. Hirth, Purdue University.

SPRING 2003: a theme issue on meaningful accountability and educational reform.
Guest edited by Cynthia J. Reed, Auburn University, and Van Dempsey, West Virginia University.

ISSUES 1990-2012 continued

FALL 2003: a theme issue on issues impacting on higher education at the beginning of the 21st century.
Guest edited by Mary P. McKeown-Moak, MGT Consulting Group, Austin, Texas.

SPRING 2004: a general issue on education topics.

FALL 2004: a theme issue on issues relating to adequacy in school finance.
Guest edited by Deborah A. Verstegen, University of Virginia.

SPRING 2005: a theme issue on reform of educational leadership preparation programs.
Guest edited by Michelle D. Young, University of Missouri; Meredith Mountford, Florida Atlantic University; and Gary M. Crow, The University of Utah.

FALL 2005: a theme issue on reform of educational leadership preparation programs.
Guest edited by Teresa Northern Miller, Kansas State University.

SPRING 2006: a theme issue on reform of educational leadership preparation programs.
Guest edited by Teresa Northern Miller, Kansas State University.

FALL 2006: a theme issue on the value of exceptional ethinic minority voices.
Guest edited by Festus E. Obiakor, University of Wisconsin-Milwaukee.

SPRING 2007: a theme issue on educators with disabilities.
Guest edited by Clayton E. Keller, Metro Educational Cooperative Service Unit, Minneapolis, Minnesota, and Barbara L. Brock, Creighton University.

FALL 2007: a theme issue on multicultural adult education.
Guest edited by Jeff Zacharakis and Gabriela Díaz de Sabatés, Kansas State University, and Dianne Glass, Kansas Department of Education.

SPRING 2008: a general issue on education topics.

FALL 2008: a general issue on education topics.

SPRING 2009: a theme issue on educational leadership voices from the field.
Guest edited by Michele Acker-Hocevar, Washington State University,Teresa Northern Miller, Kansas State University, and Gary Ivory, New Mexico State University.

FALL 2009: a theme issue on leadership theory and beyond in various settings and contexts.
Guest edited by Irma O'Dell and Mary Hale Tolar, Kansas State University.

SPRING 2010: a theme issue on the administrative structure of online education.
Guest edited by Tweed W. Ross, Kansas State University.

FALL 2010: a theme issue on educational leadership challenges in the 21st century.
Guest edited by Randall S. Vesely, Indiana University-Purdue University at Fort Wayne.

SPRING 2011: a theme issue on the National Council for Accreditation of Teacher Education (NCATE) Standard 4 – Diversity.
Guest edited by Jeff Zacharakis and Joelyn K. Foy, Kansas State University.

FALL 2011: a special issue on class size and student achievement.
Guest authored by James L. Phelps, former Special Assistant to Governor William Milliken of Michigan and Deputy Superintendent of the Michigan Department of Education.

SPRING 2012: a special issue of selected papers from the inaugural National Education Finance Conference (NEFC) held in 2011.

Subscribe *TODAY!*

to

educational considerations

Educational Considerations *is a leading peer-reviewed journal in the field of educational leadership and policy.*

Educational Considerations *is published twice yearly by the College of Education at Kansas State University.*

Educational Considerations *invites subscribers for only $13.00 annually. Subscribers receive paper copy and electronic copy.*

OR

Save 20% on the regular subscription price when you select electronic copy only!

✱✱✱ ORDER FORM ✱✱✱

Please send me:

❏   Paper copy $13.00 (electronic copy included at no extra cost) for one year subscription

❏   Electronic copy only $10.40 for one year subscription

Name _____

Address _____

_____

City _____ State _____ Zip _____

Make checks payable to *Educational Considerations*.
Mail with order form to:
Editor, Educational Considerations, Bluemont Hall, 1100 Mid-Campus Drive,
Kansas State University, Manhattan, KS 66506

Visit Us Online at:
http://coe.ksu.edu/EdConsiderations/