Educational Considerations

Volume 40 | Number 1 Article 3

9-1-2012

A Theory of School Achievement: A Quantum View

James L. Phelps

Recommended Citation

Phelps, James L. (2012) "A Theory of School Achievement: A Quantum View," *Educational Considerations*: Vol. 40: No. 1.
https://doi.org/10.4148/0146-9282.1077

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

http://newprairiepress.org/edconsiderations/vol40/iss1/3
DOI: 10.4148/0146-9282.1077

X. A Mathematical Model of School Achievement
XI. Changing the Paradigm

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

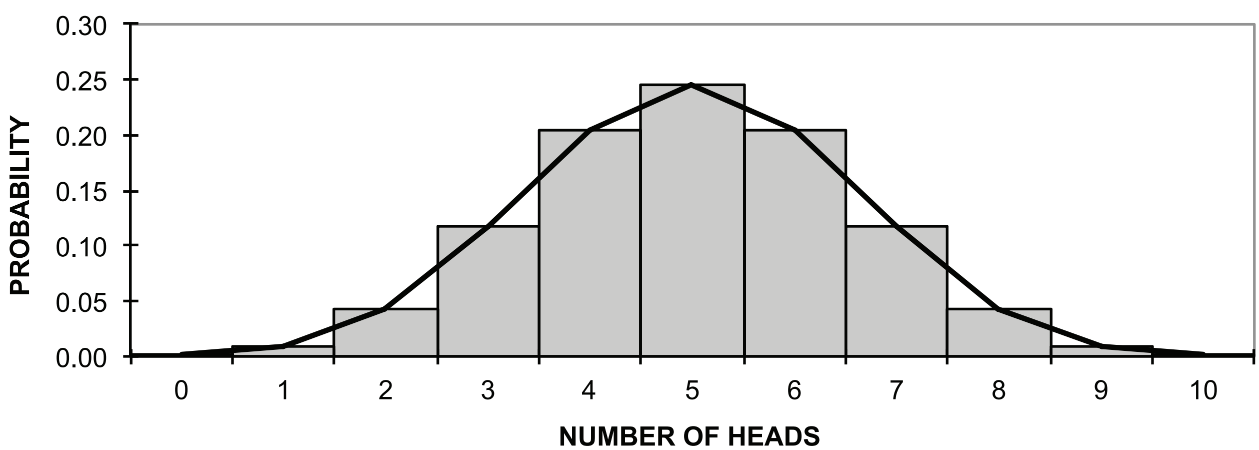
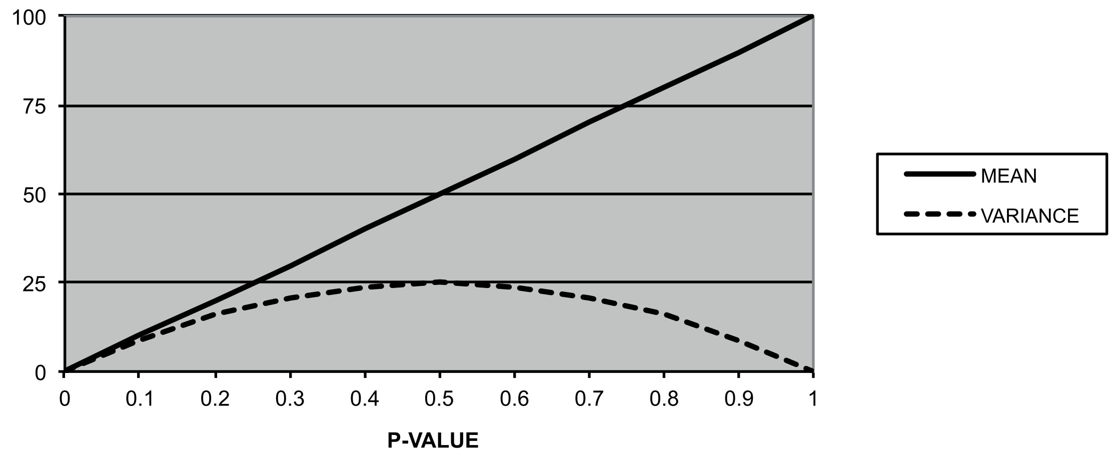


Figure 2
Effect of Changing P-Values on Binomial Statistics
(n=100)



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.

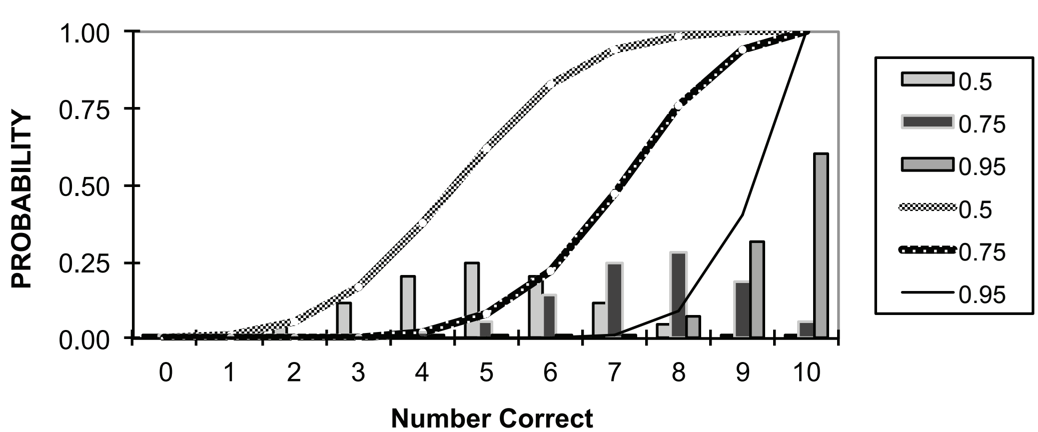
As Newton's and Leibniz's calculus became more sophisticated, a solution emerged. The concept is straightforward although the mathematics is rather sophisticated. Each point of the histogram (e.g., see Figure 1) is converted to an X-value with the height of the histogram as the Y-value. Once this step 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]

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

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



DOI: 10.4148/0146-9282.1077

Again relying on calculus, the X-values and Y-values were integrated (summed) from minus infinity to positive infinity.

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

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

$y = A\ e\^(-x^2/(2\delta^2))$

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:

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

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.

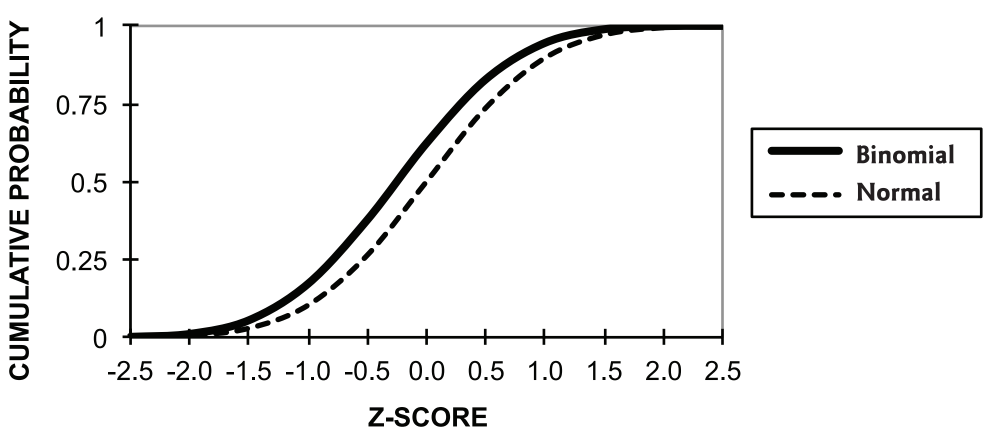
Table
Changes as Z-Scores Increase

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



Postulate 2: There is a point at the upper and lower extremes where the probability of an increase or decrease for all practical purposes is 0. Z-scores of ± 3 are used in the remainder of this article as the cut-off points. The upper and lower limits are consistent with the binomial distribution.

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.

Postulate 3: The cumulative normal probability curve for any Z-score represents duality of (1) the probability and (2) the percentile ranking. Inherent in the duality are the upper and lower limits of 1 and 0.

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.

Postulate 4: To estimate relationships and probabilities, achievement and explanatory variables must be measured as Z-scores.

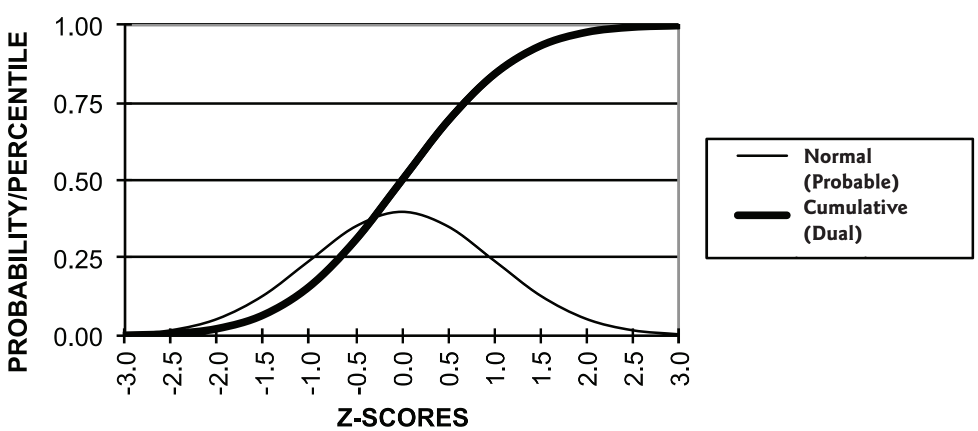
Some basic descriptive statistics are required for statistical analysis: the mean—the center point of the distribution; the variance—the area parameter of the normal curve; and the standard deviation (the square root of the variance)—the width parameter. With these three parameters—mean ($\underline{X}$), 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



http://newprairiepress.org/edconsiderations/vol40/iss1/3 6
DOI: 10.4148/0146-9282.1077

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.

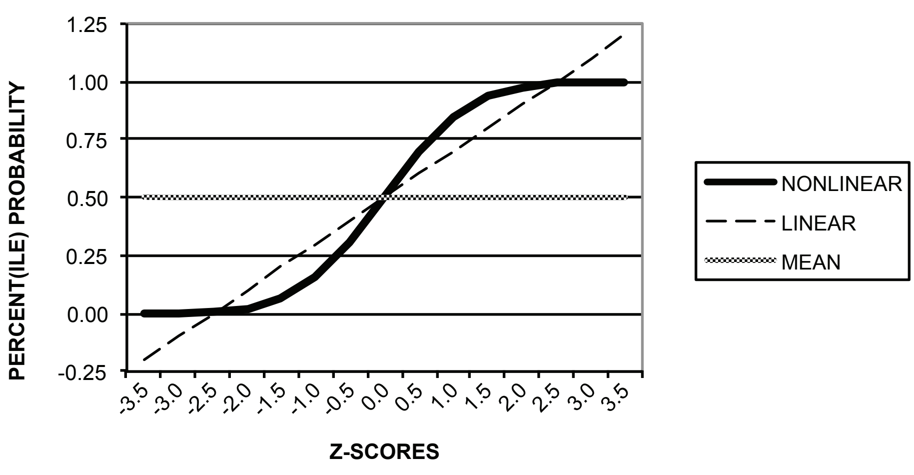
Distinction between Linear and Nonlinear Interpretations

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

http://newprairiepress.org/edconsiderations/vol40/iss1/3
DOI: 10.4148/0146-9282.1077

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*β) 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*β) 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*β), each variable represents the unique contribution to the total explanation of achievement.

> Postulate 13: Because each variable is based on the variance (r*β), 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

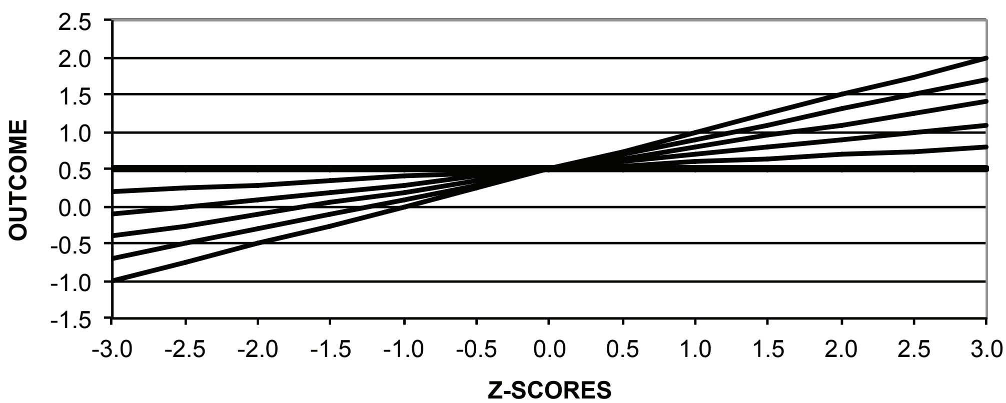
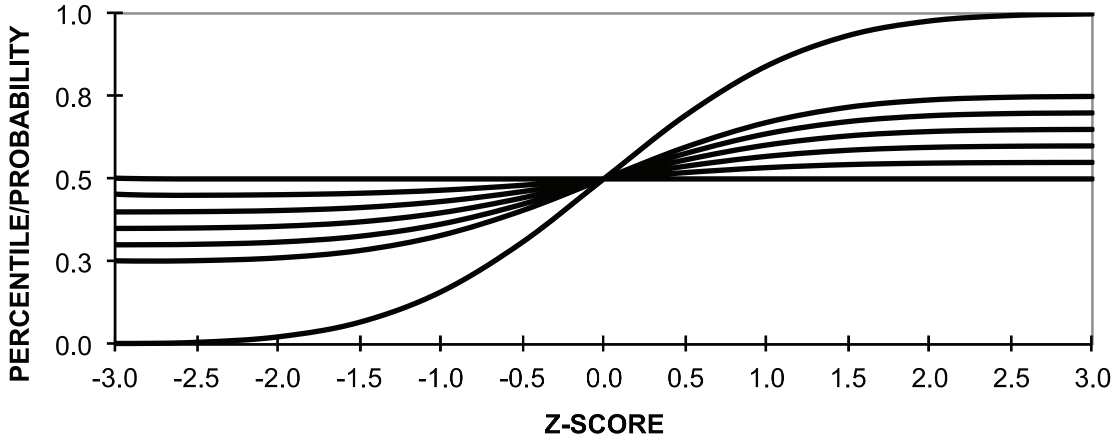


Figure 8
Explained Variance with Various Variances, Moving Upward from 0 to 1.0



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.

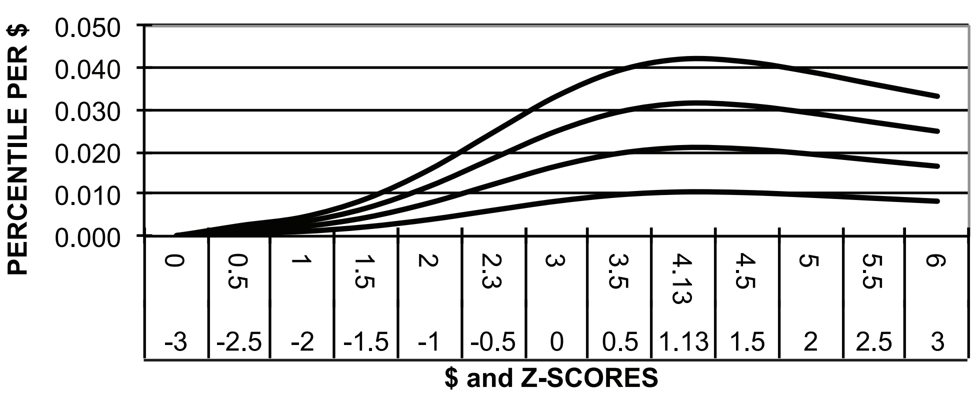
When cost-effectiveness is considered, the difference between the linear and nonlinear interpretations becomes even more striking. Once the most cost-effective variable is identified for the linear interpretation, there is every reason to invest all available funds into that single variable. In contrast, 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.

> <u>Postulate 14</u>: 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



DOI: 10.4148/0146-9282.1077

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:

$$A_i = \beta * SES_i + e \quad \textit{or} \quad A_i = (\beta{*}r) * SES_i + e$$

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:

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

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

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:

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

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

> 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.

http://newprairiepress.org/edconsiderations/vol40/iss1/3 12
DOI: 10.4148/0146-9282.1077

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.

- A 1994 study by Hedges, Laine, and Greenwald using several explanatory variables including funding. 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.
- Walberg's 1984 study of explanatory variables in the categories of instruction, curriculum, organization, homework, and time. 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.

http://newprairiepress.org/edconsiderations/vol40/iss1/3
DOI: 10.4148/0146-9282.1077

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**	**Effective-ness**	**Staff Quantity**	**Staff Character-istics**	**Instruc-tional Materials**	**Instruction**	**Curriculum**	**Time**	**Other**	**Predicted Row Sum**	**Sum of Factor R^2**
1 =	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	Σ	ΣR^2
2 =	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	Σ	ΣR^2
N =	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	Σ	ΣR^2
Factor Sum	Σ	Σ	Σ	Σ	Σ	Σ	Σ	Σ	Σ	**$\Sigma\Sigma$**	
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 =	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	Σ	
2 =	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	Σ	
N =	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	$R^2f(z)$	Σ	
Optimized Factor Sum	Σ	Σ	Σ	Σ	Σ	Σ	Σ	Σ	Σ	**$\Sigma\Sigma$**	**Target Cost**
Optiized Cost	n/a	n/a	$	$	$	$	$	$	$	**Σ $**	**$$**
Change in Achievement	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	**$\Sigma\Delta$**	

- 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-

http://newprairiepress.org/edconsiderations/vol40/iss1/3
DOI: 10.4148/0146-9282.1077

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.

http://newprairiepress.org/edconsiderations/vol40/iss1/3
DOI: 10.4148/0146-9282.1077