Modeling Achievement by Measuring the Enacted Instruction

JOHN R. WALKUP*
BEN S. JONES*

* The Standards Company LLC, 575 W. Alluvial Ave., Clovis, CA 93611

_____

This article presents a mathematical algorithm that relates student achievement with directly observable, quantifiable teacher and student behaviors, producing a modified form of the Walberg model. The algorithm (1) expands the measurable factors that comprise the quality of instruction in a linear basis of research-based teaching components and techniques and (2) incorporates the quantity of instruction that is related to the timing of classroom processes. The role such a model can play in education reform and potential research is also discussed, as well as its potential use in classroom observation rubrics programmed on mobile devices. Researchers should use the algorithm to shift reform efforts toward the improvement of input processes (e.g., use of classroom time and use of effective questioning strategies) rather than output processes (e.g., assessments). (Contains 22 references.)

Keywords: enacted instruction, classroom productivity index, classroom observation, time on task, academic engagement time, Walberg productivity model

_____

SOURCE FILE: http://www.standardsco.com/publications/

1.   INTRODUCTION

The recent focus on school accountability has renewed interest in classroom observations, especially the manner in which observation results can drive data-based decision making. Increased use of mobile applications now allows classroom observers to collect immense real-time data, which researchers can disaggregate and analyze to gauge teacher and school effectiveness. Unfortunately, the numerous factors influencing student achievement has produced non-standardized observation rubrics, confounding attempts to coalesce data collected by different research teams. Even worse, much of the vast quantities of data cannot drive teacher professional development because it correlates to factors beyond the control of the teacher. The productivity index presented in this article attempts to remedy this situation by establishing (1) criteria for selecting factors that pose the best chance of driving classroom

reform, (2) the conditions under which a researcher would credit each factor during classroom observations, and (3) a mathematical formula that can not only serve as a measure of classroom productivity but can also be programmed in a mobile application in a straightforward manner.

2. WALBERG PRODUCTIVITY MODEL

Productivity models that mathematically connect student achievement and environmental factors typically fall into two major categories: a sum of terms or a product of terms. Referring to the product-of-terms variety, education researchers have focused much of their attention on the Walberg model,

$$Q = aQ_L^j Q_N^k E, \tag{1}$$

a variation of the Cobb-Douglas function used in economics [1928], where Q_L measures the quality of instruction that a teacher exhibits, Q_N measures the quantity of instruction a student receives, and E encompasses various external influences such as classroom environment, student motivational/aptitude, and so on [Walberg 2004, McRel 2000, Walberg 1980]. The constant a scales the productivity Q between 0 and 1. As with the Cobb-Douglas function, the (non-integral) exponents weight the contribution of each term to the overall function.

In some situations, expressing the index as a product of terms provides a distinct advantage over the sum-of-terms variety: a zero value for any coefficient produces a zero value for the index. For instance, the condition $Q_N = 0$ indicates that teachers never provided their students any time to learn; the condition $Q_L = 0$ indicates that teachers delivered utterly worthless lessons. In either case, an overall $Q = 0$ index likely describes the situation fairly adequately.

Although we could examine the condition $E = 0$, this article dismisses use of this term for reasons we will now discuss.

3. CHOOSING APPROPRIATE FACTORS

In his 1980 article on educational productivity, Walberg identified a troublesome aspect of using mathematical functions to measure pedagogy:

> "If true experiments cannot be done, with random manipulation of the factors, then the best possible hope for causal inference is to include all possible causes in a regression." [Walberg 1980]

The worth of any productivity model, however, requires the ability to model student achievement, predict results, and suggest changes for raising student achievement—a productivity model that incorporates intangible factors outside the control of educational agencies may offer value as a theoretical construct, but poses little practical use.

To establish baseline measurements and set agendas for professional development—perhaps the primary vehicle for school reform—researchers need to acquire useful data. Not all data equally serves this purpose. To simplify the Walberg model for maximal effectiveness, we choose to incorporate factors based on four qualities:

1. Alterable—teachers seeking to improve their own processes have little need for factors that lie outside that which they can directly control, even if they greatly influence student achievement. For example, the E term in the Walberg model incorporates a variable pertaining to the student's home environment [Walberg 1980, McRel 2000]. Although the home lives of students affect how well they perform academically, teachers can do little to improve this situation. As another example, Walberg included the length of the school year in the variable *amount of time* [Walberg 1980, McRel 2000]; the length of the school year impacts student achievement to some extent, but school calendars remain largely in control of state legislatures and local school boards, not individual teachers.

2. Precise—even when directly controllable by a school or teacher, some factors fail to provide measurable targets. The environmental climate of a school serves as one example [Schereens et al. 1996]. Certainly, teachers and administrators can affect school climate, but no quantifiable ideal climate exists that can serve as a useful baseline or target. In more mathematical terms, we cannot normalize factors that feature no distinct upper or lower numerical bound.

3. Independent—we chose variables that researchers can measure independently with respect to each other. Two commonly cited variables, *pressure to achieve* and *leadership*, demonstrate this need [McRel 2000]. The factors that influence *pressure to achieve* and *leadership* overlap, so measuring one variable inadvertently measures (to some extent) the other variable. This superfluous "mixing" of measurements conceals the impact of each measurement on the overall productivity index.

4. Potent—focusing reform efforts around variables that drive student achievement maximizes the effectiveness of professional development. Therefore, we chose variables that correspond to conditions that *directly* affect learning. Consider the variable *strength of leadership*, identified by Weber [1971] as correlating strongly to student achievement. Strong leadership can certainly correlate to high test scores, but strong leadership in itself does not teach academic content.

Although recognizing that numerous factors affect student achievement, those incorporated in the *E* term of Equation (1) suffer from a school reform viewpoint in that they fail to meet all four criteria.

4. MODIFICATION OF THE WALBERG MODEL

With the above criteria in mind, we retain the Q_L and Q_N factors related to the quality and quantity of instruction but drop the *E* term to produce the classroom productivity index

$$Q = Q_L^j Q_N^k \tag{2}$$

that quantifies the relationship between student achievement and the enacted instruction. Walberg previously noted the importance of focusing on Q_L and Q_N, but retained the variable *home environment* for consideration because "it influences the large amounts of time students spend outside school and because it can be affected by outreach programs," [Walberg 2004] although not necessarily the teacher. However, we drop *home environment* from consideration because it satisfies only one (*potent*) of the four selection criteria. (We will later scale Q_L and Q_N between 0 and 1, therefore negating the need for the scaling constant *a*.)

The optimal values of the exponents in Equation (2) would produce the highest correlation to state exam achievement or district benchmark assessments. Expanding Equation (2) as a sum of logarithms, we find

$$\log Q = j \log Q_L + k \log Q_N. \tag{3}$$

With numerical data from future research, the familiar method of multiple regression using logarithmically-scaled values for each factor in Equation (3) would generate the expansion coefficients *j* and *k*.

In this modified form of the Walberg model, both Q_L and Q_N involve measurements of the enacted instruction, that is, the teacher behaviors and classroom environment actually experienced by students. Measuring the enacted instruction requires direct classroom observation. Adhering to our criteria for selecting factors to incorporate in the modified Walberg model, we redefine Q_L and Q_N according to the following:

1. Quality of instruction (Q_L)—the frequency in which teachers employ effective, research-based instructional techniques.

2. Quantity of instruction (Q_N)—the percentage of classroom time in which students mentally engage academic content.

Although we will describe both factors in more detail shortly, we should note that these two instructional factors measure the behaviors of different people—teachers in the case of Q_L and students in the case of Q_N. Classroom observations can focus on measuring one of the two factors, all the while ignoring the other; therefore, Q_L and Q_N correspond to independent measurements, a key requirement as noted previously.

5. MEASURING QUALITY OF TIME

To express the functional form of Q_L, we must first explain the manner in which we distinguish instructional components from instructional techniques. As defined in this article, instructional *components* refer to portions of a lesson devoted to a particular activity; teachers use instructional *techniques* during an instructional component to enhance its effectiveness. Throughout this article, we designate instructional components with the variable C and instructional components with T.

We linearly expand Q_L in terms of these lesson components,

$$Q_L = c_1 C_1 + c_2 C_2 + c_3 C_3 + \cdots + c_n C_n,$$ (4)

where each C_i refers to a specific instructional component and each c_i represents its expansion coefficient. The expansion coefficients c_i sum to 1 to normalize the expansion.

6. ILLUSTRATIVE EXAMPLE

Although only one of myriad instructional models, the direct-instruction model of Madeline Hunter [1982] serves as a useful application of Equation (4). Although variations abound, a Madeline Hunter lesson typically comprises seven steps:

1. Lesson objective (C_1)—the lesson informs the students of what they will learn.

2. Relevance (C_2)—the lesson teaches the relevance or importance of learning the lesson.

3. Activation of prior knowledge (C_3)—the lesson refamiliarizes students with previous academic content to prepare them for the learning of new content.

4. Conceptual knowledge development (C_4)—the lesson teaches the concepts associated with the lesson.

5. Procedural knowledge development (C_5)—the lesson teaches students how to perform the independent work.

6. Guided practice (C_6)—the lesson gradually teaches students to perform the lesson assignment on their own.

7. Closure (C_7)—the lesson summarizes the lesson or attempts to determine whether students learned the lesson.

At this time, we can only guess at the values of the weighting coefficients; however, future research could establish the c_i values using observation data, student achievement results, and the statistical method of multiple regression. Considering the relative importance of developing conceptual and procedural knowledge, we propose that

$$Q_L = (0.1)C_1 + (0.1)C_2 + (0.1)C_3 + (0.3)C_4 + (0.2)C_5 + (0.1)C_6 + (0.1)C_7 \tag{5}$$

provides a reasonable starting point. (Note that a product of terms would not work. After all, failing to teach a lesson objective or forgetting to summarize a lesson does not equate to zero learning.)

The direct instruction model presented here only serves as an example. Although the terms in Equation (5) would likely change in description if applied to (say) a constructivist approach, the basic mathematical structure would remain.

7. INSTRUCTIONAL TECHNIQUE

For each instructional component, teachers can employ certain techniques to enhance its overall effectiveness. The following list does not exhaust all possible techniques.

1. Basic instruction (T_1)—the teacher initiates learning during the component through a teacher-centered [Hunter 1982; Hunter 2004] or student-centered [Bruner 1961] approach.

2. Questioning (T_2)—the teacher questions students to check for understanding, promote critical thinking, or enhance student engagement.

3. Cognitive strategies (T_3)—the teacher employs special techniques developed to help students regulate their own learning. Such techniques fall into five main types: rehearsal, elaboration, organization, comprehension monitoring, and affect [Nolet and McLaughlin, p. 44; Wolsey and Fisher 2009].

4. Vocabulary development (T_4)—the lesson delivery incorporates effective vocabulary development strategies.

5. Communication strategies (T_5)—the teacher employs special techniques designed to help those who struggle to understand the verbal and written communication taking place inside the classroom [Hansen-Thomas 2008]. Such strategies include Specially Designed Academic Instruction in English (SDAIE) [California Commission on Teacher Credentialing 1995, Sobul 1995] and Sheltered Instruction in Observation Protocol [SIOP 2008].

We then linearly expand each instructional component C_i in terms of the instructional techniques T_i used to deliver the component. Each instructional component expands in a basis of these five instructional techniques. For example, the development of conceptual knowledge in the Madeline Hunter method expands as

$$C_4 = t_1 T_1 + t_2 T_2 + t_3 T_3 + t_4 T_4 + t_5 T_5, \tag{6}$$

where each T_i equals 0 or 1 depending on whether we observe the respective technique. The rest of the instructional components expand in much the same fashion.

The value of each t_i depends on its importance in learning; for example, we could weight T_2 rather heavily by using a relatively large value for t_2 since questioning strategies so significantly impact learning [Bond 2008, Cecil 1995]. We can also adjust the expansion coefficients in

Equation (6) to suit special circumstances, such as setting t_5 to relatively large values when observing classes with large populations of English learners.

We should note that Q_L pertains strictly to the teachers' behaviors, irrespective of their students' behaviors. For example, if a teacher summarizes a lesson using cognitive strategies, we would credit him or her appropriately, even if the students pay no attention—the quantity of time factor Q_N described later already accounts for student disengagement. Again, we strive to maintain independence between all measurements.

To summarize the quality of instruction index mathematically,

$$Q_L = \sum_{i=1}^{N} c_i C_i = \sum_{i=1}^{N} c_i \left(\sum_{j=1}^{M} t_{ij} T_{ij} \right) = \sum_{i=1}^{N} \sum_{j=1}^{M} c_i t_{ij} T_{ij}, \qquad (7)$$

where researchers would establish values for the expansion coefficients c_i and t_{ij} prior to observations and would assign each T_{ij} the value 0 or 1 depending on what they see transpiring during the classroom session.

8. SAMPLE CALCULATION

To increase student achievement, teachers need to develop conceptual knowledge:

> Consider a teacher teaching a mathematics lesson on calculating the mean. After writing the definition on the board, she elaborates on the definition verbally and presents clarifying examples. The teacher writes the word *mean* on the board, and then points out that the definition used in math differs from the word used to describe a malicious person.

We can capture the impact of developing conceptual knowledge with the expansion in Equation (6). For the vignette above, a classroom observer would credit the teacher with the direct instruction of conceptual knowledge and vocabulary development of the word *mean*. Also, distinguishing a newly learned word from more common homographs constitutes a communication strategy. Therefore $T_1 = T_4 = T_5 = 1$. However, $T_2 = T_c = 0$ because the teacher employed no cognitive strategies to teach the concept and failed to question students about their knowledge of the concept. If we choose to place the majority of emphasis on basic instruction (say $t_1 = 0.5$) and questioning strategies (say $t_2 = 0.2$), then the expansion of conceptual knowledge development C_4 assumes the form

$$C_4 = (0.5)T_1 + (0.5)T_1 + (0.5)T_1 + (0.5)T_1 + (0.5)T_1$$

$$= (0.5)(1) + (0.2)(0) + (0.1)(0) + (0.1)(1) + (0.1)(1) = 0.7 \,. \tag{8}$$

The observer would therefore assign the teacher seven-tenths "credit" for conceptual knowledge development.

Note that we did not judge whether students responded positively to the teacher's techniques. Again, when measuring factors related to Q_L, we advise focusing on teacher behaviors, not student reactions to such behaviors. This criteria remains even when the teacher employs a student-centered instructional approach, such as discovery learning, where teachers receive credit for placing students in activities where they can learn concepts on their own (with sufficient guidance).

9. MEASURING QUANTITY OF TIME

In regards to the quantity of instruction, Q_N, recent literature thoroughly distinguishes between academic engagement time (the time in which students engage academic content) and the allocated instruction time (the time between classroom bells) [Berliner 1990, Gettinger 1985; Karweit et al. 1981]. Results indicate that academic engagement time serves as the better bellwether of academic achievement [Fisher et al. 1978, Gettinger 1985], which explains why we use it in this model.

Measuring Q_N requires direct classroom observation. A trained observer can judge instances when students appear academically engaged and time these portions of the observation session with a stopwatch [Walkup et al. 2011]. The Q_N factor simply measures the ratio of the total time students appear to mentally engage academic content to the total observation time. A value $Q_N = 0.70$ would indicate that "students appeared academically engaged 70% of the time between the official beginning of a classroom session and its end."[1]

Unlike Q_L, the Q_N index measures the behavior of the students, not the teacher. For example, academic engagement applies to students completing an exam, even while the teacher performs

---

[1]One factor that complicates such timings is partial engagement, in which a lesson only academically engages a portion of the class. We can incorporate partial engagement, however, by defining an effective engagement index that weights the proportion of students engaged at any time during the classroom session [Walkup et al. 2011]. Note that we can never know for certain whether a student is truly academically engaged; daydreaming and other off-task behaviors will always confound such measurements.

non-academic functions such as reading a newspaper. Although crediting a teacher for reading a newspaper during classroom hours may sound absurd, the Q_L index will adequately reflect the teacher's inactivity.

10. DISCUSSION

After choosing a random selection of public schools, researchers could perform in-class observations to establish the optimal values of the weighting coefficients j and k in Eq. (3). Once established, the classroom productivity index can then serve as a professional development utility at all levels of school governance.

Education researchers at some point will need to establish reasonable targets for both Q_L and Q_N. Based on their experience performing classroom timings, the authors believe schools can achieve the target $Q_N = 0.9$ with some concerted effort on part of the school staff; on the other hand, a suitable target for Q_L remains unclear.

Since T_{ij} in Equation (6) forms an $N \times M$ matrix, a mobile application could capture classroom data using touch-screen hot spots arranged in an $N \times M$ grid, therefore capturing the data necessary to measure the *quality* of instruction Q_L. An application that combines such functionality with a stopwatch for measuring the *quantity* of instruction Q_N could then capture in real-time all the data necessary to measure the classroom productivity index Q.

Used strategically, the classroom productivity index could shift the focus of school reform to input processes (the independent variables in the model), such as a teacher's questioning strategy, rather than output processes (the dependent variable Q), such as student achievement on assessments. Professional development can then focus on improving processes that raise the values of both Q_L and Q_N. In this paradigm, districts do not hire teachers to increase student achievement on state tests per se, but rather to employ methods designed to increase the values represented by the independent variables; high test scores manifest merely as an artifact of the process.

REFERENCES

BERLINER, D. 1990. What's all the fuss about instructional time? In *The Nature of Time in Schools,* M. BEN-PERETZ and R. BROMME, R., Eds. Teachers College Press, New York, 3-35.

BOND, N. 2008. Questioning strategies that minimize behavior problems. *Education Digest: Essential Readings for Quick Review 73*, 41-45.

BROOKS, J. G. and BROOKS, M. G. 1993. *In Search of Understanding: The Case for Constructivist Classrooms*. Association for Supervision and Curriculum Development, Alexandria, VA.

BRUNER, J. S. 1961. The act of discovery. *Harvard Educational Review 31*, 21–32.

CALIFORNIA COMMISSION ON TEACHER CREDENTIALING. 1995. *Guidelines for Staff Development Programs for Teachers of English Learners*, Guide. CCTC, Sacramento, CA.

CECIL, N. 1995. *The Art of Inquiry: Questioning Strategies for K–6 Classrooms*. Peguis Publishers, Winnipeg, Manitoba, CN.

COBB, C. W. and DOUGLAS, P. H. 1928. A theory of production. *American Economic Review* (supplement), 139–165.

FISHER, C., FILBY, N., MARLIAVE, R., CAHEN, L., DISHAW, M., MOORE, J. and BERLINER, D. 1978. Teaching and learning in the elementary school: A summary of the Beginning Teacher Evaluation Study. Technical Report. Far West Lab for Educational Research and Development, San Francisco, CA.

GETTINGER, M. 1985. Time allocated and time spent relative to time needed for learning as determinants of achievement. *Journal of Educational Psychology 77*, 3-11.

HANSEN-THOMAS, H. 2008. Sheltered instruction: Best practices for ELLs in the mainstream. *Kappa Delta Pi Record 44*, 165–169.

HUNTER, M. 1982. *Mastery Teaching*. TIP Publications, El Segundo, CA.

HUNTER, R. 2004. *Madeline Hunter's Mastery Teaching*. Corwin Press, Thousand Oaks, CA.

KARWEIT, N. and SLAVIN, R. 1981. Measurement and modeling choices in studies of time and learning. *American Educational Research Journal 18*, 157-71.

MID-CONTINENT RESEARCH FOR EDUCATION AND LEARNING (McREL). 2000. A new era of school reform: Going where the research takes us. McRel, Aurora, CO.

NOLET, V. and MCLAUGHLIN, M. J. 2005. *Accessing the General Curriculum*. Corwin Press, Thousand Oaks, CA.

SCHEERENS, J. and BOSKER, R. J. 1997. *The Foundations of Educational Effectiveness*. Elsevier, New York.

THE SIOP INSTITUTE. 2008. http://www.siopinstitute.net.

SOBUL, D. 1995. Specially Designed Academic Instruction in English. Document Reproduction Services No. ED391357.

WALBERG, H. J. 2004. Improving educational productivity: An assessment of extant research. Mid-Atlantic Regional Educational Laboratory, Washington, DC. In *The Scientific Basis of Educational Productivity*. R. F. SUBOTNIK and H. J. WALBERG, Eds., Philadelphia, PA.

WALBERG, H. J. 1980. A psychological theory of educational productivity. National Institute of Education, Washington, DC. In *Generalized Theory of Design and Analysis of Educational-Effectiveness Research*.

WALKUP, J. R., FARBMAN, D. and MCGAUGH, K. 2001. Bell to bell: Measuring time usage in the classroom. In preparation.

WOLSEY, D. and FISHER, D. 2009. *Learning to Predict and Predicting to Learn: Cognitive Strategies and Instructional Routines*. Allyn & Bacon, Columbus, OH.

CONTACT
John R. Walkup, Ph.D.
President
The Standards Company LLC
575 W. Alluvial Ave.
Clovis, CA 93611
(888)777-1417
 jwalkup@standardsco.com