Toward a Mathematical Model of Learning

Martin Buoncristiani
Thinking and Learning in Concert
mbuoncri@aol.com
http://ThinkingAndLearningInConcert.org

Abstract

Research points to the importance of a student's understanding the organization of his or her knowledge and for learners to control their own learning to insure that it is adaptable. We introduce a model describing the progress of learning to explore the acquisition of knowledge and its conceptual understanding. This model attempts describes the process of learning over a long period of time, typically the time taken for a course or a school term. It assesses the level of adaptive expertise achieved throughout this period. Teachers who intend to teach metacognitive skills need to be able to assess how their student's thinking skills are changing. This model provides a means for visualizing changes in student's learning.

Introduction

A summary of recent research in education by the U.S. National Research Council (NRC) emphasizes the importance of a student's understanding the organization of his or her knowledge (NRC, 2000, 2005):

> To develop competence in an area of inquiry, students must
> a. have a deep foundation of factual knowledge;
> b. understand facts and ideas in a context of a conceptual framework; and
> c. organize knowledge in ways that facilitate retrieval and application.

This NRC summary further stresses that skillful use of metacognition offers a real opportunity for learners to evaluate, manage and improve their own learning skills. The more learners know about their individual learning process, the more they are able to think about their own thinking, the better is their control of their thought processes.

Some students treat each learning experience as an isolated event, disconnected from their prior knowledge. They are unable to describe how they learn. They do not question the processes by which they learn. Their learning is episodic, rote memorization. They often acquire knowledge for a specific purpose such as a test, an essay, or a project and consequently, that knowledge is not easily transferred to different situations. Because they have minimal understanding of the process of learning, when confronted with situations that confuse them they have no resources to resolve the confusion.

In this paper we are taking steps towards developing a mathematical model of learning with understanding. We are not speaking here about simple learning as in the memorization of increasingly complex non-sense syllables in the seminal work of (Ebbinghaus, 1885), but rather a learning process that leads to conceptual understanding. We distinguish between learning of facts in isolation or episodic learning and learning with understanding (Costa & Kallick, 2008). This distinction has been made by others. Hatano distinguishes between two types of expertise used by problem solvers (Hatano & Inagaki, 1986). *Routine expertise* is characterized by a high degree of procedural efficiency in a specialty area. *Adaptive expertise* exhibits both the core efficiencies of routine experts and the willingness and ability to modify skills to fit new contexts and to articulate the concepts and principles underlying these skills.

Edward de Bono makes a similar distinction between two types of thinking (deBono, 1994). *Operancy thinking* involves the skills needed for doing things and is similar to routine expertise, while *design thinking*, involving the skills needed to think creatively and innovatively, is similar to adaptive expertise. This distinction, whether expressed as routine versus adaptive expertise or operancy versus design thinking, is important because it marks the distinction between solving problems using usual or familiar methods and attacking them with nonstandard, creative, innovative techniques.

Teachers who can impart an understanding of the learning process to their students will enhance those students' ability to develop adaptive expertise. A teacher's own expertise involves his or her personal framework of knowledge. The teacher can transfer the facts that make up that knowledge framework but not the organization of it. Each student must construct that organization for themselves. The process of study is mainly concerned with the student's individual efforts to organize factual knowledge so that it can be retrieved and used easily and in differing situations.

A Metaphor for Learning with Understanding

We begin by introducing a simple but useful metaphor for learning with understanding. Without specification of the subject matter, we explore how the state of a learners mind might change with conceptual understanding of knowledge acquired. This provides a measure of the level of understanding beyond the simple knowledge of a set of facts.

Imagine that each bit of knowledge a student learns is a piece of a jigsaw puzzle. The puzzle may be intact and complete in the teacher's mind but it is delivered to each student one piece at a time. Some students collect the pieces as they arrive and keep them in that order. Other student's realize how some of the pieces go together and begin to reconstruct the puzzle. They may not get all of the puzzle pieces in their proper places immediately but by attempting to organize them they get an improved picture as their learning continues.

So we consider each piece of the puzzle as an isolated fact. By thinking about each fact and how it relates to the other facts a student can link them together to find a deeper order among the collection of facts than they possess in isolation. Learning with understanding is similar to finding that deeper order in puzzle pieces as they come together. It pays attention to the organization and context of knowledge.

We can push this metaphor a little bit further and indeed quantify it. If we suppose that the puzzle has N pieces, we can think about the many ways the puzzle could be partially solved; that is the many possible intermediate states of the puzzle between the pieces being totally separate and they all being connected in one picture. This is equivalent to a famous mathematical problem of finding the partition of an integer, that is, finding all of the ways the number N can be expressed as a sum of whole numbers (where the order of numbers in the sum is not important). For example, the number seven can be expressed as 15 different sums of positive integers. Each of these sums is equivalent to a different possible state of solving the puzzle; thus the sum $7 = 3+2+1+1$ corresponds to one combination of three pieces fitted together, and another with two pieces together along with two separate pieces. Each of these states represents a different degree of order.

This situation is represented in Figure 1. In that graph the point at the origin represents the sum 7 = 1+1+1+1+1+1+1, which in turn is equivalent to the puzzle with separate pieces. The next point represents 7 = 2+1+1+1+1+1, equivalent to only two pieces being joined. As the puzzle is solved the order is increasing until the last point with the sum 7=7 representing the completed puzzle. Returning to the metaphor, the degree of order increasing as the puzzle is solved represents the increasing understanding achieved as the individual facts become organized. The idea of this example arises from the notion of entropy as a measure of the disorder of a physical system.

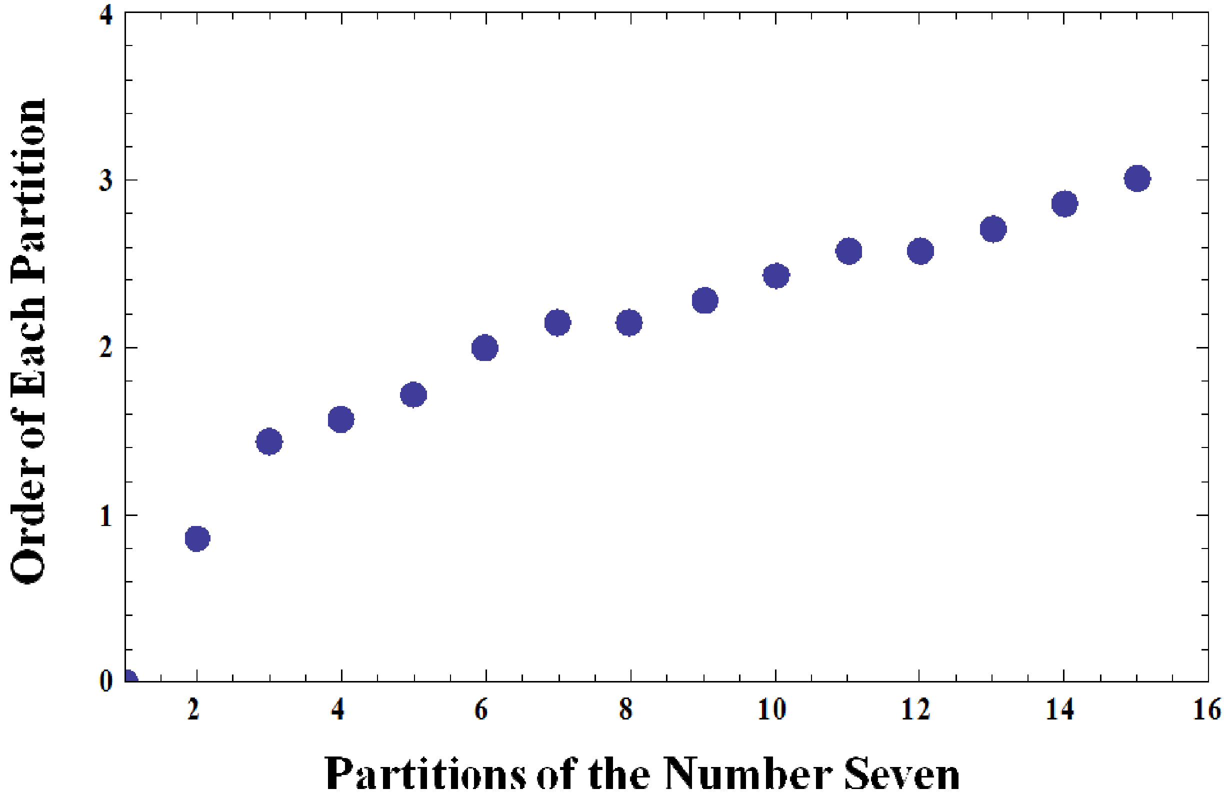

Figure 1. This graph represents the increasing order in sequence of partitions of the number 7. This mathematical problem of long standing is used as a metaphor for learning with understanding.

Thus, metaphorically, understanding of new knowledge comes from piecing ideas together, increasing their organization and order. Like all metaphors this one is suggestive but imperfect. Bits of information do not fit together tightly like jig saw puzzle pieces but rather more like a network. That is, a collection of nodes representing isolated facts connected in a particular way by links. This results in a pattern of nodes and links forming a personal mental construct. The manner in which knowledge is organized or linked becomes a personal information network and determines the ease of accessing that information (Buoncristiani & Buoncristiani, 2007).

A Model for Learning with Understanding

We introduce a model describing the progress of learning specifically to explore the acquisition of knowledge and its conceptual understanding. This model attempts to describes the process of learning over a long period of time, typically the time taken for a course or a school term. It is intended to assess the level of adaptive expertise achieved at the end of this period.

It is based on a constructivist philosophy of learning and so deals with how a student enhances learning through his or her own efforts to construct an understanding of the subject area. The learning processes involve the integration of previously learned knowledge with new knowledge. The thoroughness of understanding of previously learned knowledge influences the efficiency with which new knowledge is assimilated.

No particular teaching or learning strategy is assumed, rather it is anticipated that over the period covered many different teaching activities are employed and the learner employs many strategies for understanding. The main purpose of the model is assess the how an individual learner's skill at learning affects the overall efficiency of learning. Advocates of metacognition assert that teaching students efficient learning techniques will improve their performance. This is an attempt to quantify that assertion so that we may begin to explore ways of assessing it in the classroom.

Over the long period of learning the learner may employ many different strategies for learning (note taking, note revision, reading, outlining, rote memorization and so forth). During that time many external factors may also affect the efficacy of actions taken by the learner (tiredness, current mood, study environment and so forth). To account for all of these variables we use a statistical model. Learning will occur incrementally with increments determined by probability.

The first stage of the model describes the process of learning. We assume that learning takes place through a sequence of states designated by s_k where produces k=1, 2, 3, ... These states are analogous to the different states attained in completing a jig saw puzzle. Action taken by the learner may cause a change in the state or possibly no change. We assume that a learning action beginning with the learner at a state s_k will lead to a new state which is one of the five adjacent states (s_{k-2}, s_{k-1}, s_k, s_{k+1}, s_{k+2}). Which of the five outcomes of the learning action actually occurs is given by a probability distribution (P_{k-2}, P_{k-1}, P_k, P_{k+1}, P_{k+2}) where the sum of the probabilities is one

$$\sum_{k=-2}^{2} P_k = 1 .$$

The distribution of probabilities determines the rate at which the learning occurs.

To see how this works we initially choose the probability distribution to be linear, that is

$$P_i(x) = x\frac{i}{10} + \frac{1}{5}, \text{for } i = -2, -1, 0, 1, 2 \text{ and } 0 \le x \le 1.$$

For x = 0 each individual probabilities are equal, implying that it is equally likely that the learner moves up or down on the learning curve, while for x = 1 the probabilities are skewed toward advancing along the learning curve. So, in this case, x is a parameter indicating learning skill. Figure 2 compares the learning achievement for three values of x: (0.7, 0.5, 0.3) and for 500 learning increments.

While the linear probability distribution is a bit naïve it does illustrates the halting and incremental nature of learning. In what follows we use a more realistic distribution (0.0, 0.1, 0.3, 0.4, 0.2). This distribution favors learning progress. There is a 10% chance of going backward and a 60% chance of moving forward. It was arrived at after some experimentation with different distributions. It represents the progress of a good student generally intent on learning.

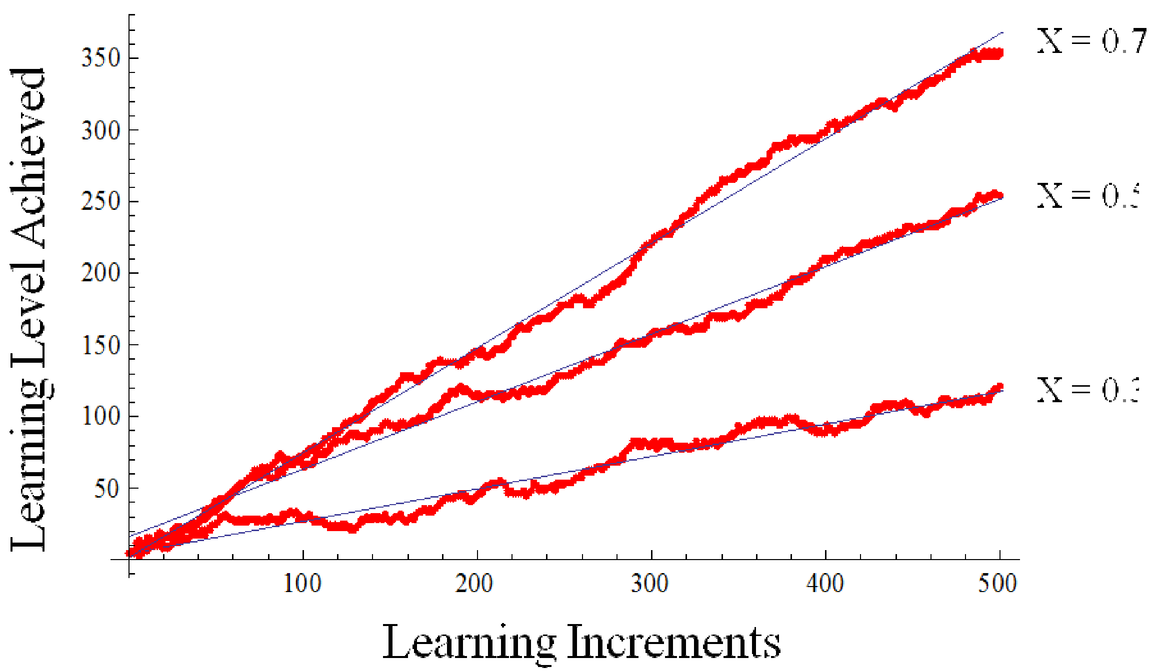

Figure 2. This graph compares of the learning level achieved with three different linear probability distributions, $P_i(x) = x\frac{i}{10} + \frac{1}{5}$ for x: (0.7, 0.5, 0.3).

With this simple statistical model of individual learning one can visualize knowledge acquisition. The next stage is to introduce the effect of a conscious effort by the learner to organize learning so that it can be used more flexibly. Howard Gardner has defined understanding as "... the ability to apply facts, concepts and skills in new situations, where they are appropriate." Understanding comes from work by the learner to contextualize new knowledge. Again, this part of learning can be done in many ways and in many circumstances so we will characterize this progress statistically.

We describe level of acquired understanding as a sequence of discrete states. Not every attempt to understand achieves an incremental change in understanding so we will use a probability function UP(x, x_0) equal to one if the randomly chosen number $0 \le x \le 1$ is smaller than x/x_0 and zero if it is not. Then, as learning proceeds as described with the model above, each step will be accompanied by a attempt to understand which will be successful or not as determined by UP(x,x_0). Figure 3 shows how the level of understanding increases as knowledge is acquired with a value of x_0 equal to eight.

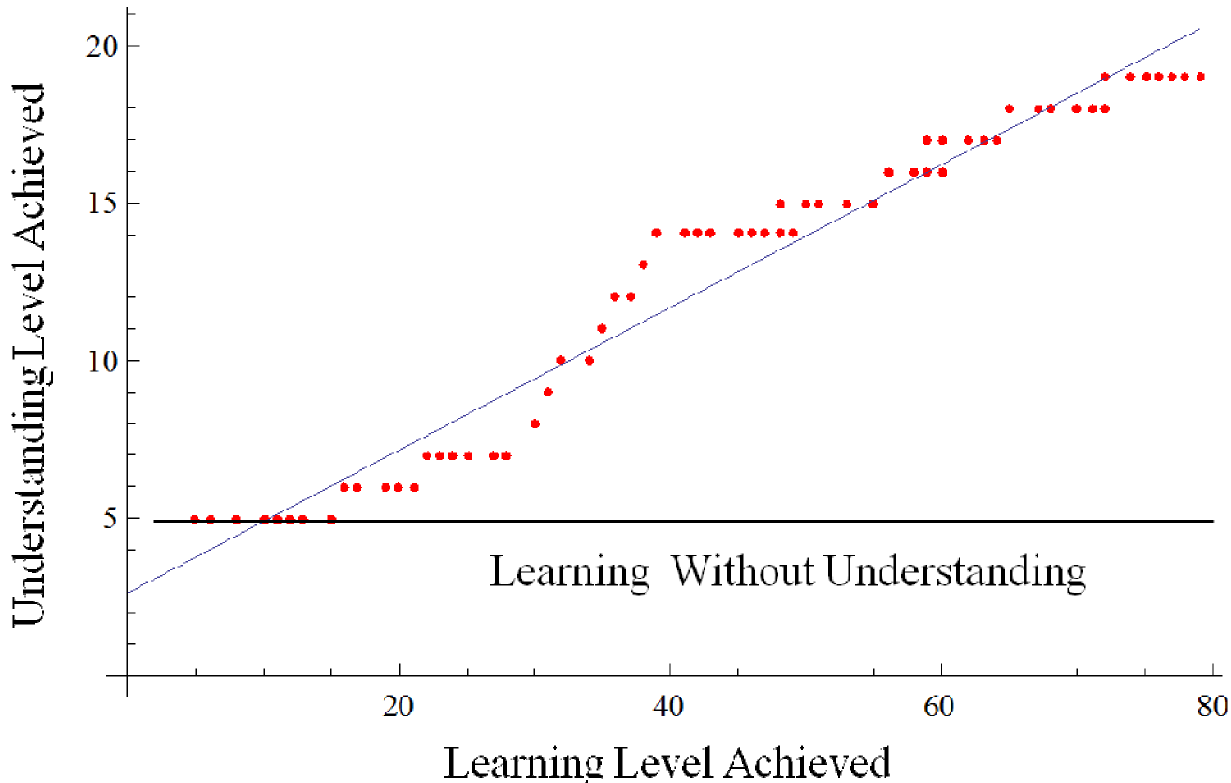

Figure 3. This figure shows how the level of understanding increases as knowledge is acquired. Understanding increases incrementally above the horizontal line representing learning without understanding.

The model was constructed to illustrate how learning with understanding might proceed over an extended period; to show how the level of expertise acquired over a course or term might evolve? An important attribute of this model is its inclusion of a means to describe a student's learning skill. This is done through the probability distribution P_k. Students learn differently. One factor in this difference is an individual student's metacognitive ability. Students who have gained control of their learning and are aware of the need to organize knowledge learn more efficaciously (Buoncristiani & Buoncristiani, 2012). This effect compounds itself throughout a course.

A Hidden Bimodal Distribution

The difficulty in assessing student's learning skill can be seen in a simulation of the grade distribution in a class with a conjectured bimodal distribution. Assume that one particular class is composed of two groups of students each with a different average learning ability. Just because a student is a skillful learner does not mean that he or she will apply that skill to the fullest. When an assessment of the class is made the grades vary with some measured parameter scaled for convenience to vary between 0 and 1. This parameter is introduced to allow some variability in the effectiveness with which a student's learning skill is applied in this assessment; it could be, for example, the time taken to study for the assessment, time taken to perform the assessment or weekly study time. The grades for each group are assumed to fall along a linear regression line.

Groups are labeled 1 and 2 and each group is assumed to have the same number of students, 20 in this example. The linear regression parameters are slope and y-intercept shown as the straight lines in Figure 1 (Grades 0 – 100 on the vertical axis and the nominal parameter value 0 – 1 on the horizontal axis). In this simulation, values of the parameter for each student are distributed normally about a median parameter value MuPar = 0.7, with a variance SigPar = 0.1. Grades for each group are distributed normally around the group's regression line with median value MuGrade = 0 and standard deviation SigGrade = 7.

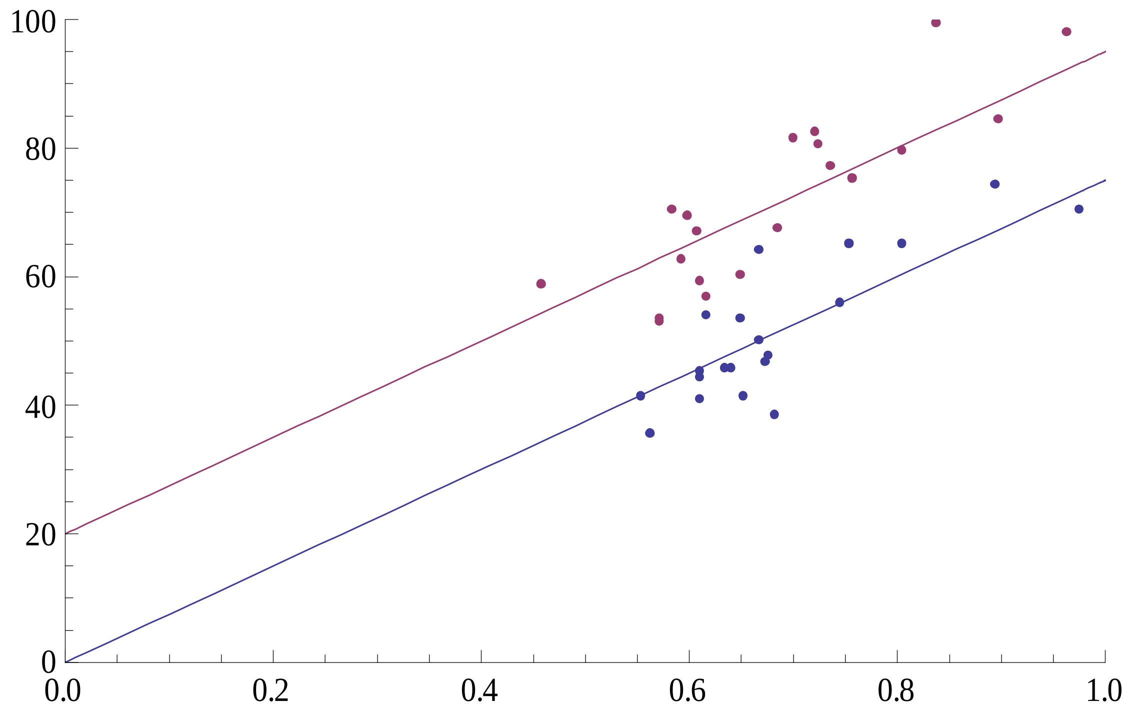

Figure 4. A simulated grade distribution for performance (0 to 100) by each of two groups of students with different learning skills versus a parameter (0 to 1) measuring the effectiveness of student's use of their learning skill.

A histogram of the grade distribution (Figure 5) does not show the bi-modal character from which it was constructed, but asking an additional question related to learning ability might just tease this information out.

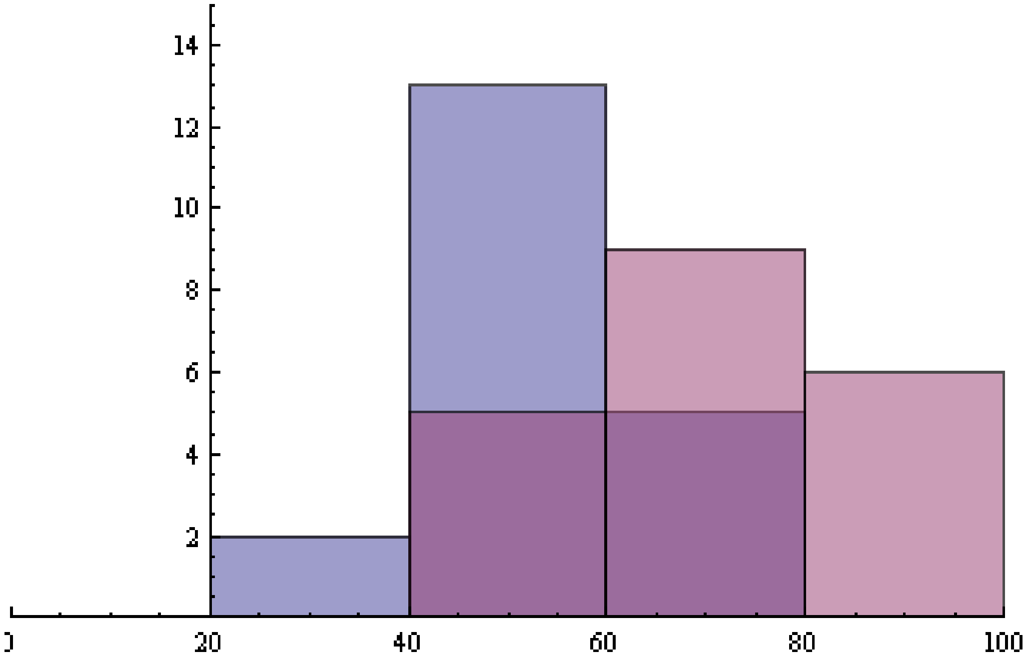

Figure 5. A histogram of the simulated grade distribution does not show the bimodal character of the student's learning skill assumed when the grade distribution was constructed.

As an example of how student learning skill can be quantified in the classroom we offer the results of a single non-discipline question in a first year university class (Buoncristiani & Buoncristiani, 2012). The following question on placed on each exam:

During this middle part of the course I have been spending, on average,

a) one hour
b) two hours
c) three hours
d) more than three hours each week

studying for this course.

This question was graded as correct for all students who answered it assuming that students were generally truthful. The results of this question are shown in Figure 6. The average study time was 2.3 hours per week. It is clear from the graph that generally grades increase with study time. Basically spending two hours per week in study instead of one hour will increase your grade by about 6 points.

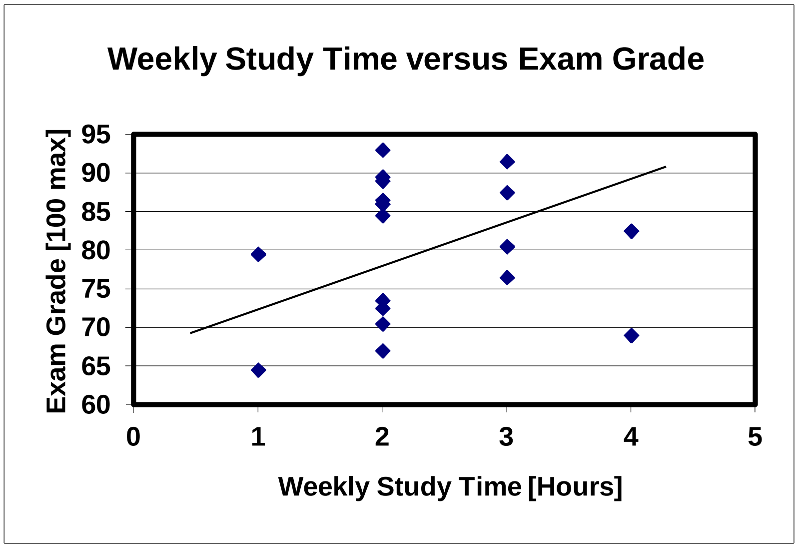

Figure 6. This graph shows the exam grade for a test on material covered in a five week period vs. the average number of hours spent in study over that period. The line is drawn to emphasize the two groups of data.

The conclusion from this graph is not surprising. We all know that if you want to increase your grade, you work harder. What they give us is the ability to make quantitative statements about the improvement that a specific increase in effort will produce.

Another interesting conclusion emerged from this data. The graph shows the students separated into two groups: the students in the group above the lines were more efficient learners than those below. For example, referring to Figure 6 where the distinction between the groups is clearer, the top group scores about 15 points higher than the lower group. The group of better students seems to know how much time they needed to devote to study and test taking. We suggest that the difference between the two groups lies in their ability to think productively about how they learn, a difference we may call the "metacognition gap." Students in the lower group are episodic learners, in the terminology of (Costa & Kallick,2008), and they carry this lack of metacognition to their thinking about how they study.

We are suggesting that these preliminary results indicate that using non-discipline questions in testing can assess learning skill. For example, when students are asked, in various forms, how much time they spend in study the results correlate with their grade as is illustrated in Figure 6.

A good area to apply this assessment of individual student's learning skill is in Computer Science, a field where student must first learn a new vocabulary and syntax of a programming language and then apply that learned skill to solve problems creatively (Chatzopoulou, 2009). Often one finds that the distribution of final grades in beginning Computer Science courses is bimodal and this, in turn, is often explained with the hypothesis that some students are good programmers and others are not. Evidence does not seem to support this (Robins, 2010). A more reasonable hypothesis is that the difference is performance is due to student's ability to learn and this problem if identified can be addressed directly. A similar bimodality of grade distributions also appears in introductory mathematics and science courses.

Conclusions

Accepting the importance of teaching students how to think one has then to face the issue of assessing how well students think. Teachers aiming at teaching metacognitive skills needs to answer three questions:

What does this student know?
How does this student think?
How do these change over time?

The first of these questions is answered by the usual assessment instruments that deal with measuring the factual knowledge and procedural skill the student has acquired. The second is very difficult to assess because it depends on the student's behavioral characteristics and mental capacity. Thinking skills may be assessed indirectly by asking students to apply what they have learned in new and different contexts. This is usually done at the end of the course in a summative assessment. A real problem faced by teachers trying to teach metacognitive skills is to assess student's success along the way. The model we have created is intended to simulate how learning with understanding might proceed over the time it takes to complete a course. It puts forward a reasonable expectation of how a student with a fixed level of learning skill might progress.

The information that is entered into the model is the probability distribution for learning, P_k and the probable rate of growth of understanding, UP. One possible use of the model is to identify a few students of differing learning skills and assign input values for each student and see how the model tracks their actual progress.

We end with this observation – courses, at every level, are usually organized with a specific pace. The syllabus or subject pacing guides dictate what is learned and when it is learned. However, students learn at different rates and consequently at any time during a course there will be some students who are apace with the syllabus while others are falling behind. We hope that understanding how an individual student's learning skill affects the efficiency and thoroughness of their learning will aid in the reformation of the curriculum to accommodate for this.

References

Buoncristiani, A. M., & Buoncristiani, P. (2007). A network model of knowledge acquisition. Proceedings of the 13th International Conference on Thinking. Norrköping, Sweden. Retrieved from http://devisa-hb.se/thinkingconference/index13.html.

Buoncristiani, M., & Buoncristiani, P. (2012). *Developing mindful students, skillful thinkers,thoughtful schools*. Thousand Oaks, CA: Corwin Press.

Chatzopoulou, D. I. (2009). Adaptive assessment in the class of programming. Retrieved from http://conta.uom.gr/conta/publications

Costa, A. L., & Kallick, B. (2008). *Learning and leading with habits of mind.* Alexandria, VA: Association for Supervision and Curriculum Development.

De Bono, E. (1994). *Thinking course.* New York: Barnes & Noble.

Ebbinghaus, H. (1885). *Über das Gedchtnis. Untersuchungen zur experimentellen Psychologie*. Leipzig: Duncker & Humblot; the English edition is Ebbinghaus, H. (1913). *Memory. A Contribution to Experimental Psychology*. New York: Teachers College, Columbia University (Reprinted Bristol: Thoemmes Press, 1999).

Hatano, G., & Inagaki, K. (1986). Two courses of expertise. In H. Stevenson, H. Azuma, & K. Hakuta (Eds.), *Child development and education in Japan* (pp. 262–272). New York: Freeman.

National Research Council. (2000). *How people learn: Brain, mind, experience, and school.* Committee on Developments in the Science of Learning. J. D. Bransford, A. L. Brown, & R. R. Cocking (Eds.). Commission on Behavioral and Social Sciences and Education. Washington DC: National Academy Press. Retrieved from http://www.nap.edu/openbook.php?isbn=0309070368.

National Research Council. (2005). *How students learn: History, mathematics, and science in the classroom.* Committee on *How People Learn:* A Targeted Report for Teachers. M. S. Donovan & J. D. Bransford (Eds.). Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Press. Retrieved from http://www.nap.edu/catalog.php?record_id=10126.

Robins, A. (2010). Learning edge momentum: A new account of outcomes in CS1. *Computer Science Education*, *20*(1), 37-71.