International Electronic Journal of

Mathematics Education

Volume 3, Number 2, July 2008 www.iejme.com

MATHEMATICS TEACHERS' INTERPRETATION OF HIGHER-ORDER THINKING IN BLOOM'S TAXONOMY

Tony Thompson

ABSTRACT. This study investigated mathematics teachers' interpretation of higher-order thinking in Bloom's Taxonomy. Thirty-two high school mathematics teachers from the southeast U.S. were asked to (a) define lower- and higher-order thinking, (b) identify which thinking skills in Bloom's Taxonomy represented lower- and higher-order thinking, and (c) create an Algebra I final exam item representative of each thinking skill. Results indicate that mathematics teachers have difficulty interpreting the thinking skills in Bloom's Taxonomy and creating test items for higher-order thinking. Alternatives to using Bloom's Taxonomy to help mathematics teachers assess for higher-order thinking are discussed.

KEYWORDS. Higher-order Thinking, Bloom's Taxonomy, U.S. High School Mathematics Teachers, Algebra, Assessment.

INTRODUCTION

Mathematics teaching in the U.S. has traditionally relied on factual recall and a focus on the use of standardized algorithms with little effort to teach or assess for higher-order thinking (Kulm, 1990; Battista, 1994). As a result, students generally learn mathematics without being able to use their knowledge to solve problems in diverse or non-familiar situations (de Lange, 1987; Schoenfeld, 1988). Although there has been an effort to reform mathematics education in the U.S. over the past two decades (NCTM, 1989; 2000), teaching mathematics has changed little since Schoenfeld (1988) characterized the typical mathematics classroom:

> All too often we focus on a narrow collection of well-defined tasks and train students to execute those tasks in a routine, if not algorithmic fashion. Then we test the students on tasks that are very close to the ones they have been taught. If they succeed on those problems, we and they congratulate each other on the fact that they have learned some powerful mathematical techniques. In fact, they may be able to use such techniques mechanically while lacking some rudimentary thinking skills. To allow them, and ourselves, to believe that they "understand" the mathematics is deceptive and fraudulent. (p. 30)

ISSN: 1306-3030

This notion of teaching mathematics stands in contrast to teaching for higher-order thinking skills where students are able to meaningfully apply methods and concepts to situations previously unfamiliar to them (Donovan & Bransford, 2005; Hiebert, et al, 1997). However, research indicates that many teachers have a weak conception of higher-order thinking (Harpster, 1999) and that teaching for higher-order thinking is difficult for teachers to sustain as an integral part of classroom instruction and assessment (Henningsen & Stein, 1997).

Characterizing Lower- and Higher-Order Thinking

Resnick (1987) noted that thinking skills resist precise forms of definition, but lower- and higher-order thinking can be recognized when each occurs. Lower-order thinking (LOT) is often characterized by the recall of information or the application of concepts or knowledge to familiar situations and contexts. Schmalz (1973) noted that LOT tasks requires a student "… to recall a fact, perform a simple operation, or solve a familiar type of problem. It does not require the student to work outside the familiar" (p. 619). Senk, Beckman, & Thompson (1997) characterized LOT as solving tasks where the solution requires applying a well-known algorithm, often with no justification, explanation, or proof required, and where only a single correct answer is possible. In general, LOT is generally characterized as solving tasks while working in familiar situations and contexts; or, applying algorithms already familiar to the student.

In contrast, Resnick (1987) characterized higher-order thinking (HOT) as "non-algorithmic." Similarly, Stein and Lane (1996) describe HOT as "the use of complex, non-algorithmic thinking to solve a task in which there is not a predictable, well-rehearsed approach or pathway explicitly suggested by the task, task instruction, or a worked out example." (p. 58) Senk, et al (1997) characterized HOT as solving tasks where no algorithm has been taught, where justification or explanation are required, and where more than one solution may be possible. In general, HOT involves solving tasks where an algorithm has not been taught or using known algorithms while working in unfamiliar contexts or situations.

Teaching for HOT

According to the National Center for Education Statistics (1996), teaching for HOT along with professional development in HOT were found to be two of the top five variables positively associated with improved student achievement. Students of teachers who teach for both LOT and HOT outperform students whose teachers only teach for LOT (Wenglinsky, 2002). However, national and international assessments in mathematics indicate that U.S. students are not doing well on items requiring the use of HOT (Mullis, et al, 2004). U.S. students performing poorly on national and international exams is reflected in the research which indicates that most U.S. teachers do not teach and assess for HOT (Kulm, 1990; Senk, et al, 1997).

U.S. teachers not assessing for HOT is well-documented and is not limited to mathematics. Reynolds and Menard (1980) found that teachers' classroom tests were heavily reliant on LOT tasks (interpreted as knowledge, comprehension and application in *BT*). Similarly, Fleming and Chambers (1983) used *BT* to analyze over 8000 test items from K-12 teacher-developed tests and found that consistently across all grade levels over 90% of test items measured LOT. In a study of high school grading practices in mathematics, Senk, et al (1997) found that the percentage of LOT items on mathematics teachers' tests ranged from 53% to 90% with a mean of 68%. These studies analyzed existing tests created by teachers; in contrast, Harpster (1999) found that when mathematics teachers were specifically asked to write a single task representative of HOT, sixty-percent of teachers created a task that assessed LOT.

In general, teacher assessments tend to focus on LOT even when teachers say they want to teach and assess for HOT; however, teachers are often unaware of this inconsistency (Black & Wiliam, 1998). Any endeavor to improve assessing for HOT depends on whether or not teachers can identify and create items that assess for HOT (Costa, 2001; Kulm, 1990; Resnick & Resnick, 1992). Numerous efforts to improve assessing for HOT include *Bloom's Taxonomy* which is often used to evaluate the level thinking required on tasks. How is HOT characterized in *Bloom's Taxonomy?*

Bloom's Taxonomy and higher-order thinking

For over 50 years, *Bloom's Taxonomy (BT)* (Bloom, 1956) has heavily influenced teaching and assessment throughout the world (Anderson & Sosniak, 1994) and is still commonly used in mathematics education. For example, Kastberg (2003) and Vidakovic, Bevis, & Alexander (2003) provide examples of how high school and collegiate mathematics teachers can use *BT* to develop test items. Numerous studies have used *BT* as the standard for judging whether test items are LOT or HOT. The thinking skills in *BT* considered LOT include knowledge and comprehension, while the thinking skills of analysis, synthesis and evaluation are considered HOT. Application often falls into both categories.

In *BT*, for a test item to be at the level of application or higher, a "new situation" (to the student) is required. Bloom emphasized in his original work in 1956 and subsequent discussions on this issue (Bloom et al, 1971, 1981) that application and higher levels in the taxonomy do not refer to test items where only minor changes are made, but otherwise, the procedure was the same to that practiced in class. As Bloom, et al (1981) stated:

> By 'new problems and situations' we mean problems and situations which are *likely* to be new to the student. These are similar to those which were included in the instruction but have some element of newness or unfamiliarity for the student. Students should not be able to solve the new problems and situations merely by remembering the solution to or the precise method of solving a similar problem in class. It is not a new problem or situation if it is exactly like others solved in class except that new quantities or symbols are used (as in mathematics or physics). (p. 233)

Despite the widespread use of *BT*, it is not well understood how mathematics teachers' interpret *BT* or whether *BT* facilitates the development of HOT test items in mathematics.

RESEARCH QUESTIONS

As discussed above, teachers' test items have been investigated regarding whether they assess for LOT or HOT; however, researchers used *BT* or variations of *BT* (Cooney, 1992; Wilson, 1971), or other frameworks (e.g., Quellmalz, 1987) to determine whether test items were HOT or LOT. In these studies, teachers were not using *BT* as a guide in the development of their own HOT test items (e.g., Stiggins, Griswold, & Wikelund, 1989; Fleming & Chambers, 1983; Harpster, 1999). Therefore, this study was designed to investigate the following questions:

1. How do high school mathematics teachers define lower- and higher-order thinking?
2. Which thinking skills in *Bloom's Taxonomy* do high school mathematics teachers associate with higher-order thinking?
3. What type of Algebra I final exam items do high school mathematics teachers perceive are representative of thinking skills in Bloom's Taxonomy?

According to Bloom (1956), *BT* was designed to classify types of thinking likely required on a test item *after* students have been taught the specific objectives being tested. Therefore, for research question 3, a final exam context was chosen to ensure teachers developed test items based on the type of thinking likely used by a student *after* s/he had been taught the material. An Algebra I context was chosen to provide uniformity in the types of test items being analyzed.

METHOD

Participants

Thirty-two high school mathematics teachers from four school districts in a state in the southeast U.S. participated in this study. Teachers were asked to volunteer for this study while attending workshops at their school during professional development days. Participants were from rural and semi-urban communities, predominately female and Caucasian with ages ranging from early twenties to late fifties. No additional demographic information was collected (e.g., gender, race, years teaching, highest degree earned, etc.).

Data Collection

The high school mathematics teachers were first asked to write a definition of lower- and higher-order thinking. Second, after writing their definitions, teachers were given a handout summarizing the thinking skills in *BT* (see Appendix). The thinking skills were listed in alphabetical order to avoid biasing teachers' perceptions of whether a thinking skill might be LOT or HOT depending on its order in the handout. Third, teachers were asked to identify whether they were "very familiar," "somewhat familiar," or "not familiar" with *BT*, and to briefly describe where, if at all, they had learned about *BT*. Fourth, teachers were asked to classify each thinking skill in *BT* as either LOT or HOT. And lastly, teachers were asked to write an Algebra I final exam item for each thinking skill.

Using the characterizations from the literature described earlier, test items in this study were classified as LOT and HOT as follows:

a) LOT – "*Algorithmic Thinking*": Solving tasks that require the recall of information or the application of well-known algorithms in situations and contexts likely familiar to the student.

b) HOT – "*Non-algorithmic Thinking*": Solving tasks where no specific algorithm has been taught to the student; or, using known algorithms in contexts or situations likely unfamiliar to the student.

As noted earlier, Bloom (1956) stated that when testing for higher levels of thinking, "…the problem situation must be new, unfamiliar, or in some way different from those used in the instruction" (p. 238). In this study, the differences between LOT and HOT are consistent with the differences in *BT* between comprehension and application. Comprehension is characterized in *BT* as using *known algorithms*. In contrast, application involves using knowledge in new situations and solving problems without being told which algorithm or method to use. For

Bloom, “the emphasis in writing application items is on getting situations new to the student” (p. 130) and if a “problem or one just like it had been solved in classwork, [it] would be a comprehension rather than an application item” (p. 133).

Data Analysis

Teachers’ categorization of thinking skills as LOT or HOT were analyzed using descriptive statistics (frequencies and percentages). The test items created by teachers were classified as HOT or LOT using the definitions above. A graduate research assistant also classified each test item as LOT or HOT independent of the researcher; the agreement in classifying items was 92%. Disagreements over a test item being LOT or HOT were resolved in favor of the teacher’s classification. The research focused on teachers’ interpretation of HOT while using *BT*; test items were not independently analyzed as to whether they “fit” with a thinking skill in *BT* – only whether an item matched the LOT or HOT classification of the teacher.

RESULTS

Seventy-five percent of teachers indicated that they were either very familiar or somewhat familiar with *BT*. Teachers familiar with *BT* indicated they used the taxonomy while in their undergraduate or graduate teacher education programs to develop test items as part of an educational psychology, methods, or evaluation course.

Research Question 1: How do high school mathematics teachers define lower- and higher-order thinking?

Many teachers’ definitions of LOT and HOT included characteristics of LOT and HOT often discussed in the literature. A summary of teacher descriptions are included in Table 1.

Table 1: Teachers’ descriptions of LOT and HOT

LOT	HOT
Following rules	Discovering patterns
Performing computations	Solving word problems
Definitions / vocabulary	Interpreting information
Simple applications	Complex applications
Procedural knowledge	Conceptual understanding
“Copies” teacher / rote learning	Critical thinking / analyzing

However, not all teachers viewed HOT in these terms; approximately one – third of teachers included as part of their definitions of LOT or HOT (a) level of difficulty, (b) number of steps “required” to solve a task, or (c) solving tasks involving “higher math”. One teacher wrote, "A math problem is higher order thinking if it takes 3 or more steps to solve." She later provided as an example of HOT, "Simplify (3x3 – 6x + 9x2) ÷ 3x" while writing in the margin of her paper "three steps." Several teachers wrote that LOT problems are easier than HOT problems (e.g., "Higher order thinking involves solving difficult or challenging math problems.") However, there are many mathematics tasks that are computational / algorithmic in nature that are quite difficult or challenging; therefore, although HOT items tend to be more difficult, level of difficulty is not a characterization in the literature on HOT and LOT (de Lange, 1987).

Regarding using "basic" versus "advanced" mathematics as a key characteristic of LOT and HOT, one teacher wrote that "Lower order thinking involves solving basic math problems e.g., adding, multiplying integers / fractions, etc., while higher order problems are beyond basic math; e.g., algebra and calculus." When writing their definitions of HOT, none of the teachers included the concept of *familiarity* with an algorithm or a problem situation. However, this is a fundamental concept in distinguishing LOT and HOT in the literature and in *BT*.

Research Question 2: Which thinking skills in *Bloom's Taxonomy* do high school mathematics teachers associate with higher-order thinking?

Using the descriptions in the handout or their prior experiences with *BT*, teachers were asked to identify each thinking skill in *BT* as either LOT or HOT. The results are presented in Table 2.

Table 2: Teachers classification of thinking skills

Bloom's Taxonomy	LOT	HOT
LOT	**# (%)**	**# (%)**
Knowledge	29 (91%)	3 (9%)
Comprehension	13 (40%)	19 (60%)
LOT or HOT		
Application	9 (28%)	23 (72%)
HOT		
Analysis	3 (9%)	29 (91%)
Synthesis	6 (19%)	26 (81%)
Evaluation	18 (57%)	14 (43%)

Teachers' classification of thinking skills in *BT* indicates that a description of a thinking skill by itself (or least the summary provided for teachers in this study) is not sufficient to help teachers differentiate LOT and HOT. Over 90% of teachers correctly identified knowledge as LOT and analysis as HOT. However, comprehension (LOT) was interpreted as HOT by 60% of mathematics teachers. Evaluation and synthesis are considered HOT in *BT*; however, over half of the mathematics teachers interpreted evaluation as LOT and approximately one-fifth of teachers indicated synthesis was LOT.

Many teachers correctly identified the levels of thinking in *BT* as either HOT or LOT; however, teachers who defined HOT by (a) number of steps, (b) level of difficulty, or (c) algebra as a "higher-order subject" tended to list all thinking skills (except occasionally knowledge) as HOT. Given teachers' interpretation of HOT in *BT*, do mathematics teachers write HOT test items for thinking skills in *BT* they classify as HOT?

Research Question 3: What type of Algebra I final exam items do high school mathematics teachers perceive are representative of the thinking skills in *Bloom's Taxonomy*?

Teachers were asked to create Algebra I final exam items for each thinking skill in *BT*. Test items fell into one of three categories: Items classified as LOT by both the teacher and researcher (LOT – LOT) [Table 3]; items classified as HOT by both the teacher and researcher (HOT – HOT) [Table 4]; and items classified as HOT by the teacher, but LOT by the researcher (HOT – LOT) [Table 5]. No items classified as LOT by teachers were classified as HOT in this study. For the LOT items in table 3, it is likely that students have been taught an algorithm or general procedure to answer each question. Items in Table 4 are HOT under the assumption that students were previously not taught procedures or algorithms to solve these or similar problems.

Table 3: Sample LOT-LOT items

State the quadratic formula (K)
Solve x + 2 = 6 (K)
Simplify: 3x – 7y + 5 – x + 8y (C)
Change 3x + 4y = 12 to the form y = mx + b (C)
Simplify: (3*b*2*c*)(8*b*3*c*6) (Ap)
Multiply: (2x – 5)(x + 8) (Ap)
What is the greatest common factor of 3*x*2 - 9*x*4 + 6*x*3? (An)
Simplify 3x + 7xy – 2x + 3(x – y) – xy (S)
Solve 4(x – 7) + 5 = - x – 3; check you answer (E)
If x = -2 and y = 10, what is 2x + 3y = 26? (E)

Note: Teachers' categorization in BT: K = Knowledge; C = Comprehension; Ap = Application; An = Analysis; S = Synthesis; E = Evaluation

Table 4: Sample HOT – HOT items

Write a problem where the expression 2x – 1 can be used to solve the problem (C)
4, 7, 10, 13, __ Find the next term in the sequence. 50th term? Nth term? (Ap)
For 2x2 – bx + 3, what integral values of b will the equation factor? Explain your reasoning (Ap)
John stated that (x + 5)2 = x2 + 25. Explain why John is or not correct. (An)
Explain the differences between $\sqrt{x}$, $-\sqrt{x}$, and $\sqrt{(-x)}$ (An)
How many lines with slope m = 2 go through point (1, 3)? Explain (S)
Given a graph of a real-world linear relationship, find the slope of the line and explain what the slope means in this particular situation (S)
Use you calculator to find a decimal approximation of $\sqrt{3}$. (a) Describe a situation where this approximation would be more useful. (b) Describe a situation where the exact value (i.e., $\sqrt{3}$) would be more useful. Explain your reasoning. (E)

Note: Teachers' categorization in BT: K = Knowledge; C = Comprehension; Ap = Application; An = Analysis; S = Synthesis; E = Evaluation

Test items in table 5 were labeled as HOT by teachers, but they were not classified as HOT in this study since they likely involve the use algorithms or procedures for which most students have been taught and therefore should be familiar on a final exam. Of the 114 items developed by teachers for the thinking skills they labeled as HOT in *BT*, only 51 items (or 45%) were classified as HOT in this study.

Table 5: Sample HOT – LOT items

Write out a set of steps to show how to solve the equation 2x – 7 = 5 (K)
Solve for a: ax + by = az (C)
Given a table of values, write an equation and graph (C)
Find the distance between points (5, 7) and (9, 10) (Ap)
John can mow a yard in 2 hours; Sam can mow the same yard in 3 hours. How long would it take them to mow the yard working together? (Ap)
Which of the following is the graph of 2x – 5y =10? (several choices given) (An)
A line has a slope of 4 and a *y*-intercept of - 3. Which of the following is the equation of the line? (several choices given) (An)
Given the equation x2 + x – 6, make a table of values and graph the equation (S)
The perimeter of a rectangle is 120 feet. If *x* is the width of the rectangle, write a function that represents the area of the rectangle (S)
Evaluate 2x + 3y if x = -1/2 and y = 2/3 ; show your work (E)

Note: Teachers' categorization in BT: K = Knowledge; C = Comprehension; Ap = Application; An = Analysis; S = Synthesis; E = Evaluation

In the analysis of test items, several patterns were apparent. First, it was common among teachers to classify as HOT those items that asked students to "explain" their answers regardless of students' familiarity with the task. Second, several teachers included multiple choice items as representative of HOT but rarely for LOT. In discussing the results with teachers at a later date, several commented that these items might have been labeled HOT by teachers since students first have to "analyze" or "evaluate" the choices before they can answer a question. Third, test items that teachers would likely identify as difficult or take several steps to solve were often classified as HOT regardless of students familiarity with the algorithm or solution methods. And finally, the term "evaluation" was defined or interpreted by many teachers as "finding the value of" instead of how it is defined in *BT*. Thus, despite being given the definitions of the thinking skills in *BT*, many teachers continued to use their own interpretations of these terms or defined these terms as they are used in teaching mathematics.

DISCUSSION

The findings in this study are consistent with research that indicates teachers tend to over-estimate the level of thinking required on test items. For example, Senk, et al (1997) found that the percentage of LOT items on mathematics teachers' tests averaged 68%. Harpster (1999) found that when mathematics teachers were specifically asked to write a single task representative of HOT, 60% developed a task that assessed LOT. However, in the studies cited above, teachers were not using *BT* as a guide to write HOT test items. This study indicated that although mathematics teachers were often able to identify various characteristics of LOT and HOT, many teachers often did not write HOT items. Approximately 55% of test items labeled as HOT by teachers were categorized as LOT in this study. Although none of the teachers listed familiarity in their definitions of HOT, teachers who defined HOT as problem solving, discovering patterns, interpreting information, and conceptual understanding were much more likely to write HOT items than teachers who did not use these terms. In contrast, teachers who defined HOT based on characteristics such as (a) number of steps "required" to solve a task, (b) level of difficulty, or (c) algebra as a higher order subject created LOT items almost 100% of the time.

It is worth noting that teachers who were more familiar with *BT* were no more likely to write HOT items than teachers who were not familiar with *BT*. One teacher, in particular, who was unfamiliar with *BT* wrote and correctly identified HOT items consistently (e.g., table 4, #8). Overall, familiarity with *BT* did not appear to affect teachers' interpretation of *BT* or their ability to write HOT test items. Overall, this research indicates that a key concept in the literature on HOT and in Bloom's (1956, 1971, 1981) discussion on this issue missing from mathematics teachers' interpretations of HOT was the level of *familiarity* students have with the algorithms, methods of solving a problem, or the context / situation of the task needed in a test item. As a result, teachers often misinterpreted *BT* and over-estimated the amount of HOT in test items they created. In this study, *BT* did not appear to affect teachers' perception of HOT or their efforts to write HOT test items for their students.

Although *BT* can be used effectively by mathematics teachers (Kastberg, 2003; Vidakovic, Bevis, & Alexander, 2003), the study lends support that *BT* might not be an effective method of helping mathematics teachers assess for HOT. One alternative is to use a modified version of *BT* specifically developed for mathematics (e.g., Wilson, 1971; Cooney, 1992 cited in Harpster, 1999). It is unknown if these modified taxonomies are more effective in teaching and assessing for HOT, but Bloom, et al (1971) noted that *BT* would likely need to be adapted to meet the needs of individual disciplines. Another alternative would be for mathematics teachers to supplement *BT* (or the modified versions of *BT*) with the definition of HOT used in this study. This will help teachers consider students' familiarity of a procedure or situation when selecting or identifying test items as either LOT or HOT.

And lastly, mathematics teachers can use thinking skill frameworks apart from *BT* specifically designed to assess for HOT in mathematics. For example, the thinking skills framework developed by Smith and Stein (1998) uses four categories of cognitive demands to classify mathematics tasks based on the type of thinking required of the students. Mathematics teachers can also use the 2005 National Assessment of Education Progress mathematics framework. This framework classifies tasks as low, moderate or high complexity (U.S. Department of Education, 2001). Both frameworks are mathematics specific with descriptors consistent with the characterization of HOT as used in this research and as found in the literature on HOT. Although more professional development on teaching for HOT in mathematics is needed, teaching and assessing for HOT is very difficult even with extensive professional development (Henningsen & Stein, 1997; Harpster, 1999). Therefore, more research is needed on creating models of professional development that support teachers' effort to assess for HOT.

LIMITATIONS

There are several limitations in this study. First, although most teachers were familiar with *BT*, given greater training in its use or given illustrative mathematics examples for each thinking skill, mathematics teachers might interpret HOT within *BT* differently. Second, teachers were restricted to developing items only for Algebra I; in the U.S., Algebra I is often taught procedurally and thus emphasizes LOT. Asking teachers to write test items for other subject areas (e.g., Geometry) might have yielded different results. And lastly, teachers were not randomly selected and were restricted to a small geographic area of one state in the southeast U.S. Results may differ for a larger, more diverse sample of teachers. Despite these limitations, this study provides insights into mathematics teachers' interpretations of *BT* and the complexity of assessing for HOT in mathematics.

REFERENCES

Anderson, L., & Sosniak, L. (1994) *Bloom's Taxonomy: A forty-year retrospective*. Ninety-Third Yearbook of the National Society for the Study of Education. Chicago: University of Chicago Press.

Battista, M.T. (1994). Teacher beliefs and the reform movement of mathematics education. *Phi Delta Kappan*, 75, 462– 467.

Black, P., & Wiliam, D. (1998). Inside the black box. *Phi Delta Kappan*, 80(2) 139-147.

Bloom, B. (Ed.) (1956). *Taxonomy of educational objectives: Book I, cognitive domain*. New York: Longman Green.

Bloom, B. S., Hastings, J.T., & Madaus, G.F. (Eds), (1971). *Handbook on formative and summative evaluation of student learning.* New York, NY: McGraw – Hill

Bloom, B., Madaus, G., & Hastings, J. (1981). *Evaluation to improve learning.* McGraw-Hill: New York, NY.

Cooney, T.J. (1992). *A survey of secondary teachers' evaluation practices in Georgia*. Athens, GA: University of Georgia.

Costa, A. (Ed) (2001). *Developing minds: A resource book for teaching thinking.* Alexandria, VA: ASCD.

de Lange, J. (1987). Mathematics, insight, and meaning. Utrecht, the Netherlands: OW & OC, Rijksuniversiteit te Utrecht.

Donovan, M. S., & Bransford, J. D. (Eds.). (2005). *How students learn: Mathematics in the classroom.* Washington, DC: National Academies Press.

Fleming, M., & Chambers, B. (1983), Teacher-made tests: Windows to the classroom. In W.E. Hathaway (Ed*.), Testing in the schools* (pp. 29-47). San Francisco, CA: Jossey-Bass.

Harpster, D. L. (1999*). A study of possible factors that influence the construction of teacher-made problems that assess higher-order thinking skills*. (Doctoral Dissertation, Montana State University-Bozeman, 1999) UMI 9927881.

Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K., Wearne, D., Murray, H., Olivier, A,, & Human, P. (1997). *Making sense: Teaching and learning mathematics with understanding*. Portsmouth, NH: Heinemann.

Henningsen, M., & Stein, M. K. (1997). Mathematical task and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, 28(5), 524-549.

Kastberg, S. (2003). Using Bloom's Taxonomy as a framework for classroom assessment. *Mathematics Teacher*, 96(6), 402-405.

Kulm, G. (Ed.) (1990). *Assessing higher-order thinking in mathematics*. Washington, DC: Association for the Advancement of Science

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., & Chrostowski, S. J. (2004). *TIMSS 2003 international mathematics report: Findings from IEA's Trends in International Mathematics and Science Study at the eighth and fourth grades*. Chestnut Hill, MA: Boston College.

National Center for Education Statistics. (1996). *High school seniors' instructional experiences in science and mathematics*. Washington, D.C.: U.S Government Printing Office.

National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics. Reston*, VA: Author.

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Quellmalz, E. S., (1987). Developing reasoning skills. In J. R. Baron, & R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice*. New York: Freedman Press.

Resnick, L. (1987). *Education and learning to think*. Washington, DC: National Academy Press.

Resnick, L. B., & Resnick, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. R. Gifford & M. C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement and instruction* (pp. 37–75). Boston: Kluwer.

Reynolds, W. M., & Menard, K. A. (1980). *An investigation of teachers' test construction practices*. Paper presented at the annual meeting of the National Council of Measurement in Education. Boston.

Schmalz, R. (1973). Categorization of questions that mathematics teachers ask. *Mathematics Teacher*, 66(7). Reston, VA: NCTM

Senk, S. L., Beckmann, C. E., & Thompson, D. R. (1997). Assessment and grading in high school mathematics classrooms. *Journal for Research in Mathematics Education*, 28(2), 187-215

Schoenfeld, A.H. (1988). When good teaching leads to bad results: The disasters of 'well-taught' mathematics courses. *Educational Psychologist*, 23(2), 145-166.

Smith, M., & Stein, M. K. (1998). Selecting and creating mathematical tasks: From research to practice. *Mathematics Teaching in the Middle School*, 3(5), 344–350.

Stein M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation*, 2(1), 50-80.

Stiggins, R. J., Griswold, M. M., & Wikelund, K.R. (1989). Measuring thinking skills through classroom assessment. *Journal of Educational Measurement*, 26(3), 233-246.

US Department of Education (2001). *Mathematics assessment framework for 2005: National Assessment of Educational Progress*. Washington, DC: National Assessment Governing Board.

Vidakovic, D., Bevis, J., & Alexander, M. (2003). Bloom's Taxonomy in developing assessment items. *Journal of Online Mathematics and its Applications*. Retrieved 7/26/2007, from http://mathdl.maa.org/mathDL/4/?pa=content&sa=viewDocument&nodeId=504

Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives*, 10(12).

Wilson, J. W. (1971). Evaluation of learning in secondary school mathematics. In B.S. Bloom, J. T. Hastings, & G. F. Madaus (Eds), *Handbook on formative and summative evaluation of student learning*. New York, NY: McGraw – Hill

APPENDIX

Teacher Handout: Thinking skills in Bloom's Taxonomy*

In Alphabetical Order

ANALYSIS — Ability to breakdown concepts into their component parts so that the hierarchy of ideas is clear; clarify existing information by examining parts and relationships; identify relationships and patterns; identify errors and logical fallacies and where possible, correcting them.

APPLICATION — Ability to use prior knowledge within a new situation. This involves bringing together the appropriate information, procedures, generalizations, or principles that are required to solve a problem without being told to do so or without any specific or immediate cues.

COMPREHENSION – Ability to understand what is being communicated and make use of the material or idea being communicated without necessarily relating it to other material or seeing its fullest implications. This may be shown by (a) translating concepts from one form to another, (b) interpreting and summarizing facts, and (c) contrasting, comparing, or predicting consequences.

EVALUATION—Ability to judge or assess the value of material and methods for given purposes as well as to compare and discriminate between ideas; assessing the reasonableness and quality of ideas including establishing criteria (setting standards for making judgments) or verifying (confirming the accuracy of claims).

KNOWLEDGE –Ability to recognize or recall facts, methods, processes, patterns, structures, basic concepts, conventions, principles, and theories.

SYNTHESIS —Ability to work with pieces, parts, elements, etc., and arranging and combining them in such a way as to form a whole or constitute a pattern or structure not clearly there before. This may also include generalization from given facts, relating knowledge from several areas, predicting, and drawing conclusions.

* Summarized from Bloom, B. (Ed.) (1956). *Taxonomy of educational objectives: Book I, cognitive domain*. New York: Longman Green

Author : **Tony Thompson**

E-mail : thompsonan@ecu.edu

Address : Department of Mathematics and Science Education
Flanagan Bldg 327, East Carolina University
Greenville, NC 27858

Phone : 1- 252-328-9358