AC 2009-307: ANALYZING RIGOR AND RELEVANCE IN SCIENCE AND MATHEMATICS CURRICULA

Doug Kueker, Vivayvic

The Instructional Design and Curriculum Evaluation Lead for Vivayic, Inc. Prior to joining Vivayic, Inc., in September 2006, Doug worked for the National FFA Organization as a Project Management Specialist. In his professional career, Doug has led and participated in more than 25 national curriculum design, development, implementation, and evaluation projects. He received his M.S. Ed from Purdue University in December 2007and holds a Bachelors Degree in Agricultural Science Education from the University of Missouri. Address: 69 Eagle Crest Road, Lake Ozark, MO 65049 Telephone: (573) 286-0597 E-mail: dkueker@vivayic.com

Pam Newberry, Project Lead the Way

The Director of Strategic Curriculum Initiatives for Project Lead The Way, Inc. Prior to joining Project Lead The Way, Inc., in July 2002, she served as the Associate Director for the International Technology Education Association?s Technology for All Americans Project for five years. She taught technology education and mathematics for 10 years. During that time, she was an Albert Einstein Fellow in 1996 and received the Presidential Award for Excellence in Mathematics Teaching in 1994.

Address: 177 Stone Meadow Lane, Wytheville, VA 24382

Telephone: (276) 228-6502

E-mail: pampltw@embarqmail.com

Analyzing Rigor and Relevance in Science and Mathematics Curricula

Introduction

To be successful in today's educational climate schools must select and implement rigorous and relevant science and mathematics curricula. Since the publication of a *Nation at Risk* [19] (1983), schools across the U.S. have sought to meet a growing demand from business and government to increase the level of academic rigor in disciplines, such as mathematics and science for all students. Enhancing the quality of mathematics and science curricula remains as a priority for our nation's educational institutions. In fact, accountability legislation, such as *No Child Left Behind* [36] (2001) creates additional pressure for educators to take a serious look at curricula, especially in the areas of mathematics and science, selected to support local, district, state, and national learning priorities. Creating and implementing both rigorous and relevant mathematics and science curricula is also necessary to support U. S. business and industry in meeting employment and training needs for science, technology, engineering, and mathematics (STEM) careers (NSF[22], 2004).

Despite the standards based movement to improve science and mathematics curricula in schools, the nation continues to lag behind others. The *1999 Trends in International Mathematics and Science Study* looked at the ways that mathematics and science instruction differs among seven countries. High-performing countries avoided reducing mathematic and scientific tasks to mere procedural exercises and they placed greater cognitive demand on students by encouraging them to focus on concepts, explaining connections among those concepts, as well as explaining their reasoning when solving a problem (Hiebert[10], 2003; Roth[28], 2006). Improving the cognitive rigor and relevance of the instructional tasks in U. S. curricula plays an important role in the quest to provide a high-quality, globally-competitive educational system that enhances students' educational career options and meets the needs and priorities established by business and government.

The Role of Rigor and Relevance

Daggett[7] (2005) suggests that lasting gains in student achievement come from applying high-rigor expectations in relevant, real-world settings. Daggett's framework for improving school curricula, considers both cognitive rigor and relevance. In his framework, rigor is defined as the level of cognitive demand, or the quantity and quality of the cognitive processes, required to complete an instructional or assessment task. Relevance, on the other hand, deals with the context in which the content is applied. Context of application varies based upon the degree to which the context in which the content is to be applied, or transferred, approximates the real world.

Educational research supports Daggett's assumptions about the importance of rigorous and relevant instructional tasks to student learning. Students demonstrate gains on measures of reasoning and problem-solving when instructional tasks are set up and enacted at a high level of cognitive demand (Henningson & Stein[9], 1997). Encouraging students to explain connections among concepts and offering justification are just two examples of practices that foster high cognitive demand (AERA [3,4], 2006, 2007; Stein, Grover, Henningsen[31], 1996). Further, Lave &

Wenger[16] (1991) and Brown, Collins and Duguid[5] (1989) demonstrates that efforts to make learning academic concepts relevant through application of knowledge in real-world contexts improves knowledge retention and transfer and enhances students' motivation to learn. Carefully designed instructional tasks that mirror real or simulated investigations can substantially improve understanding of complex ideas and lead to long-term understanding of discipline related knowledge (Kimbell[12], 1997; Kimbell, Stables, Wheller, Wosniak, & Kelly[13], 1991; Minner, Levy, & Century[18], 2007; Schauble, L., Klopfer, L.E., & Raghavan, K.,[29] 1991). These experiences may also have the added advantage of attracting and supporting a diverse group of science learners to meet the need for an educated workforce (Linn, Lee, Tinker, Husic, & Chiu[17] 2006).

A Theoretical Framework for Analyzing Rigor and Relevance

Despite pressure to implement rigorous and relevant science and mathematics curricula, consistent methods for analyzing and describing the rigor and relevance of any curriculum are still emerging. A framework for analyzing our curricula is described by Andrew Porter[26] (2004). Porter defines curricula analysis as the systematic process of isolating and analyzing targeted features of a curriculum. Curricula analysis most commonly involves describing and isolating a particular set of content (e.g., mathematics content, science content, or language arts content) in a curriculum and then analyzing the performance expectations, or cognitive demand, that describe what students are to know and do with the content. Creating a consistent method for analyzing curricula is made complex because there are multiple viewpoints from which one may examine a curriculum. Porter[25, 26] (2002, 2004) makes distinctions regarding the four levels at which curricula analysis may occur. Table 1 reflects the focus of curricula analysis at each of the four levels.

Table 1. Primary Focus of Curricula Analysis at Each Dimension of a Curriculum

Level	Primary Focus of Curricula Analysis
Intended Curriculum	Analysis is concerned with examining the content (e.g., declarative, procedural, tactile, and situative knowledge) and the performance expectations, which is the level at which a student is expected to know and use the content as it is communicated in the documents and materials created to guide instruction and assessment.
Enacted Curriculum	Analysis is concerned with examining the content and the performance expectations as the instructor enacts it.
Assessed Curriculum	Analysis is concerned with examining the content and performance expectations represented by the tasks, questions, and performance tasks contained in assessment materials.
Learned Curriculum	Analysis is concerned with measuring the content and level at which learners enact the performance expectations in a targeted context.

Porter[25, 26] (2002, 2004) contends that building a comprehensive understanding of curricula in any discipline requires some form of analysis at each of these four levels. Currently, most of our understanding of the science and mathematics curricula implemented in our schools comes from evaluating the assessed curriculum (Webb[33,34], 1997, 2002). Recent efforts have started to build our understanding of the enacted curriculum in disciplines, such as mathematics, science and language arts (Surveys of the Enacted Curriculum[32], 2009). However, few efforts currently focus on employing systematic and scientific approaches to analyze and understanding the intended curriculum in any discipline.

Stein, Grover & Henningsen[9] (1996) demonstrate that the intended instructional tasks, as presented in curricular materials, influence how teachers enact those instructional tasks and ultimately how students carry out the tasks in the classroom. As Porter[25, 26] (2002, 2004) points out, analysis of the intended curriculum will not ensure students learn the targeted concepts. However, Porter also points out that defining high quality, rigorous, and relevant instructional tasks in the intended curricula is a critical input that influences students' opportunity to learn.

Additional studies that apply this theoretical framework to examine each of the four levels are needed. The authors suggest that the framework may be applied to generate and report data regarding the level of rigor and relevance in high priority areas, such as science and mathematics content in science, technology, engineering, and mathematics (STEM) curricula. Reports that examine the various levels of a curriculum can provide educators with the data they need to sort among the options in today's instructional materials market. Further, applying this framework encourages accountability among those who develop and distribute ready-to-use curricula.

Purpose, Scope, and Researcher Role

The method and data discussed in this paper are derived from an initial project in which Project Lead the Way Inc. commissioned a third-party evaluator to investigate claims regarding the rigor and relevance of mathematics and science content included in the Project Lead The Way® *Introduction to Engineering*™ course. The present paper is a derivative of the initial evaluation study commissioned by Project Lead the Way Inc. The goal of this paper is to illustrate, with interested audiences, how Porter's[25, 26] (2002, 2004) theoretical framework may be applied to evaluate claims regarding the rigor and relevance of mathematics and science content found in commercially available STEM curricula.

The method and data introduced in this paper are only concerned with examining the intended curriculum. In other words, this paper addresses an analysis of the performance expectations and context of application for science and mathematics content as they are represented in the documents and materials created to guide instruction and assessment in the STEM curricula selected for the analysis. Applying Porter's[26] (2004) framework to analyze the similarities and differences among the other facets of a curriculum would certainly be a worthy endeavor; however it is beyond the scope of this paper.

The primary author for this paper served as the primary investigator from the third-party evaluator's organization. The second author for this article served as a representative of the Project Lead the Way organization during the initial evaluation study. The primary author and colleagues in his organization were solely responsible for the method design, data collection, data analysis, and initial reporting. As a third-party evaluator, the primary author bore the burden of maintaining integrity and transparency throughout the evaluation process. The representative from Project Lead The Way Inc. was only involved in providing access to the curricula and answering questions of clarification from the third-party evaluators before and after the process. The second author also served in a K-12 STEM content expert capacity for the development of this paper.

Methodological Paradigm and Claims Evaluated

The evaluation study described in this paper is conducted within a reality-oriented correspondence theory perspective. Studies conducted within a reality-oriented paradigm are concerned with testing claims of effectiveness by bringing data, including qualitative data, to bear on those claims to determine if evidence exists to corroborate the claims (Patton[24], 2002). As such, the evaluation study was designed to determine the degree to which objective, reliable, and valid data corroborated two claims made regarding the Project Lead The Way® Introduction to Engineering Design™ course and the underlying instructional methodology. The claims evaluated were:

Claim 1: The instructional tasks are cognitively rigorous and promote application of content in relevant, real-world contexts.

Students use a problem-solving model to improve existing products and invent new ones. They learn how to apply this model to solve problems in and out of the classroom. Emphasis is placed on analyzing potential solutions and communicating ideas to others. This approach is called activities-based learning, project-based learning, and problem-based learning or APPB-learning. Research shows that schools practicing APPB-learning experience an increase in student motivation, an increase in cooperative learning skills and higher-order thinking, and an improvement in student achievement (Newberry & Hughes[23], 2006).

Claim 2: The course objectives integrate mathematics and science content with technology and engineering instruction through rigorous and relevant instructional tasks.

PLTW's curriculum makes mathematics and science relevant for students. The curricular objectives integrate math and science content as defined by NCTM[20] (2000), NRC[21] (1996), and AAAS[1] (1993) and with standards for technological literacy as defined by the ITEA[11] (2000) and enduring engineering concepts as identified by ABET, Inc. Criterion 3-Outcomes A-K (ABET[2], 2007) By engaging in hands-on, real-world projects, students understand how the mathematics and science skills they are learning in the classroom can be applied in real-world engineering and technological design problems (PLTW[27], 2007).

Methods and Procedures

Content analysis was selected as the research technique to guide the collection of data used to evaluate the claims made about the course. Content analysis is the systematic classification, tabulation and interpretation of key symbols in various forms of communication, including documents and texts (Krippendorff[14], 2004). Content analysis as a method is particularly well matched to the demands of evaluating the intended curriculum. The intended curriculum as pointed out by Porter[26] (2004) usually consists of documents containing standards, objectives, instructional tasks, and related support materials. As such, the curriculum analysis framework lends itself to analyzing the content of these documents in a systematic and scientific way.

Content analysis, as defined by Krippendorff[14] (2004), emphasizes the following phases:
- Phase 1: Designing a coding scheme based upon the purposes of the analysis
- Phase 2: Identifying the scope of the units to be analyzed
- Phase 3: Selecting a sample of the units to be analyzed
- Phase 4: Analyzing the units using a pre-determined and validated coding scheme
- Phase 5: Checking for reliability in the coding of the units
- Phase 6: Reducing the coded data to manageable representations using statistical techniques
- Phase 7: Narrating the findings

The following sections describe, in detail, the specific procedures used to enact the content analysis framework outlined here. Since the study was conducted from a reality testing paradigm special attention was paid to issues of validity, reliability, and objectivity. As such, measures to reduce bias in the method selected are discussed at length throughout.

Designing and Validating a Coding Scheme

The first claim made by Project Lead the Way, Inc. regards the rigor and relevance of instructional tasks in the curricula. To bring data to bear on this particular claim, the evaluator made it a priority to identify a coding scheme for evaluating instructional tasks that accounted for both cognitive demand (rigor) and the context for applying the content (relevance). Webb's[33, 34] (1997, 2002) Depth of Knowledge model provided the basis for the coding scheme. It should be noted that the term knowledge, as it is used here, might be somewhat of a misnomer. The Depth of Knowledge model was chosen because the descriptions for each level in the model encompass both cognitive demand (rigor) and context for application (relevance). Additionally, the model has been adapted to analyze science and mathematics content in a number of statewide curricula (Webb[34], 2002).

The model was adapted for use as the coding scheme in this analysis. The adaptation was necessary to ensure that the coding scheme reflected the terms, concepts, and context of application outlined in an integrated STEM curricula. To increase objectivity in the coding scheme content, the coding scheme and descriptions were defined before the evaluator reviewed any of the course materials. The version of the model used was adapted from the definitions of the Depth of Knowledge Levels for Science and for Mathematics (Webb[34], 2002). The evaluators collected input from experts in the field of engineering and technology to refine the descriptions for each level of the model. A final version of the coding scheme was validated for face validity with a second group of engineering and technology professionals before use in the study.

The following table reflects an overview of the adapted model used for the coding scheme. In the model, as the levels increase so does the cognitive demand (rigor) and proximity to real-world application (relevance). Thus, an objective, instructional task, and assessment item assigned to level one will have the lowest cognitive demands and will likely reflect application to tasks in the domain. Whereas, an item assigned to level four will have the highest cognitive demands and will likely reflect application in real-world unpredictable contexts. A detailed description of the cognitive demands and the context of application represented by each level is located in Appendix A.

Webb's Depth of Knowledge Model	
DOK Level	**Title of Level**
1	Recall and Reproduction
2	Skills and Concepts
3	Short-term Strategic Thinking
4	Extended Strategic Thinking

The second claim made by Project Lead the Way, Inc. regards the integration of mathematics and science content into rigorous and relevant technology and engineering based tasks. To bring data to bear on this particular claim the evaluator chose to employ nationally recognized standard frameworks for both mathematics and science to guide the isolation and identification of mathematics and science content in the *Introduction to Engineering Design*™ course. Mathematics concepts are defined by *Principles and Standards for School Mathematics* as published by the National Council of Teachers of Mathematics[20] (NCTM). Science concepts are defined by *National Science Education Standards* as published by the National Research Council[21] (NRC).

Defining the Units Included in the Analysis

The performance and assessment objectives, as stated in the course lesson plans, were selected as the primary unit for the analysis. In Project Lead The Way® curricula the concepts and performance objectives are statements of expectation regarding what students are to know and be able to do as well as the context in which the knowledge and skill are to be applied following completion of each lesson. The performance objectives are based upon a set of overarching concepts for each course. These concepts represent what students are to understand at the end of the course and are translated into a set of performance objectives, which then guide development of the instructional activities, projects, and problems used to teach the concepts. In order to understand fully the content described in the performance objectives, it was necessary to examine thoroughly the related course concepts. Likewise, to understand the performance expectations expressed in the performance objectives it was necessary to examine closely the course activities, projects, and problems as these artifacts also share an integral linkage.

The assessment objectives are statements of expectation regarding how students are to demonstrate their understanding of the course concepts. In the lesson, the statements are written in the form of behavioral objectives and both course content and performance expectations are presented. The statements are located under a section called "Assessment" in the Project Lead the Way® curricula. The assessment objectives are organized by Wiggins and McTighe's[35] (2005) six facets of understanding and are provided as an expectation for students to demonstrate real-world application and critical thinking with regard to the course concepts as students complete each lesson. Throughout the method described here, the performance and assessment objectives are generically referred to as course objectives.

The course objectives were selected as the unit of analysis for three primary reasons. First, the study at hand was concerned with isolating science and mathematics content. The objectives are a proxy for the science and mathematics related content to be integrated into instruction. Second, the study at hand sought to measure the level of cognitive expectations expressed in the course materials. The objectives identify expectations regarding the kinds of thinking processes students are expected to employ as they complete the instructional tasks. Finally, the study at hand sought to isolate and measure the various contexts in which students were expected to apply the content. The objectives also outline expectations regarding the context of application.

Sample of Units Included for Analysis

Since the purpose of the analysis was to generate data to evaluate the claims about the entire course the entire population of performance and assessment objectives were included in the sample analyzed. Thus, 168 course objectives were included as the sample for the analysis. The entire population of objectives were included in the analysis to increase confidence that the results reported were in fact a representation of the entire course, not a smaller sub-section of the course or an inadvertently biased sample of the course objectives.

Applying the Coding Scheme to Analyze the Units

A reviewer familiar with using the Depth of Knowledge Levels began the analysis by coding a Depth of Knowledge Level to each objective in the *Introduction to Engineering Design*™ (IED) course. As the reviewer coded each course objective, the following pre-determined guidelines were employed. These guidelines were outlined by Webb[33,34] (1997, 2002) as advice to other evaluators using the model to code objectives.

- The Depth of Knowledge Level assigned should reflect the level of work students are most commonly required to perform in order for the response to be deemed acceptable.
- The Depth of Knowledge Level should reflect the complexity of the cognitive processes demanded by the task outlined by the objective, rather than its difficulty. Ultimately, the Depth of Knowledge Level describes the kind of thinking required by a task, not whether or not the task is "difficult."
- If there is a question regarding which of two levels, a statement addresses, such as Level 1 or Level 2, or Level 2 or Level 3, it is appropriate to select the higher of the two levels.
- The Depth of Knowledge Level should be assigned based upon the cognitive demands required by the central performance described in the objective.
- The objective's central verb(s) alone is not sufficient information to assign a Depth of Knowledge Level. Evaluators must also consider the complexity of the task and

information, conventional levels of prior knowledge for students at the grade level, and the mental processes used to satisfy the requirements set forth in the objective.

The reviewer worked chronologically through the curriculum starting with the first unit and continued working through each lesson in order. Performance objectives for the lesson were evaluated first followed by a review of the lesson's assessment objectives. The level assigned represented the highest level of cognitive processing demanded for a student to satisfactorily demonstrate attainment of the objective. To understand the level of expectation for "satisfactory attainment" the reviewer used the *Introduction to Engineering Design*™ course materials to investigate the nature of the content, the instructional treatments used, and the assessment rubrics related to each objective. Once the supporting information was reviewed, a level was assigned. Periodically, the evaluator reviewed previously evaluated objectives to ensure consistent application of the four levels among objectives with a similar focus. If a discrepancy occurred, the reviewer re-evaluated the information available for both objectives and made a final assignment.

As the reviewer coded each objective using the Depth of Knowledge model, each objective was also coded with respect to its emphasis on mathematics or science content standards. "Emphasis on mathematics or science," for this analysis was defined as active employment or use of the mathematics or science skill and knowledge as listed in the standards to meet some established expectation. Thus, both the subject matter and the expectation established for employing that subject matter were considered as each objective was analyzed.

Objectives were assigned to one of three categories: direct match, possible match, and no match. Direct matches between concepts addressed in the objectives and standards were flagged by identifying the specific standard(s) in the respective national framework that are addressed by the objective. Objectives that did not have exact key word matches, but shared a relationship to concepts in either set of standards were flagged for further analysis. For these objectives, the reviewer used the course materials and both sets of standards to investigate each of those objectives flagged as potential matches in the initial pass. If the subject matter did in fact emphasize either science or mathematics content then that objective was also assigned to be included in the analysis regarding science and mathematics content in the course.

Checking for Reliability in the Coding

This method is mostly heuristic and is sensitive to the expertise of the initial curriculum reviewer. Eisner[8] (1991) discusses the role of perceptivity in conducting qualitative research. Perceptivity according to Eisner is a researcher's ability to "notice" what is important in a "setting." In part, perceptivity is developed as a result of professional experience. The third-party evaluators have extensive experience working with the Depth of Knowledge model as a coding scheme to evaluate curricular objectives. The primary investigator for this study has reliably coded more than 5,000 objectives in multiple disciplines using the model. However, to ensure the reliability of the primary investigator's application of the coding scheme a second evaluator, also experienced with the Depth of Knowledge model, was involved in the coding process.

To check the reliability of the initial coding process, a sample of 33 randomly selected course objectives out of the total 168 (containing both performance and assessment objectives) were independently analyzed by a second trained evaluator. The sample of objectives was selected using a simple randomization technique. The second evaluator assigned a Depth of Knowledge Level, one through four, to each of the 33 objectives in the sample using the same descriptions and procedures described earlier. An attribute agreement analysis was conducted using Mini-Tab® to evaluate the level of concurrence between the two evaluator's Depth of Knowledge assignments. Out of 33 statements inspected 28, or 84.85%, of the assigned levels provided by both evaluators matched. To evaluate the level of agreement Cohen's kappa (κ) and Kendall's coefficient of concordance (W) were both employed.

Cohen's kappa (κ) measures absolute agreement between two raters who each classify a set number of items into a set of mutually exclusive categories (Cohen[6], 1960). This statistic is considered more robust than calculating only a simple percent agreement as it considers chance agreement. The analysis reports a value between 0 and 1. κ values closer to one indicate stronger agreement between the raters involved. Landis & Koch[15] (1977) indicate the following rule-of-thumb for evaluating Cohen's kappa measures: $\kappa = 0.41$ to 0.60 reflects moderate inter-rater agreement, κ=0.61 to 0.80 substantial and κ=0.81-1.00 almost perfect. Following analysis of the ratings the level of agreement between the two raters on the 33 items was found to be substantial $\kappa=0.75983$ ($P < 0.000$). This represents a level of agreement significantly different (α=.05) from those that would be achieved by chance.

Kendall's coefficient of concordance (W) is another non-parametric statistic often used in combination with Cohen's kappa to evaluate the degree of concordance or discordance between independent raters (Siegel & Castellan[30], 1988). As a comparison, Cohen's kappa coefficient reflects the absolute agreement among multiple raters. κ doesn't take into account the order of the scores or the severity of misclassifications among raters (i.e. one rater assigns the objective to Level 1 while the other assigns the objective to Level 4). The coefficient of concordance statistic is, however, sensitive to ordering and to the seriousness of misclassifications among raters. Kendall's coefficient of concordance can range from 0 to 1; the higher the value of Kendall's, the stronger the association. After data analysis W= 0.911609 (df=32, $\chi 2$=58.3430, $P<0.0030$). This indicated an acceptable level of concordance between the two rater's assignment of Depth of Knowledge Levels that is significantly different (α=.05) from those achieved by chance alone.

While attribute agreement analysis provided evidence to suggest that the first evaluator's original assignments were acceptable, a review of data output from the Cohen's κ analysis was used to reveal areas where absolute agreement differed most. Based upon this analysis it was determined that the evaluators most consistently matched on assigning Level 4 (κ=1, $P<0.000$). The source of disagreement between the evaluators was mostly found with assignment of Levels 2 & 3 (κ=0.67836 and κ=0.69274, P=0.000 respectively). To address this issue, the primary investigator reviewed those objectives originally assigned a Level two and three to determine if any changes were merited based upon feedback from the second evaluator. Only a few adjustments were made to the originally assigned levels based upon this second review. Appendix B contains a sample of the objectives showing the Depth of Knowledge Level to which they were assigned.

Before finalizing the objectives to include in the analysis regarding the science and mathematics content, a sample of objectives identified in each of the three categories (direct match, possible match, and no match) were also shared with a reviewer. The reviewer selected was also familiar with both sets of standards to ensure valid identification of mathematics and science concepts. Appendix C shows example objectives included in this analysis and the standards to which they were linked. It is important to note that due to the nature of the objectives in the engineering context many of the objectives were identified to emphasize both science and mathematics concepts.

Reducing the Coded Data to Manageable Representations

Once all coding was complete the data were reduced using descriptive statistics to begin to analyze the data and form conclusions regarding the claims about the curricula. To reduce the data, we created a series of histograms to show how often the four Depth of Knowledge Levels occurred throughout all course objectives in general. Next, we ran the same analysis for objectives that were identified to emphasize mathematics. Finally, we created a series of histograms to show how often the four Depth of Knowledge Levels occurred throughout the course objectives with a science emphasis. In addition to the histograms, we also generated other measures of central tendency, such as mean, median, mode, and standard deviation.

Narrating the Findings

As discussed, the procedures employed were designed to determine the degree to which data corroborated two claims made regarding the Project Lead The Way® *Introduction to Engineering Design*™ course and the underlying instructional methodology. This section narrates the findings as they relate to each separate claim.

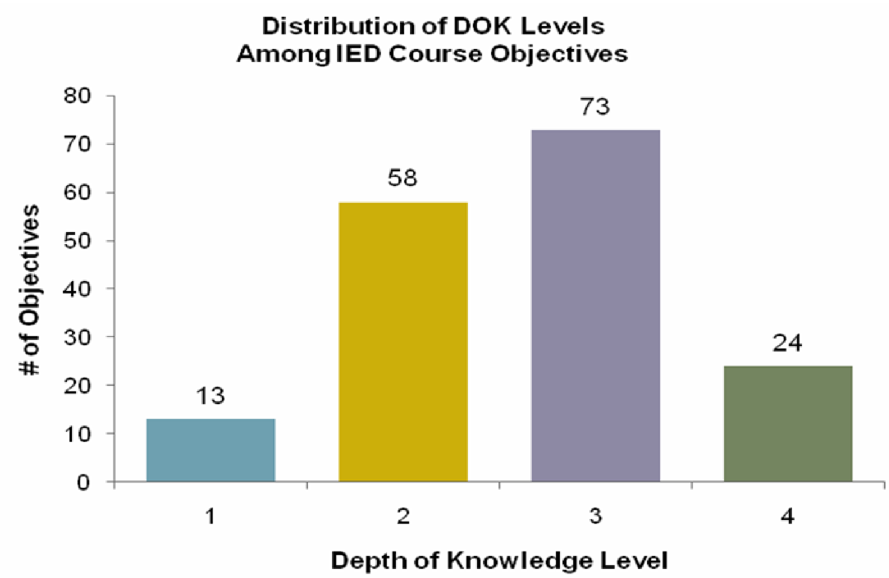
Claim 1: The instructional tasks are cognitively rigorous and promote application of content in relevant, real-world contexts.

The course objectives represented a range of cognitively rigorous tasks with application to a range of contexts. The data demonstrate the following breakdown by Depth of Knowledge level:

- 7.74% of the course objectives were assigned to Level 1 – Recall and Reproduction;
- 34.52% were assigned to Level 2 – Working with Skills and Concepts;
- 43.45% were assigned to Level 3 – Short Term Strategic Thinking and
- 14.29% were assigned to Level 4 – Extended Strategic Thinking.

Keep in mind, the Depth of Knowledge Levels defined were such that higher Depth of Knowledge Levels (Levels 3 and 4) reflect performance and assessment objectives that have a higher degree of rigor (cognitive demand) and relevance (proximity to real-world application). In this case, 57.73% of the performance and assessment objectives demonstrated a high degree of cognitive rigor and a closer proximity to real-world application. Chart 1 graphically shows the distribution of all course objectives among the four Levels in the Depth of Knowledge model.

Chart 1. Analysis of Rigor and Relevance in All Course Objectives



DOK Level	*Frequency*
1	13
2	58
3	73
4	24

Count	168
Mean	2.59
Mode	3
*SD**	0.815
Median	3.00

*SD=Standard Deviation

Claim 2: The course objectives integrate mathematics and science content with technology and engineering instruction through rigorous and relevant instructional tasks.

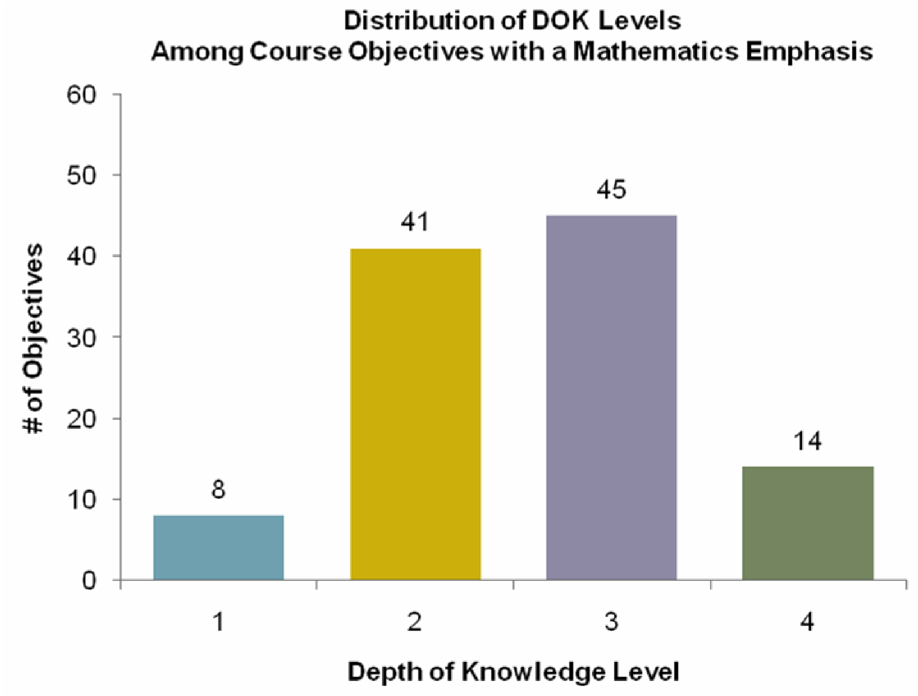
With regard to science and mathematics content, out of 168 objectives in the *Introduction to Engineering Design*™ course 88.69% (149) of the 168 total objectives in the *Introduction to Engineering Design*™ course were identified to emphasize either, or both, mathematics and science content standards. One hundred and eight (64.28%) were identified for emphasizing one or more of the mathematics standards established by NCTM[20]. One hundred fourteen (67.85%) objectives were identified for emphasizing one or more of the stated science standards established by the NRC[21]. Additionally, 47.65% (71) of the 168 course objectives included in this analysis demonstrated a dual emphasis on mathematics and science concepts.

Looking across the entire course, generally objectives that emphasize mathematics and science established an expectation for the use of short-term strategic thinking (Depth of Knowledge Level 3).

- 42.98% of objectives emphasizing science were assigned to Depth of Knowledge Level 3
- 41.67% of objectives emphasizing mathematics were assigned to Depth of Knowledge Level 3

Charts 2 and 3 demonstrate the distribution of course objectives by Depth of Knowledge Levels. As discussed before, the higher the Depth of Knowledge Level the higher the expectations for rigor and relevant application.

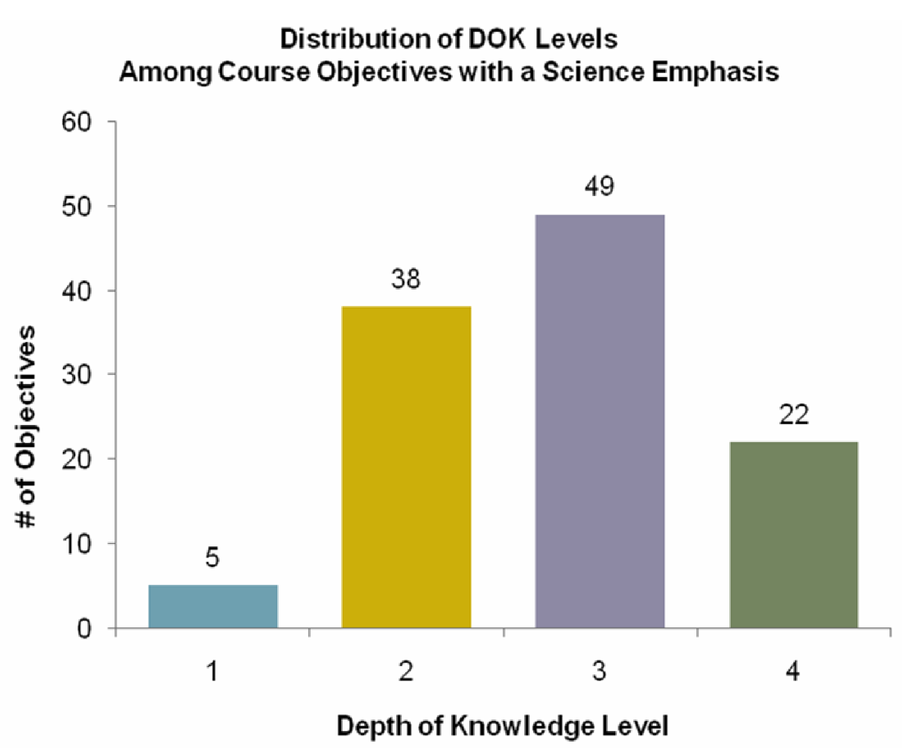
Chart 2. All Course Objectives Emphasizing Mathematics



DOK Level	*Frequency*
1	8
2	41
3	45
4	14

Mean	2.60
Mode	3
Median	3
SD	0.808
Count	108

Chart 3. All Course Objectives Emphaizing Science



DOK Level	*Frequency*
1	5
2	38
3	49
4	22

Mean	2.77
Mode	3
Median	3
SD	0.810
Count	114

A closer look at the rigor and relevance of the mathematics and science content by the two kinds of objectives (performance and assessment) outlined in the course materials, reveals additional data with which to evaluate the second claim. From the data in Table 2, one can see that performance objectives for both mathematics and science place strong emphasis on working with skills and concepts, Level 2, and short-term strategic thinking, Level 3. On the other hand, assessment objectives tended to place a stronger emphasis on integration of skills and concepts through short-term strategic thinking – Depth of Knowledge Level 3, and extended strategic thinking – Depth of Knowledge Level 4. This is likely a reflection of the course design, as students initially learn and apply concepts through the instructional focus (performance objectives) prior to being expected to employ those same skills and knowledge to solve both projects and problems as part of assessment in the course.

Table 2. Detailed Analysis of Course Objectives by Depth of Knowledge Level

	Level 1	**Level 2**	**Level 3**	**Level 4**
	Increasing rigor → Increasing relevance →			
Performance Objectives				
Emphasizing Mathematics	11.76%	44.12%	36.77%	7.35%
Emphasizing Science	8.2%	44.26%	36.06%	11.48%
Assessment Objectives				
Emphasizing Mathematics	0	27.5%	50%	22.5%
Emphasizing Science	0	20.75%	50.95%	28.30%

Table 3 demonstrates how the data may be analyzed by units of instruction within the course to provide another view of the rigor and relevance of mathematics and science content throughout the course. From the data in Table 3, one can see the trends for objectives emphasizing both mathematics and science across all four units of the *Introduction to Engineering Design*™ course. Generally, the objectives emphasizing both mathematics and science in Units 1 thru 3 spanned all four Depth of Knowledge levels, but tended to concentrate on level 2, working with skills and concepts, and level 3, short-term strategic thinking. In contrast, objectives with a mathematics and science emphasis in Unit 4 had a higher concentration of objectives at level 3 (short-term strategic thinking) and level 4 (extended strategic thinking). This also seems to be a reflection of the course design. The last unit of the course appears to employ higher-level application of all skills and concepts learned in the course in the context of solving a real-world problem where the outcomes are not predictable.

Table 3. Analysis of Science and Mathematics Content by Course Unit				
	Level 1	**Level 2**	**Level 3**	**Level 4**
Science Content	Increasing rigor → Increasing relevance →			
Unit 1 Objectives (n=32)	6.25%	34.37%	46.88%	12.5%
Unit 2 Objectives (n=34)	0	35.29%	41.18%	23.53%
Unit 3 Objectives (n=35)	8.57%	34.28%	42.86%	14.29%
Unit 4 Objectives (n=13)	0	23.08%	38.46%	38.46%
Mathematics Content				
Unit 1 Objectives (n=39)	5.13%	43.59%	41.02%	10.26%
Unit 2 Objectives (n=47)	8.51%	29.79%	48.93%	12.77%
Unit 3 Objectives (n=18)	11.11%	55.56%	27.78%	5.55%
Unit 4 Objectives (n=4)	0	0	25%	75%

Discussion and Generative Promise

We believe that the method and procedures described here demonstrate a useful application of Porter's[25, 26] (2002, 2004) curricula analysis framework to analyze a STEM related curricula. Application of this framework to analyze curricula in any discipline creates a win-win situation for educators and those who develop curricula materials. Educators and administrators win because they have access to data with which to evaluate claims made by curricula developers regarding the rigor and relevance of tasks included in the curriculum. Given the pressures on today's educational institution, a focus on analyzing rigor and relevance alongside the integration of mathematics and science concepts seems particularly useful for educators who are interested in knowing more about the quality of their STEM curricula. Through approaches like this one, those in charge of selecting curricula can make informed choices regarding the curricula they plan to implement.

In addition, those involved in curricula development also benefit from analyses similar to this one. First, they are able to apply this approach to generate data to validate the claims about their products and to make improvements to the courses they develop. Additionally, applications of this approach could lead to a consistent language and approach for evaluating the rigor and relevance of STEM related curricula. Ultimately, further applications of the approach outlined here could improve accountability on the part of all commercially available STEM curricula.

As an example of future applications, Project Lead The Way® is currently investing resources to conduct this same analysis on two more courses in the Project Lead The Way® foundation course sequence. Once analysis is complete on these two courses, data will be available to compare and contrast all three foundation courses using a consistent set of criteria.

Limitations

As with any study, there are inherent limitations to discuss. First, by no means should one assume that the method demonstrated, or its results, constitute a comprehensive "report card" by which a stakeholder may evaluate all the facets of a curriculum. Additional research using Porter's[25, 26] (2002, 2004) framework should be conducted to also evaluate the curriculum as it is enacted, assessed, and ultimately learned by students. The approach discussed in this paper is only a starting place and only addresses one facet of a curriculum's overall effectiveness. Moreover, this is only one attempt to apply the framework. Additional studies similar to this one will be able to test and improve the execution of the methods outlined here.

Another inherent limitation in any research is the potential for bias on the part of the researcher. Measures were taken to reduce bias throughout the study design. First, the coding scheme was validated with a panel of subject matter experts before it was applied. Next, the researcher did check the reliability of the coding. Acceptable levels of inter-rater reliability using standards appropriate for studies like this were met. Last, communication between the third-party evaluator and representatives of the organization that commissioned the evaluation were limited during the completion of the study. This was a precaution to ensure the researcher's viewpoint was not adversely influenced.

Conclusion

It is apparent that the United States is lagging behind in efforts to educate enough young people to fill the ever-increasing pipeline of necessary science, technology, engineering, and mathematics (STEM) talent that is demanded by business and industry in order to sustain economic growth. Research continues to point to the role high-rigor, high-relevance curricula will play in addressing this issue (e.g. Daggett[7], 2005; Henningson & Stein[9], 1997; Minner, Levy, & Century[18], 2007). One-step that may be taken to curtail this issue is to ensure shared accountability for the design and development of high-rigor, high-relevance STEM curricula. This study was designed to address the need for a clear and consistent way to describe the rigor and relevance of a particular curriculum and to highlight the importance of providing a method that will enable different educational stakeholders a means to make decisions regarding instructional materials selected to meet the needs of students of STEM education. Currently, many curriculum development processes employ strategies to develop instructional and curriculum guides that define rigorous course objectives, outline the content to be covered, and provide opportunities for students to apply their newly acquired knowledge in relevant, real-world settings. It is our hope that this approach will be used to determine the level of rigor and relevance and ties to subjects, such as mathematics and science, in other STEM curricula. We believe this form of accountability for producing high-rigor, high-relevance curricula is an important component to ongoing efforts to build a high quality and globally competitive STEM education system.

Appendix A

Detailed Depth of Knowledge Descriptions

Level 1 – Recall & Reproduction of Information or Procedures

Curricular elements that fall into this category involve basic tasks that require students to recall or reproduce knowledge and/or skills. The subject matter content at this particular level usually involves working with facts, terms, and/or properties of objects to accomplish a classroom based learning task (e.g. completing a series of problems on a worksheet). It may also involve use of simple procedures and/or formulas. There is little transformation or extended processing of the target knowledge required by the tasks that fall into this category. Key words that often denote this particular level include list, identify, and define. Example tasks at this particular level include:

- Basic calculation tasks involving only one step (i.e. addition, subtraction, etc)
- Tasks that engage students in locating or retrieving information in verbatim form
- Straight-forward recognition tasks related to identifying features, objects and/or steps that don't vary greatly in form (i.e. recognizing features of basic tools)
- Writing tasks that involve applying a standard set of conventions and or criteria that should eventually be automated (i.e. using punctuation, spelling, etc)
- Basic measurement tasks that involve one step (i.e. using a ruler to measure length)
- Application of a simple formula where at least one of the unknowns are provided
- Locating information in maps, charts, tables, graphs, and drawings

Level 2 – Working with Skills & Concepts

Curricular elements that fall in this category involve working with or applying skills and/or concepts to tasks related to the field of engineering in a predictable laboratory setting. The subject matter content at this particular level usually involves working with a set of principles, categories, heuristics, and protocols. At this level, students are asked to transform/process target knowledge before responding. Example mental processes that often denote this particular level include summarize, estimate, organize, classify, and infer. Some tasks that may fit at this particular level include:

- Routine application tasks (i.e. applying a simple set of rules or protocols to a laboratory situation the same way each time)
- Explaining the meaning of a concept and/or explaining how to perform a particular task
- Stating relationships among a number of concepts and or principles
- More complex recognition tasks that involve recognizing concepts and processes that may vary in how they "appear"
- More complex calculation tasks (i.e. multi-step calculations such as standard deviation)
- Research projects and writing activities that involve locating, collecting, organizing and displaying information (i.e. writing a report with the purpose to inform)
- Measurement tasks that occur over a period of time and involve aggregating/organizing the data collected in to basic presentation forms such as a simple table or graph

Level 3 – Short-term Strategic Thinking

Curricular elements in this category demand a short-term use of higher order thinking processes, such as analysis and evaluation, to solve real-world problems with predictable outcomes. Stating one's reasoning is a key marker of tasks that fall into this particular category. The expectation established for tasks at this level tends to require coordination of knowledge and skill from multiple subject-matter areas to carry out processes and reach a solution in a project-based setting. Key processes that often denote this particular level include analyze, explain and support with evidence, generalize, and create. Some curricular and assessment tasks that require strategic thinking include:

- Short-term tasks and projects placing a strong emphasis on transferring knowledge to solve predictable problems
- Explaining and/or working with abstract terms and concepts
- Recognition tasks when the environment observed is real-world and often contains extraneous information which must be sorted through
- Complex calculation problems presented that draw upon multiple processes
- Writing and or explaining tasks that require altering a message to "fit" an audience
- Creating graphs, tables and charts where students must reason through and organize the information with instructor prompts
- Identifying a research question and/or designing investigations to answer a question
- Tasks that involve proposing solutions or making predictions

Level 4 – Extended Strategic Thinking

Curricular elements assigned to this level demand extended use of higher order thinking processes such as synthesis, reflection, assessment, and adjustment of plans over time. Students are engaged in conducting investigations to solve real-world problems with unpredictable outcomes. Employing and sustaining strategic thinking processes over a longer period to solve the problem is a key feature of curricular objectives that are assigned to this level. Key strategic thinking processes that denote this particular level include synthesize, reflect, conduct, and manage. Example tasks include:

- Applying information to solve ill-defined problems in novel situations
- Tasks that require a number of cognitive and psychomotor skills in order to complete
- Writing and/or research tasks that involve formulating and testing hypotheses over time
- Tasks that require students to make multiple strategic and procedural decisions as they are presented with new information throughout the course of the event
- Tasks that require perspective taking and collaboration with a group of individuals
- Creating graphs, tables and charts where students must reason through and organize the information without instructor prompts
- Writing tasks that have a strong emphasis on persuasion

Appendix B

Sample IED Course Objectives and Depth of Knowledge Level Assignments

DOK Level	Objective	Location
Level 1 *Recall/ Reproduction of Information or Procedures*	Identify common geometric shapes and forms by name.	Unit 2 Lesson 2.1 Performance Objective
	Students will list the elements of design.	Unit 3 Lesson 3.1 Assessment Objective
Level 2 *Working with Skills and Concepts*	Apply engineering notebook standards and protocols when documenting their work during the school year.	Unit 1 Lesson 1.1 Performance Objective
	Students will explain the difference between one-point, two-point, and three-point perspectives.	Unit 1 Lesson 1.2 Assessment Objective
Level 3 *Short-Term Strategic Thinking*	Apply geometric numeric and parametric constraints to form CAD modeled parts.	Unit 2 Lesson 2.4 Performance Objective
	Students develop a black box model to identify the inputs and outputs associated with a system.	Unit 3 Lesson 3.2 Assessment Objective
Level 4 *Extended Strategic Thinking*	Research and construct a product impact-timeline presentation of a product from the brainstorming list and present how the product may be recycled and used to make other products after its lifecycle is complete.	Unit 4 Lesson 4.1 Performance Objective
	Students will apply the design process to solve a design problem within a virtual team.	Unit 4 Lesson 4.2 Assessment Objective

Appendix C

Sample IED Objectives Emphasizing Mathematics and Science by Depth of Knowledge Level

DOK Level	Objective	Mathematics and/or Science Standard Link
Level 1 *Recall/ Reproduction of Information or Procedures*	Identify common geometric shapes and forms by name.	*PSSM* Geometry Standard
	Students will list the elements of design.	• *PSSM Connections Standard* • *PSSM Geometry Standard*
Level 2 *Working with Skills and Concepts*	Apply engineering notebook standards and protocols when documenting their work during the school year.	• *PSSM Communication Standard* • *PSSM Representation Standard* • *NSES Content Standard A: Science As Inquiry*
	Explain the concept of fluid power, and the difference between hydraulic and pneumatic power systems	• *NSES Content Standard B: Physical Science*
Level 3 *Short-Term Strategic Thinking*	Apply geometric numeric and parametric constraints to form CAD modeled parts.	• *PSSM Geometry Standard* • *PSSM Algebra Standard* • *NSES Content Standard E: Science and Technology:*
	Students develop a black box model to identify the inputs and outputs associated with a system.	• *NSES K-12 Unifying Concepts: Form and Function* • *NSES Content Standard E: Science and Technology*
Level 4 *Extended Strategic Thinking*	Research and construct a product impact-timeline presentation of a product from the brainstorming list and present how the product may be recycled and used to make other products after its lifecycle is complete.	• *NSES* Content Standard E: Science and Technology
	Students will apply the design process to solve a design problem within a virtual team.	• *PSSM Problem Solving Standard* • *NSES Content Standard E: Science and Technology*

References

[1]American Association for the Advancement of Science (AAAS). (1993). *Project 2061: Benchmarks for science literacy*. NY: Oxford University Press.

[2]ABET, Inc. (2007). *Criteria for accrediting engineering programs*. Retrieved December 15, 2007 from http://www.abet.org/.

[3]American Educational Research Association (AERA). (2006). "Do the Math: Cognitive Demand Makes a Difference." *Research Points: Essential Information for Education Policy,* 4(2).

[4]American Educational Research Association (AERA). (2007). "Science Education That Makes Sense." *Research Points: Essential Information for Education Policy,* 5(1).

[5]Brown, J.S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning, *Educational Researcher, 18.* 32-42.

[6]Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20.* 37–46.

[7]Daggett, W. (2005). Achieving academic excellence through rigor and relevance. *International Center for Leadership in Education.* Retrieved September 25, 2008 from http://www.leadered.com/white_papers.html

[8]Eisner, E.W. (1991). *The enlightened eye.* New York: Macmillan. 227-246.

[9]Henningson, M., & Stein, M.K. (1997). "Mathematical Tasks and Student Cognition: Classroom-Based Factors That Support and Inhibit High-Level Mathematical Thinking and Reasoning," *Journal for Research in Mathematics Education,* 28, 524-549.

[10]Hiebert, J. (2003). *Teaching Mathematics in Seven Countries: Results From the TIMSS 1999 Video Study* (NCES 2003-013). Washington, DC: National Center for Education Statistics.

[11]International Technology Education Association (ITEA). (2000). *Standards for technological literacy: Content for the study of technology*. Reston, VA: Author.

[12]Kimbell, R. (1997). *Assessing technology: International trends in curriculum and assessment*. Buckingham, UK: Open University Press.

[13]Kimbell, R., Stables, K., Wheller, T., Wosniak, A., & Kelly, V. (1991). *The assessment of performance in design and technology*. London, UK: School Examinations and Assessment Council.

[14]Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology.* (2nd Edition). Thousand Oaks, CA: Sage.

[15]Landis, J. R., Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33:159-174.

[16]Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge, UK: University Press.

[17]Linn, M.C., Lee, H.S., Tinker, R., Husic, F., & Chiu, J.L. (2006). "Teaching and Assessing Knowledge Integration in Science." *Science, 313.* 1049-1050.

[18]Minner, D., Levy, A., & Century, J. (2007). *Inquiry Synthesis Project Executive Summary.* Newton, MA: Education Development Center.

[19]National Commission on Excellence in Education. (April, 1983). *An open letter to the American people: A nation at risk – The imperative for educational reform*. Retrieved October, 2008 from http://www.mat.uc.pt/~emsa/PMEnsino/ANationatRisk.pdf

[20]National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
[21]National Research Council (NRC). (1996). *National science education standards*. Washington, D.C.: National Academy Press.
[22]National Science Foundation. (2004). *Science and Engineering Indicators.* Washington, DC: Author.
[23]Newberry, P., & Hughes, E. (2006). *Activities-, project-, problem-based learning: A modality of teaching and learning*. Retrieved October 2, 2008 from http://www.pltw.org
[24]Patton, M.Q. (2002). *Qualitative research & Evaluation methods.* (3rd Edition). Thousand Oaks, CA: Sage Publications.
[25]Porter, A. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher, 31.* 3-14.
[26]Porter, A. (2004). Curriculum assessment. *Complementary Methods for Research in Education*. J. Green, G. Camilli, & P. Elmore (Eds.) Washington DC: AERA. pp. 141-159.
[27]Project Lead The Way, Inc. (PLTW). (2007). *Course descriptions*. Retrieved October 2, 2008 from http://www.pltw.org
[28]Roth, K.J. (2006). *Teaching Science in Fivve Countries: Results from the TIMSS 1999 Video Study* NCES 2006-011). Washington, D.C.: National Center for Education Statistics.
[29]Schauble, L., Klopfer, L.E., & Raghavan, K. (1991). Students' transition from an engineering model to a science model of experimentation. *Journal of Research in Science Teacher*, 28, 859-882.
[30]Siegel, S. & Castellan, N.J. (1988). *Nonparametric statistics for the behavioral sciences, 2nd ed.* New York: McGraw Hill.
[31]Stein, M.K., Grover, B.W., & Henningsen, M. (1996). "Building Student Capacity for Mathematical Thinking and Reasoning: An Analysis of Mathematical Tasks Used in Reform Classrooms." *American Educational Research Journal, 33.* 455-488.
[32]Surveys of the Enacted Curriculum, (2009). *SEC Reports*. Retrieved January 23, 2009, from http://seconline.wceruw.org/
[33]Webb, N. L. (1997). *Criteria for alignment of expectations and assessments in mathematics and science education*. Council of Chief State School Officers and National Institute for Science Education Research Monograph No. 6. Madison: University of Wisconsin, Wisconsin Center for Education Research.
[34]Webb, N.L. (2002). *Depth of Knowledge levels for four content areas.* Retrieved November 30, 2007, from http://facstaff.wcer.wisc.edu/normw/All%20content%20areas%20%20DEPTH OF KNOWLEDGE%20levels %2032802.doc
[35]Wiggins, G.P. & McTighe, J. (2005). *Understanding by Design.* New Jersey: Prentice Hall.
[36]U. S. Department of Education. (n.d.). *No child left behind* (*NCLB*). Retrieved October, 2008 from http://www.ed.gov/nclb/landing.jhtml