Results from the Third International Mathematics and Science Study (TIMSS) show American students lagging behind their international counterparts in the eighth grade mathematics (IEA, 1997). More disturbing, 4th grade mathematics achievement is above the international average but by 8th grade, achievement drops below the average. The latest findings of the National Assessment of Educational Progress (NAEP) indicate that while achievement in mathematics has continued to increase slowly during the past decade, fewer than 20 percent of students achieve above basic levels of performance (National Center for Education Statistics, 1997).

Textbooks are an important component of mathematics education and have begun to receive increased attention as a factor in low mathematics achievement. The release of the TIMSS report has raised the awareness of the impact of textbooks on mathematics achievement. The TIMSS analysis of U. S. mathematics textbooks, which consisted of a page by page count of the coverage of mathematics topics, revealed that textbooks are repetitive and unfocused, attempting to cover far more topics each year than in other countries (Schmidt, et. al., 1997). The TIMSS approach to looking at textbooks provides important data about coverage of topics but stops short of determining how well materials are aligned with specific standards or benchmarks. It also does not address the critical issue of what instructional strategies textbooks employ. Although it would help, especially in the middle grades, to reduce the number of topics in textbooks, little progress can be
made toward achieving standards-based curricula unless textbooks are aligned with the mathematics learning goals set forth in the standards agreed upon by national professional organizations and by states and school districts.

Since so many teacher rely heavily on textbooks for instructional guidance, it is also important an analysis includes an examination of the instructional characteristics of curriculum materials. Although there is no single strategy for effective mathematics instruction, there are principles that have been shown through research and practice to lead to student learning. These principles can be used to develop criteria for judging textbooks.

**Textbook Selection**

The marketing and selection of textbooks is a multi-billion dollar enterprise in the United States that sometimes has very little to do with producing the best instructional materials (Tyson, 1986). The situation has not improved with the development and wide acceptance of mathematics standards. According to one critic, “Despite the development of national and state standards, and standards-driven curricula and tests, the problem of superficial textbooks is still with us and appears to be getting worse…” (Tyson, 1997).

Teachers have several types of opportunities to influence textbook selection, including membership on state level adoption committees, school or district selection groups, or making unilateral decisions for their own department or classroom. The influence that teachers can have on the textbook problem is severely limited by the amount of time and training they receive to do the job. More importantly, the prospects of changing the system are limited by the procedures that are used in the selection process. A textbook committee is typically provided a stack of textbooks along with a checklist of criteria for “rating” each book. The time to read and examine each book is often less than an hour, with little or no discussion or training in the use of the criteria. The value of the process is limited if
evaluators have different understandings of the meaning of criteria such as “appropriate breadth and depth of content” or “clear illustrations and examples.” The validity of these kinds of criteria is even more limiting. How do we know that “clear illustrations and examples” lead to learning? If teachers and other educators are to have a voice in the improving quality of mathematics textbooks, they need a selection procedure that offers an opportunity to reflect specific, evidence-based descriptions of the strengths and weaknesses of materials. Because textbook development is a business that responds to the market, schools and teachers must learn to be more critical consumers.

Curriculum Materials Analysis

An important step in selecting high quality mathematics textbooks and other instructional materials is to learn how to evaluate the materials. Learning how to analyze and evaluate materials according to standards and benchmarks results not only in identifying better textbooks, but also engages teachers in discussions that can lead to a better understanding of mathematics content and instruction.

Project 2061 has developed a curriculum materials analysis procedure that reliably identifies science and mathematics materials that are aligned with specific content standards and reflect effective instructional strategies (Roseman, Kesidou, & Stern, 1996; Kulm, 1997b). The analysis procedure is divided into four phases: 1) Preliminary Phase, 2) Content Analysis Phase, 3) Instructional Analysis Phase, and 4) Summary/Report Phase. During the Preliminary Phase a reviewer makes an initial pass through a set of curriculum materials (a unit, single grade text, or multi-year series) in order to determine the likely standards or benchmarks that are addressed by the materials. For textbook selection, this preliminary phase can be used to eliminate from consideration the materials that have major gaps or mismatches with the set of standards that will be used as the criteria.
In the Content Analysis Phase, the most efficient approach for the purpose of textbook selection is to select a small representative set of standards or benchmarks that serve as learning goals for the analysis. In mathematics, standards from important content strands can be sampled: (e.g., Number and Computation, Algebra, Geometry and Measurement, and Probability and Statistics). Once these standards are selected, the reviewer examines a textbook to determine how well it addresses the substance and sophistication of each standard. In order to understand substance of the learning goal (rather than just the topic), a clarification procedure is used. Project 2061 recommends using various resources to help with this clarification process (Roseman, Kesidou, and Stern, 1996). In mathematics, the NCTM Standards (NCTM, 1989) and AAAS Benchmarks (AAAS, 1993) provide information about appropriate content and learning goals. The comparison of these two sets of national standards is a useful tool for clarifying the intent of standards statements (Kulm, 1997a). At the state or district level, the mathematics framework document should describe the philosophy behind its creation and the intent of the learning goals to develop students’ concepts and skills in each strand. Also useful are research findings on how children learn mathematics topics, concepts, and procedures at different age levels. Research summaries such as the one edited by Owens (1993) and Chapter 15 in Benchmarks (AAAS, 1993) can provide information to further clarify concepts and skills as well as place the learning goal at the appropriate sophistication level.

Once a standard is clarified, the reviewer is prepared to work through the material, making sightings of specific explanations, activities, exercises, assessments, and other content that addresses the standard. With a clear idea of how the standard represents a mathematics learning goal, the reviewer can judge whether each potential sighting
addresses the central concepts and skills contained in the standard. A set of materials or
textbook may adequately address a part or whole of any given standard or learning goal.
The content analysis is completed by examining the textbook and teacher’s guide page by
page, noting sightings of explanations, activities, examples, exercises, assessments, and
other material that address each identified standard or learning goal.

The Instructional Analysis Phase is best described as a procedure to determine the
approach that a textbook uses to present the mathematics contained in each standard. The
content analysis sightings are used to in this phase in which the reviewer revisits the
sightings made during the Content Analysis Phase. The reviewer analyzes the material
with respect to a list of instructional criteria divided into clusters including Sense of
Purpose, Student Ideas, Engaging Students, Mathematical Concepts and Skills, Student
Reflection, Assessment, and Support Features. A material may do quite well implementing
some instructional criteria and not as well in addressing others.

Finally, during the Summary and Report Phase, the reviewer describes the findings
and makes a judgment of the material’s treatment of the identified standards and learning
goals. The report would include tables of the sightings with the reviewers’ notes which
provide evidence for the judgments that were made. When considering several sets of
materials for textbook selection, the Summary Reports would include ratings for
comparison or ranking.

The Kentucky Teacher Network

A case study of the adaptation of curriculum analysis for textbook adoption is provided
by Project 2061’s work with the Kentucky Middle Grades Mathematics Teacher’s Network
(KMGMTN), a long-term project funded by the National Science Foundation (NSF). The
KMGMTN is designed to connect teachers and universities across the state in staff
development efforts aimed at improving mathematics instruction. After three years of work, KMGMTN teachers had an opportunity to review pre-publication versions of five new middle grades mathematics materials that had been developed with the assistance of NSF. The mathematics curriculum development projects funded by NSF have produced 13 new products, from elementary through high school levels, intended to reflect the NCTM Standards. The materials have not yet had much impact in schools, partly because there has not been time for them to be adopted in states and districts and partly because a great deal of staff development is necessary for their use.

Although the applications-oriented and teacher- and student-friendly design of the NSF materials were appealing, the Kentucky teachers were unsure how well the materials addressed learning goals described by NCTM Curriculum and Evaluation Standards for School Mathematics (1989) and Kentucky’s Core Content for Mathematics Assessment (n.d.). In these materials, the familiar mathematics topics, procedures, and exercises were embedded in units that were organized thematically rather than topically. Teachers’ prior experience in searching tables of contents and flipping through pages for favorite topics or lessons were not helpful approaches in reviewing the materials.

These Kentucky teachers had worked hard during the three year project to improve their teaching and to learn about standards-based ideas. Some of them had helped to write the Kentucky standards and develop the state assessment guidelines. This effort had paid off in improved scores by middle school students on the Kentucky state mathematics assessment. It was important to know whether the new NSF materials did live up to the Kentucky standards, and whether these teachers could make such a judgment of the materials.

Further, since statewide mathematics textbook adoption was to take place in 1997-98, it was important that a procedure be put in place that gave both the NSF materials and newly
developed commercial textbooks that claimed to address NCTM Standards a fair chance at being judged on Kentucky’s Core Content for Mathematics Assessment.

Adapting the Curriculum Analysis Procedure

The Project 2061 curriculum analysis procedure was originally developed and tested on middle grades science materials, using Benchmarks for Science Literacy (AAAS, 1993) as the criteria for content alignment. In preparing for the Kentucky training, an advisory group consisting of Kentucky mathematics educators, state department leaders, and developers of the NSF materials assisted Project 2061 in adapting the procedure (Kulm, 1997). For the Kentucky textbook selection training, the curriculum materials analysis was adapted in three ways: (1) the procedure itself was adjusted to reflect research and practice in effective content and instruction in mathematics, (2) the criteria for content alignment was the Kentucky Core Content for Mathematics Assessment, and (3) the standards statements used in the analysis were sampled across content strands.

The adaptation of in the procedure for use in mathematics did not result in substantive changes in the analysis. Terminology was changed somewhat to reflect the importance and distinction of conceptual and skill development in mathematics. Naturally, the examples and discussion of content standards were changed, using NCTM Standards statements as exemplars.

In using the Core Content statements as the analysis criteria, the main issue was the “grain size” of the statements. The procedure requires specific statements of what students are expected to know and be able to do, and works best when the content criteria are specific and concrete, as in the AAAS Benchmarks. Mathematics standards statements, including those in the Kentucky Core Content tend to be more general. This characteristic of mathematics standards opened the possibility that the procedure would fail to
discriminate between materials that appear to address the standards on a surface level and those that actually align with the standards. In applying the procedure, the Core Content statements were compared with relevant AAAS Benchmarks and NCTM Standards, and clarified by studying the essays in the national documents. This clarification made it possible for the Core Content statement to be used as effective criteria for doing the analysis.

Finally, adapting the procedure for the purpose of textbook selection required decisions about using sampling and other techniques that could make the analysis efficient while providing sufficient depth and breadth in the review. Kentucky’s standards are organized into four strands (Numeration, Geometry and Measurement, Algebra, and Probability and Statistics) and three cognitive processes (skills, concepts, relationships). Core Content statements were sampled to represent each strand and each type of cognitive process. To assure in-depth review, each sampled standard statement was carried through an entire set of materials, from grades 5 through 8. This approach provided an opportunity to judge both NSF materials and commercial textbooks fairly, even though they differed considerably on the, grade level, distribution and type of learning activities, and depth at which the sampled standards were addressed.

Curriculum Analysis Training

A two-week summer session was developed and conducted to train a selected group of middle grades teachers in the analysis procedure and prepare them to present textbook selection workshops in their regions. Thirty mathematics educators (25 middle grades teachers, 2 resource teachers, 3 university faculty) participated in the two-week training on the campus of the University of Kentucky.
The focus of the work during the first week of the workshop was on learning the analysis procedure by applying it to four NSF-funded middle grades mathematics materials and four recently published commercial textbooks. As an outcome of the training, participants produced documentation and evidence of each material’s alignment with their selected standard.

The activities used to teach the procedure were as follows:

- **Overview.** Presentation of a brief overview of the full analysis procedure,
- **Core Content clarification.** The participants were each assigned two Kentucky Core Content statements to clarify for use as criteria for content and instruction analysis of each material. This approach enabled participants to become “experts” on specific mathematics standards. In a jigsaw activity, participants examined “their” Core Content standard, using comparisons with NCTM Standards, AAAS Benchmarks, Kentucky Transformations, and summaries of research findings on student mathematics learning.
- **Content analysis.** Using their selected standards, participants located specific sightings of activities that addressed the standards in two separate curriculum materials (one NSF-funded, one commercial). This strategy was an effective way to emphasize to the participants the importance of using specific standards statements as analysis criteria. It also helped them to see contrasts in the way instructional materials address mathematics content. The sightings were especially revealing for the NSF materials in which concepts were defined and practice was provided within investigations and activities rather than as separate sections.
- **Instructional analysis.** In a carousel activity, participants located examples of activities that addressed the selected Core Content standards with respect to each of
the seven instructional clusters, across all eight materials (4 NSF-funded, 4 commercial). This approach enabled the participants to learn how each of the eight sets of curriculum materials differed in the instructional approach used to address specific standards.

The second week was devoted to adapting the procedure for workshops that the participants could use in their own districts. Participants made decisions about which activities to include and which to modify or eliminate in developing 1-hour and 2-day workshops. Working in teams, participants produced a notebook that contained handouts, transparency masters, and detailed workshop outlines for these workshops. Each group had an opportunity to do a test run of workshop activities, receive feedback, and discuss advantages, disadvantages and issues raised. During the 1997-98 school year, the participants presented these workshops in their schools, districts, and regions as part of the textbook selection process. In addition, some of the participants adapted the workshops for three 90-minute presentations on the Kentucky Educational Television network.

Changes in Ideas About Selecting Materials

On the first day, a questionnaire was given to determine the priorities and criteria that participants used to form judgments about curriculum materials, and to assess their own estimates of ability to judge materials. At the end of the second week, the participants completed the same questionnaire to determine how their judgments had changed. Of the 30 participants, 29 completed both the pre- and post-questionnaire.

In interpreting the results, it is important to note that the participants were a select group who had a great deal of prior experience with standards and the notion of alignment of standards with curriculum and assessment. Nevertheless, the questionnaire yielded informative data about the effects of the training and indications of changed perceptions.
Participants were asked in an open-response format what criteria they would use to select middle grades mathematics materials. Before the training, 23 of the 29 participants noted alignment with standards (NCTM and/or Kentucky standards) as important. After the training, only 14 of the 29 cited alignment. While this result seems contradictory in light of the emphasis on content alignment, there was a pre-post increase from 2 to 12 participants who mentioned the importance of depth of content or concepts. This result reflects a fine-tuning of understanding what it means “to align with standards.” The criterion “concept development” increased from 1 to 7 responses, reflecting the instructional analysis training. In contrast, before training, surface-level “teacher-friendly” criteria were mentioned by 6 of 29 participants, but only by one person after training.

Participants were asked how much time they thought is needed to review materials for adoption. Before training, answers ranged from “½-1 hour” to “months” and 18 respondents said “at least three days” would be needed. After training, 21 people said “at least 3 days” and 7 people increased their judgments from ‘hours’ to ‘at least 3 days.’ It was not surprising that the time estimates increased but the size of the increase reflected the impact of the training on the importance of in-depth analysis.

The final section of the questionnaire asked participants to rate their ability to make judgments about instructional strategies and design issues. Five of the questions were drawn directly from the instructional analysis criteria. The others were about technology, format features, and teacher support. Participants were asked to rate their ability to make judgments on each criterion on a 5-point Likert scale (5 - Very capable, 1 - Less capable). Table 1 presents the mean ratings for each of the five questions related to Instructional Analysis clusters, along with sample statements participants gave for how they make judgments, before and after training.
Because these were very capable teachers who had overall confidence in their abilities, the pre-training scores were in the 3 to 4 range. Even so, improvement was made not only in the quantitative ratings but also in the way participants said they would make judgments after training. Many of them referred directly to the analysis procedure and the idea of using sightings in the material as evidence for judgments.

Discussion and Conclusions

Overall, the training appears to have been successful. Several participants expressed considerable respect for the level of professional development achieved regarding standards alignment with curriculum materials and implications for adoption. Many participants commented that it will be impossible for them in the future to merely flip through a set of curriculum materials without considering how well it addresses specific learning goals (in terms of substance, sophistication and instructional practices). The experience gained during the workshop raised additional issues related to adapting the procedure to mathematics materials using Kentucky’s learning goals and in using the analysis for textbook adoption. The two most common concerns were the amount of time required to do an adequate job and the terminology used in describing mathematics content and instruction.

Selecting a small sample of standards for the analysis appears to be a valid way to significantly reduce time without damaging the integrity of the procedure. It allows local districts or schools to identify the standards and learning goals that they feel are most important for textbooks to include and develop. The instructional analysis is
comprehensive enough to obtain a good picture of how a textbook addresses learning, through the use of a few sampled standards. In the analyses, it was clear that textbooks had consistent instructional approaches, so that analyzing a relatively small number of standards was sufficient.

For teachers who often consider materials for adoption in addition to their regular workload, a method of streamlining the procedure was needed. The problem was to maintain attention to aligning content with specific learning goals expressed in the standards, as well as addressing the instructional clusters, while reducing the time required for detailed analysis. The participants developed a checklist requiring reviewers to consider the quality of sightings in terms of the substance of learning goals in each strand and in terms of instructional criteria. Reviewers would still perform the complete analysis, but would have a method for summarizing and rating the materials more efficiently.

Some of the analysis procedure terminology, developed with science materials in mind caused confusion for math teachers. For example, one of the cluster criteria asks reviewers to consider how the material introduces “technical” vocabulary. A suggested revision led to a statement about “how the material introduces mathematical terms and algorithms,” as the focus for judging whether materials used terminology appropriately and developed justification and understanding of algorithms.

Another example of the adaptation for mathematics involved the idea of students having hands-on experience with a concept or skill. This criteria arose from the concern in science that students need to engage in inquiry with actual scientific phenomena, rather than simply memorizing vocabulary and scientific methods. In mathematics, direct experience with ideas can include using real-world data or materials, working with manipulatives, or solving meaningful problems. “Minds-on” might be a better term,
implying that students would experience a variety of meaningful activities, not restricted to
natural phenomena.

The model used in this project of providing in-depth training for master teachers, along
with guidance for adapting the procedure for participants’ own use is promising. Following
a week’s training the participants understood the procedure well enough to be able to
develop their own workshops plans. Further, they were able to identify the key components
of the analysis to include and find ways to make the procedure useful for their colleagues.
Although further information from the participants’ own workshops will be important in
determining the effectiveness of the approach, several participants have already made
presentations and received positive responses from their audiences.

There is value in learning the analysis, even if its application to textbook selection is
limited. Teachers make curricular decisions daily as they plan activities and adapt
materials. A deep understanding of goals and the criteria that enhance students’ learning of
them extends beyond textbook selection. In the questionnaire results, there was evidence
that teachers changed their thinking about goal and instructional activities related to them.

Reform can continue and succeed if teachers, administrators, and parents are willing to
consider and choose carefully the materials being used to teach mathematics concepts and
skills. Educators at all levels need to clarify the learning goals students are expected to
achieve and to make connections within and between strands or topics in mathematics and
between mathematics and other disciplines (science or social studies, for example). The
appropriate sophistication level must be used and concepts need to be developed logically
and sequentially. The curriculum analysis procedure is a tool that can be used to develop
more reflective teachers and more critical consumers of mathematics instructional
materials.
References


Kentucky Department of Education. (n.d.). *Core Content for Mathematics Assessment.* Lexington, KY.


### Table 1.

Mean Scores and Sample Participant Comments for Five Instructional Analysis Questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Mean Score</th>
<th>Sample answers: How judgment is made (Self-rating)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ability to judge how a material helps teachers address students’ difficulties in understanding mathematical ideas.</strong></td>
<td></td>
<td>Pre-training</td>
</tr>
<tr>
<td></td>
<td>2.9</td>
<td>Post-training</td>
</tr>
<tr>
<td></td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>…depends on the text. Hard to find until text is used. (2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Look for section/workbook for remediation or various approaches to introducing it. (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Based on instruments we developed [in the workshop]. (5)</td>
</tr>
<tr>
<td><strong>Ability to judge how a material engages students.</strong></td>
<td>3.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Through classroom experiences and involvement in professional development activities. (4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I really think you need to use it to make this judgment. You never know how students will accept things. (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Student engagement portion of analysis. (5)</td>
</tr>
<tr>
<td><strong>Ability to judge how a material builds mathematical ideas on students prior knowledge.</strong></td>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>I have taught 4th grade &amp; have KMGMTN &amp; K-4 math program under my belt. This has helped to see a broader scope. (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>This should be reflected if you look at one strand at a time. (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Searching the materials for the last 2 weeks has greatly increased the process for this area. Using the 3 basic support materials: Standards (Core Content) clarification, content analysis, instructional analysis. (5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Look for substance in sightings - Look through several levels of material. (4)</td>
</tr>
<tr>
<td>Question</td>
<td>Pre-training</td>
<td>Post-training</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------</td>
<td>--------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Ability to judge how a material provides students with opportunities to</td>
<td>3.2</td>
<td>4.0</td>
</tr>
<tr>
<td>reflect on experiences and knowledge.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ability to judge how a material embeds assessment activities within</td>
<td>3.7</td>
<td>4.3</td>
</tr>
<tr>
<td>instruction.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>