Alignment Content Analysis of TIMSS and PISA Mathematics and Science Assessments Using the Surveys of Enacted Curriculum Methodology

February 2009

A CCSSO paper
Prepared for National Center for Education Statistics
and American Institutes for Research

Council of Chief State School Officers
The Council of Chief State School Officers (CCSSO) is a nonpartisan, nationwide, nonprofit organization of public officials who head departments of elementary and secondary education in the states, the District of Columbia, the Department of Defense Education Activity, and five U.S. extra-state jurisdictions. CCSSO provides leadership, advocacy, and technical assistance on major educational issues. The Council seeks member consensus on major educational issues and expresses their views to civic and professional organizations, federal agencies, Congress, and the public.

State Collaborative on Assessment and Student Standards
The State Collaborative on Assessment and Student Standards (SCASS) Project was created in 1991 to encourage and assist states in working collaboratively on assessment design and development for a variety of topics and subject areas. SCASS projects accomplish a wide variety of tasks identified by each of the groups including examining the needs and issues surrounding the area(s) of focus, determining the products and goals of the project, developing assessment materials and professional development materials on assessment, summarizing current research, analyzing best practice, examining technical issues, and/or providing guidance on federal legislation.

Alignment Content Analysis of TIMSS and PISA
Mathematics and Science Assessments
Using the Surveys of Enacted Curriculum Methodology

February 2009

A CCSSO paper
Prepared for National Center for Education Statistics
and American Institutes for Research

Conducted by Surveys of Enacted Curriculum (SEC)
State Collaborative Project
Rolf K. Blank, Director

Research design and analysis
John Smithson
Wisconsin Center for Education Research
University of Wisconsin-Madison

Table of Contents

**Alignment Content Analysis of TIMSS and PISA
Mathematics and Science Assessments
Using the Surveys of Enacted Curriculum Methodology**

In Fall 2008, the Council of Chief State School Officers (CCSSO) conducted an alignment content analysis of the 2007 TIMSS Mathematics and Science education assessments for students at grades 4 and 8 and the 2006 PISA Mathematics and Science Literacy assessments for students at age 15 (i.e., TIMSS--Trends in Mathematics and Science Study, PISA--Program for International Student Assessment). The content analysis was completed using the methodology and content frameworks for the Surveys of Enacted Curriculum (SEC). A goal of the project for CCSSO was to allow states to analyze the alignment of state content standards and assessments in relation to TIMSS and in relation to PISA. The project was supported by the American Institutes for Research through a contract with the National Center for Education Statistics (NCES). The results of the alignment analysis are now being reported. The purpose of this paper is to summarize the analysis procedures, to highlight several analyses in relation to States, and to explain the use of the website for accessing the TIMMS and PISA analysis data.

SEC Content Analysis Method and Procedures

The SEC instruments include a two-dimensional content framework for each subject that was designed to collect, analyze and report data on curriculum that has been taught and to analyze curriculum content in relation to standards (intended curriculum) as well as assessments that determine what has been learned. In Fall 2008, CCSSO arranged for four-person teams of mathematics and science educators to use the SEC method and content frameworks to analyze the TIMSS and PISA mathematics and science assessment items. About 75 percent of the team members had experience with the SEC content analysis method. CCSSO together with our research contractor Wisconsin Center for Education Research (WCER) provided a one-half day training session and the team members conducted the content analysis with the complete set of mathematics and science assessment items over a two-day period in November 2008. (See Appendix for description of the content analysis process and reliability statistics.) The secure items sets were organized and provided to the teams by staff of American Institutes for Research, a contractor of the NCES.

The Surveys include a K-12 subject "content framework." The complete Survey design and the content frameworks for mathematics and science were developed by CCSSO and the Wisconsin Center for Education Research through a collaborative project with state departments of education involving educators, researchers, and subject area specialists (see Blank, Porter, & Smithson, 2001; Porter, 2002). The SEC analysis method has been used to analyze the content of standards and assessments in mathematics, science, English language arts, and social studies in over 30 states. (See www.SEConline.org for description of the SEC alignment analysis methodology and to view the content framework; see www.SECsurvey.org for a listing of states participating and conducting content analyses and further references to SEC research and development).

Accessing Alignment Data through SEConline.org

We are reporting the TIMSS and PISA content analysis results through the internet on the www.SEConline.org webpage to provide access to a broad range of users of the analysis data. The analyses can be viewed through the two reporting formats used in the SEC online system –either a) contour maps, or b) tile charts. The user selects the content results to be viewed and each selected content map or chart can be compared directly with one other selected map or chart (e.g., TIMSS grade 4 math assessment by NAEP math grade 4).

Subject content analysis data are reported in two dimensions—Topics and Expectations for learning (or "cognitive demand"). Each selected chart is first displayed at the Coarse-grain (large) Topic level. A Fine-grain topic chart under each Coarse-grain topic can be viewed by clicking on the name of the topic (shown in green font). (For examples of the Chart formats go to http://www.SECsurvey.org/ "How are Data Reported?")

Using the SEConline.org website, results of all prior content analyses conducted through the SEC methodology and frameworks can be compared with the TIMSS or PISA content analyses.   The paired analyses might include NAEP Frameworks, NAEP assessments, State standards or assessments, College Board standards, and other documents that have been entered into the SEConline.org system. An alignment index is reported for each pair of standards or assessments that are selected and compared, with the index varying from 0 to 1 according to the degree of consistency of match in the content topics and expectations for student learning of the documents being compared.

Content Analysis Results
On the attached charts, the "alignment index" refers to the degree of consistency or match between the content (2 dimensions) for the standards/assessment on the left side with the content of the standards/assessment on the right side. The "coarse grain" statistic refers to the alignment or consistency of the main topics and expectations.   The "re-centered" statistic refers to the alignment of the standards/assessment content at the fine grain level.

TIMSS compared to NAEP Mathematics Assessments
Content alignment can be compared for TIMSS and PISA assessments with US national assessments and state standards and assessments. The TIMSS mathematics assessment for grade 4 is aligned to NAEP mathematics assessment at the coarse grain level (large topic categories) of .75.   The overall alignment index (categories and fine grain topics) is .55. The level of alignment indices for grade 8 NAEP and TIMSS are very similar.

Update Map

0% - 0.99%	1% - 1.99%	2% - 2.99%	3% - 3.99%	4% - 4.99%
5% - 5.99%	6% - 6.99%	7% - 7.99%	8% - 8.99%	9% - 9.99%

Contour Interval - 1% of Content Coverage

Number Sense / Properties / Relationships
Operations
Measurement
Consumer Applications
Basic Algebra
Advanced Algebra
Geometric Concepts
Advanced Geometry
Data Displays
Statistics
Probability
Analysis
Trigonometry
Special Topics
Functions
Instructional Technology
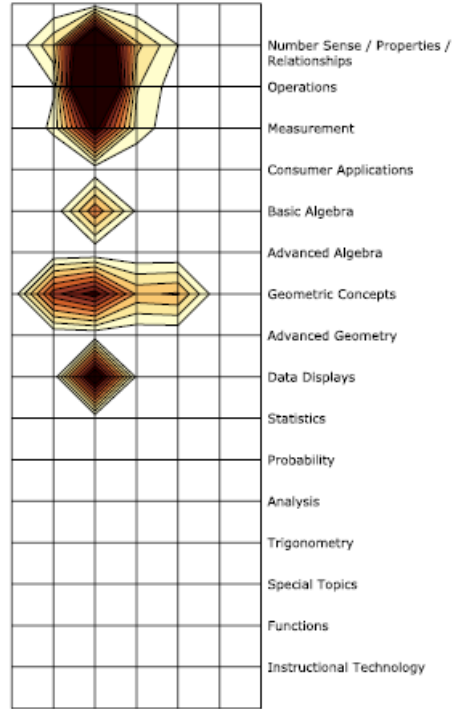
Memorize Facts, Definitions, Formulas
Perform Procedures
Demonstrate Understanding
Conjecture, Analyze, Generalize, Prove
Solve Non-Routine Problems/Make Connections

Number Sense / Properties / Relationships
Operations
Measurement
Consumer Applications
Basic Algebra
Advanced Algebra
Geometric Concepts
Advanced Geometry
Data Displays
Statistics
Probability
Analysis
Trigonometry
Special Topics
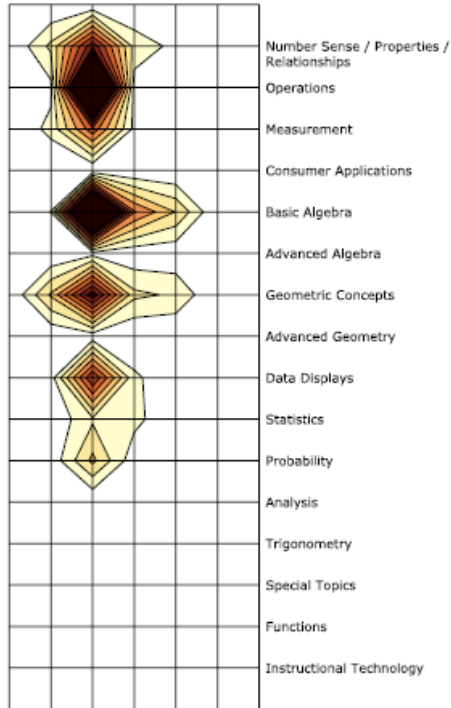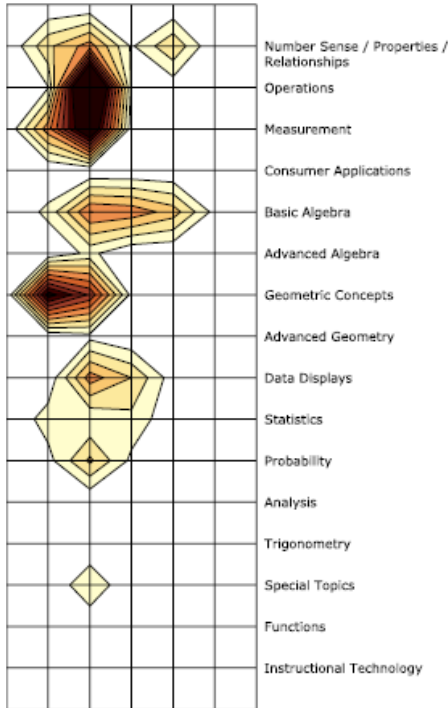Functions
Instructional Technology

Memorize Facts, Definitions, Formulas
Perform Procedures
Demonstrate Understanding
Conjecture, Analyze, Generalize, Prove
Solve Non-Routine Problems/Make Connections

6
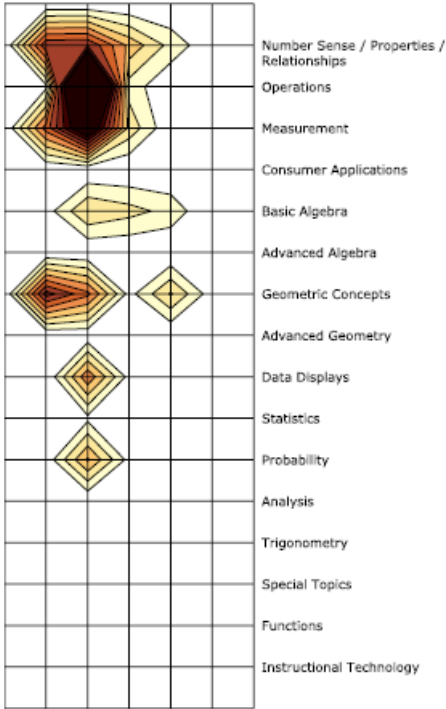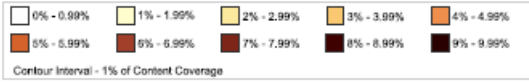
Administration Year: 2008
Viewing: NAEP 2007 Gr. 8 Data
Data Cut: All Data
Count: 1

Administration Year: 2008
Viewing: 2008 TIMMS Gr. 8 Data
Data Cut: All Data
Count: 1

Update Map

0% - 0.99%	1% - 1.99%	2% - 2.99%	3% - 3.99%	4% - 4.99%	
5% - 5.99%	6% - 6.99%	7% - 7.99%	8% - 8.99%	9% - 9.99%	

Contour Interval - 1% of Content Coverage

Number Sense / Properties / Relationships
Operations
Measurement
Consumer Applications
Basic Algebra
Advanced Algebra
Geometric Concepts
Advanced Geometry
Data Displays
Statistics
Probability
Analysis
Trigonometry
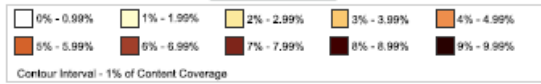Special Topics
Functions
Instructional Technology

Memorize Facts, Definitions, Formulas
Perform Procedures
Demonstrate Understanding
Conjecture, Analyze, Generalize, Prove
Solve Non-Routine Problems/Make Connections

7

As a point of comparison, this level of alignment is high and comparable to the best alignment results seen for state framework/standards to state assessment alignment results. Based upon available mathematics alignment data from state documents (n = 59 comparisons, or 118 documents), the mean overall alignment for state standards/frameworks to state assessments is .26, with the highest level of alignment among state standards and assessments at .50.

We have attached content maps from several states to show the alignment of content in state standards and assessments with TIMSS Math assessment. These examples highlight the kind of analysis that is possible for state level standards and assessments with the TIMSS and PISA content analyses.

State Standards compared to TIMSS Mathematics

The first Mathematics Contour map attached shows the degree of alignment between a State's standards (Ohio) for grade 4 called "Indicators" and the TIMSS grade 4 assessment. The coarse grain alignment index is .63 (coarse grain= main topic categories displayed), and the overall alignment index is .29. This alignment is high. We can see this because the topic categories (displayed in rows) from Number Sense, Operations, Measurement, Basic algebra, Geometric concepts and Data Displays are included in both the TIMSS assessment and the Ohio standards, and the levels of emphasis across the five expectations levels are similar. As can be seen in the chart, the TIMSS assessment differs from the state standards in the expectations dimension for Geometric concepts, in that TIMSS asks students to complete items that call for them to be able to conjecture, analyze or generalize in answering geometry items (4th column from right). Additionally, the Ohio standards include a focus on Probability at the Perform procedures level.

A second chart for Ohio standards by TIMSS at grade 4 displays a fine-grained analysis for the topic category **Geometric concepts.** The data show that the Ohio standards emphasize instruction of a smaller number of topics in Geometry at grade 4 while the TIMSS assessment calls for students to respond to content under 16 topics and the expectations cover levels 1-4 for most topics. Thus, while the topic looks somewhat similar in emphasis at the coarse grain level, it is apparent that the TIMSS assessment is more content demanding than the Ohio standards. The "re-centered" (fine grain) alignment statistic for this particular topic is .26.
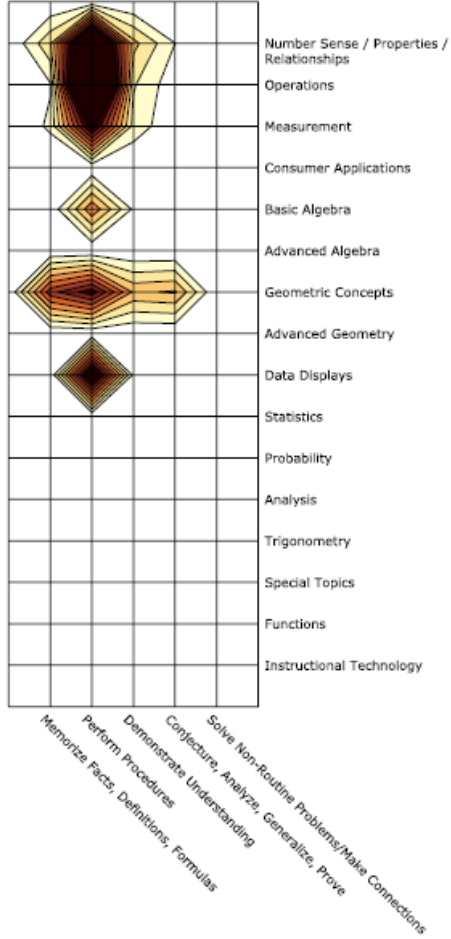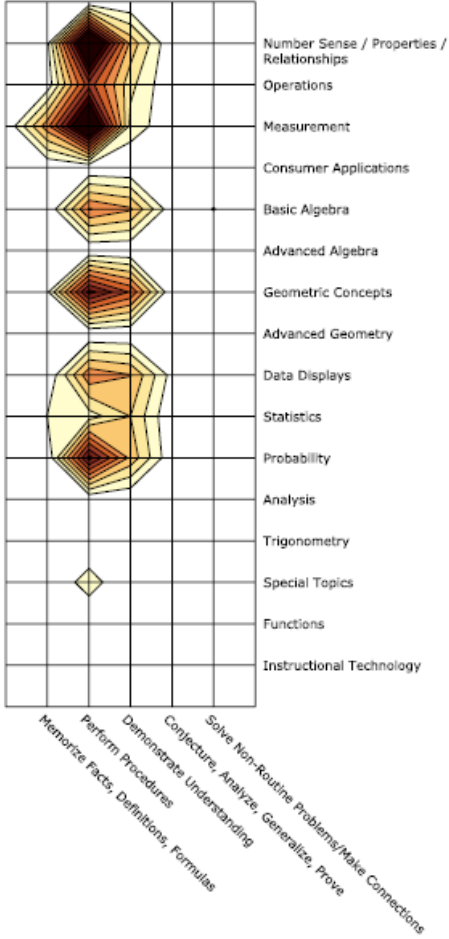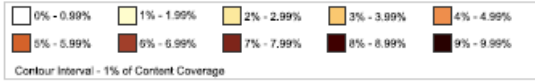
Data Cut: All Data

Data Cut: All Data

Count: 1

Count: 1

Update Maps

0% - 0.99%	1% - 1.99%	2% - 2.99%	3% - 3.99%	4% - 4.99%	
5% - 5.99%	6% - 6.99%	7% - 7.99%	8% - 8.99%	9% - 9.99%	

Contour Interval - 1% of Content Coverage

Number Sense / Properties / Relationships

Operations

Measurement

Consumer Applications

Basic Algebra

Advanced Algebra

Geometric Concepts

Advanced Geometry

Data Displays

Statistics

Probability

Analysis

Trigonometry

Special Topics

Functions

Instructional Technology

Memorize Facts, Definitions, Formulas

Perform Procedures

Demonstrate Understanding

Conjecture, Analyze, Generalize, Prove

Solve Non-Routine Problems/Make Connections

Alignment Overall: 0.288
Alignment Re-centered: 0.2629

= Not Covered
= < 0.5%
= < 1.0%
= < 1.5%
= >= 1.5%

Show Data Tables

Administration Year:

Sample Selection: 2008 TIMSS Gr. 4 | OH Indctrs Gr. 4

Report By: All Data | All Data

Update

Count: 1   1

Basic terminology		
Points, lines, rays, segments, and vectors		
Patterns		
Congruence		
Similarity		
Parallels		
Triangles		
Quadrilaterals		
Circles		
Angles		
Polygons		
Polyhedra		
Models		
3-D Relationships		
Symmetry		
Transformations (e.g., flips or turns)		
Pythagorean Theorem		

Student Expectations

I. Memorize Facts, Definitions, Formulas
II. Perform Procedures
III. Demonstrate Understanding
IV. Conjecture, Analyze, Generalize, Prove
Solve Non-Routine Problems/Make

I.   I.
II.   II.
III.   III.
IV.   IV.

10

State Science Standards by TIMSS Science

We have included an example analysis of content alignment for one State's grade 8 standards (Wisconsin) with the TIMSS grade 8 assessment. The alignment index is .24 at the coarse-grain level (the large topic categories shown), and the overall alignment index is .10. The State grade 8 standards focus emphasis more heavily on Nature of Science (at all levels of expectations) as well as on Ecology and Earth Systems. The TIMSS assessment at grade 8 includes a range of emphasis in content including Science, Health & Environment, and a number of topics in Life Sciences, Physical Sciences, and Earth and Space Sciences. To fully examine these differences, other sources of information might be examined. For example, the State standards may include additional science topics at prior grade levels (5-7) that were not content analyzed. The TIMSS assessment is designed to assess student learning in science through grade 8 including prior grades.

A fine-grained chart for the same State (Wisconsin) by TIMSS science grade 8 displays the content analyzed under the topic of **Ecology.** The data show that the State standards emphasize instruction of only four topics under Ecology at grade 8 while the TIMSS assessment calls for students to respond to content under 8 topics and the expectations cover 3-4 levels for all of these topics. Thus, while at the large grain topic level there is some similarity in content focus, it is apparent that TIMSS assessment is more content demanding than the State standards. The "re-centered" (fine grain) alignment statistic for this particular topic is only .16.
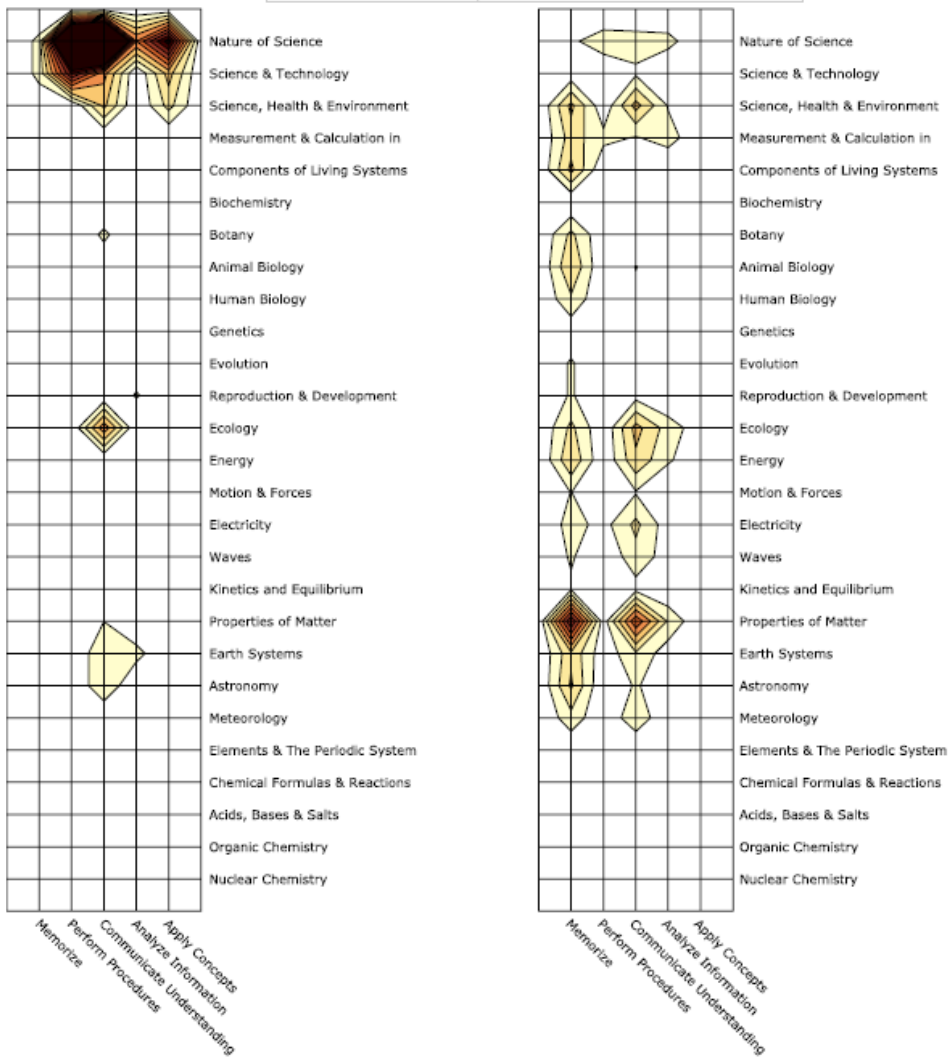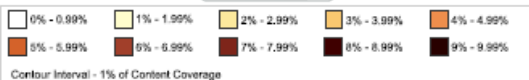
Ecology Standards vs. TIMSS grade 8

Alignment Overall: 0.0993
Alignment Re-centered: 0.161

= Not Covered
= < 0.5%
= < 1.0%
= < 1.5%
= >= 1.5%
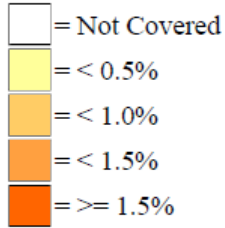
Show Data Tables

Administration Year:

Sample Selection: 2008 TIMSS Gr. 8 | WI Stnds Gr. 8

**Report By:** All Data | All Data

Update

Count: 1 | 1

- Food webs and chains
- Competition and cooperation
- Energy flow relationships
- Biotic and abiotic factors
- Ecological succession
- Ecosystems
- Population_dynamics
- Environmental_chemistry
- Adaptation and variation
- Niche populations

Student Expectations
I. Memorize
II. Perform Procedures
III. Communicate Understanding
IV. Analyze Information
V. Apply Concepts
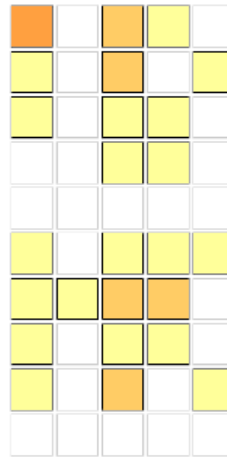
I.
II.
III.
IV.
V.

I.
II.
III.
IV.
V.

13
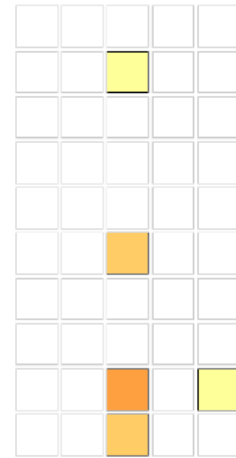
State Math Standards by PISA Math

The analysis of alignment of one State's (Idaho) high school standards for Mathematics in relation to PISA Math assessment shows a number of areas where math content is similar but there are also several areas of differences. The coarse grain alignment level (shown in the chart below) is .41, while the overall alignment statistics is .10. These statistics show that the content called for in this state at the high school level is fairly consistent with PISA at the main topic level but quite different at the fine-grain topics level and the levels of expectations.

The State standards include more mathematics topics than PISA, including advanced algebra, advanced geometry, probability, analysis, functions, and instructional technology. On the other hand, PISA focuses more emphasis on Measurement and Data displays.

The categories of expectations for student learning in both PISA and the State standards are primarily in the Perform Procedures category. Both also emphasize the conjecture/analyze expectation for geometry and statistics.   The PISA items call for a range of expectations for student learning for student assessment under the Basic Algebra topic, including solving non-routine problems, conjecture/analyze, and demonstrate understanding.

Update Map

| 0% - 0.99% | 1% - 1.99% | 2% - 2.99% | 3% - 3.99% | 4% - 4.99% |
| 5% - 5.99% | 6% - 6.99% | 7% - 7.99% | 8% - 8.99% | 9% - 9.99% |

Contour Interval - 1% of Content Coverage

Number Sense / Properties / Relationships

Operations

Measurement

Consumer Applications

Basic Algebra

Advanced Algebra

Geometric Concepts

Advanced Geometry

Data Displays

Statistics

Probability

Analysis

Trigonometry

Special Topics

Functions
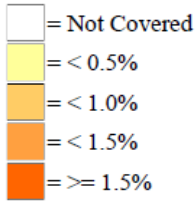
Instructional Technology
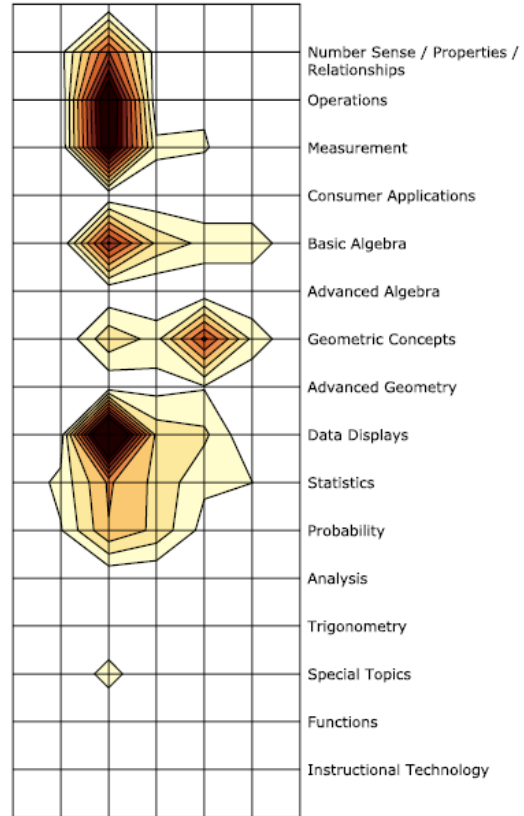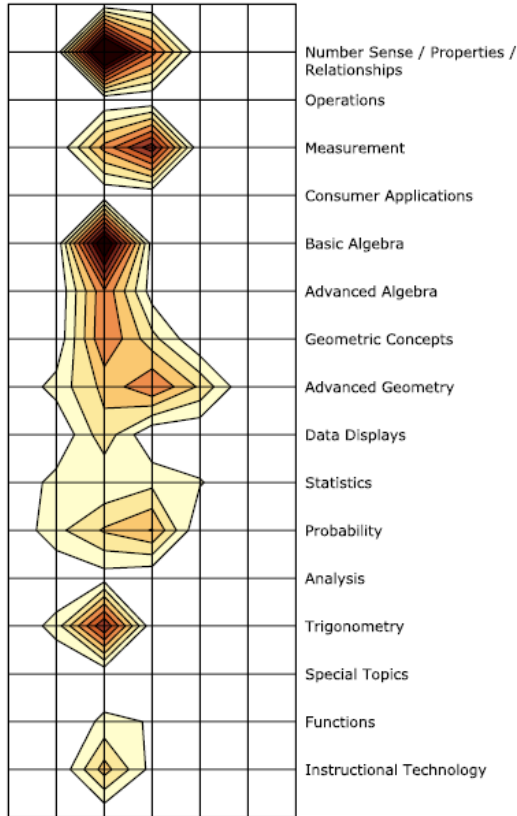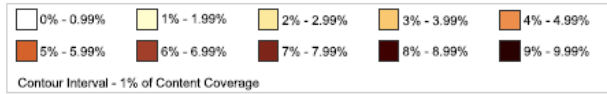
Memorize Facts, Definitions, Formulas

Perform Procedures

Demonstrate Understanding

Conjecture, Analyze, Generalize, Prove

Solve Non-Routine Problems/Make Connections

State Geometry Standards by PISA

Alignment Overall: 0.133
Alignment Re-centered: 0.1017

= Not Covered
= < 0.5%
= < 1.0%
= < 1.5%
= >= 1.5%

Show Data Tables

Administration Year:

Sample Selection: ID Stnds Gr. 9_12   2006 PISA Gr. 15 y.o.

Report By: All Data   All Data

Update

Count: 1   1

Basic terminology
Points, lines, rays, segments, and vectors
Patterns
Congruence
Similarity
Parallels
Triangles
Quadrilaterals
Circles
Angles
Polygons
Polyhedra
Models
3-D Relationships
Symmetry
Transformations (e.g., flips or turns)
Pythagorean Theorem

Student Expectations
I. Memorize Facts, Definitions, Formulas
II. Perform Procedures
III. Demonstrate Understanding
IV. Conjecture, Analyze, Generalize, Prove
V. Solve Non-Routine Problems/Make Connections

State Science Standards by PISA Science Assessment
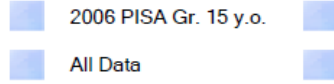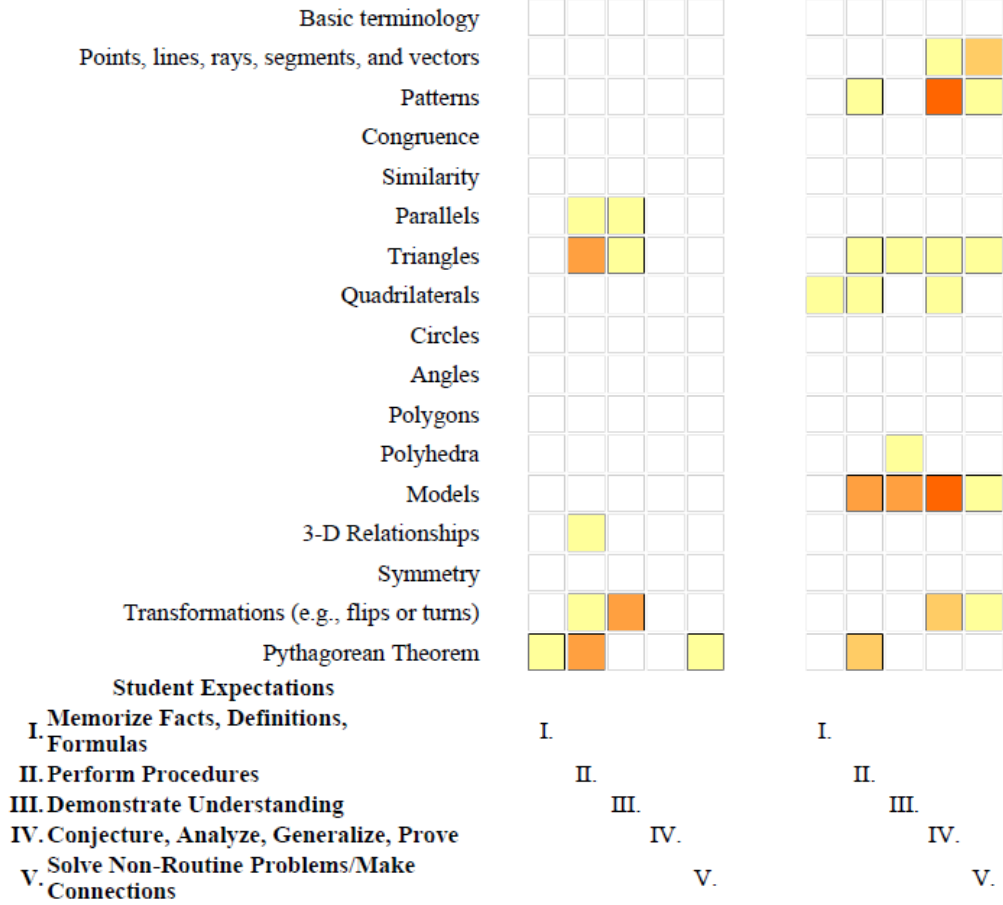
The analysis example of content alignment of one State (Ohio) grade 10 high school science standards and the PISA science age 15 assessment shows an index of .35 for coarse grain alignment (the large topic categories shown). The overall alignment index is .18. The Ohio grade 10 standards focus heavily on Nature of Science (at expectations levels 1 thru 4) as well as Living Systems, Evolution, Ecology, and Earth Systems. The expectations for learning are largely at levels 1-3 (Memorize, Perform procedures, Communicate understanding).

The PISA assessment emphasizes Nature of Science (levels 3-5), Science, Health & Environment, Measurement and Calculation, Human Biology, Evolution, Ecology, Properties of Matter and Earth Systems and Astronomy. Student expectations for PISA assessment are primarily at levels 3 (Communicate), 4 (Analyze) and 1 (Memorize).
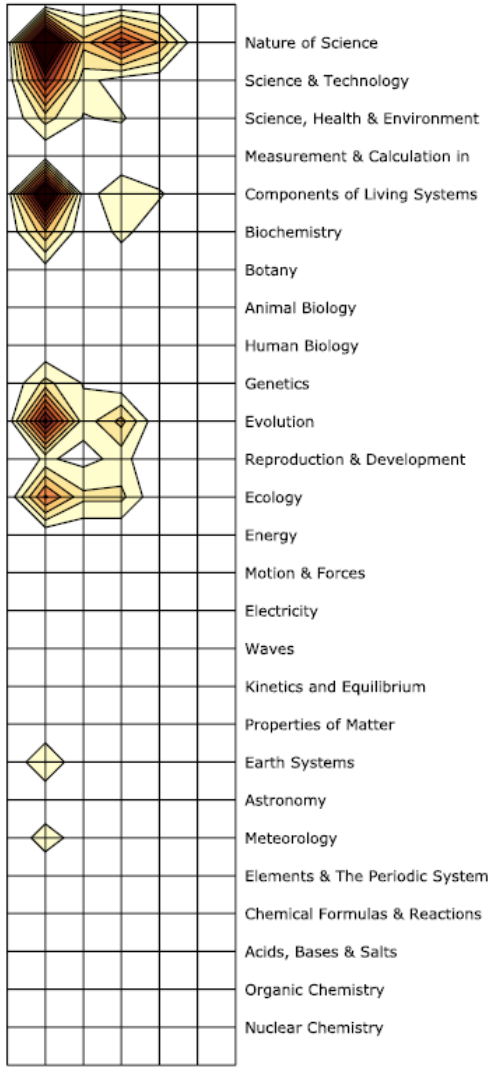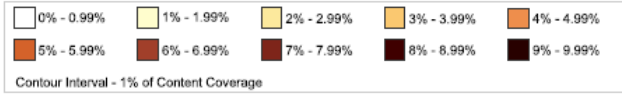
Data Cut: All Data

Count: 1

Data Cut: All Data

Count: 1

Update Map

0% - 0.99%	1% - 1.99%	2% - 2.99%
3% - 3.99%	4% - 4.99%	
5% - 5.99%	6% - 6.99%	7% - 7.99%
8% - 8.99%	9% - 9.99%	

Contour Interval - 1% of Content Coverage

Nature of Science
Science & Technology
Science, Health & Environment
Measurement & Calculation in
Components of Living Systems
Biochemistry
Botany
Animal Biology
Human Biology
Genetics
Evolution
Reproduction & Development
Ecology
Energy
Motion & Forces
Electricity
Waves
Kinetics and Equilibrium
Properties of Matter
Earth Systems
Astronomy
Meteorology
Elements & The Periodic System
Chemical Formulas & Reactions
Acids, Bases & Salts
Organic Chemistry
Nuclear Chemistry

Memorize
Perform Procedures
Communicate Understanding
Analyze Information
Apply Concepts

Summary

Based on the initial review of data by CCSSO and WCER and analysis of the content analysis and coding carried out by specialist teams, we are confident that the content analyses of TIMSS and PISA assessment items provide a fair and valid description of the content assessed by those items. We have provided some initial examples of how the content analyses can be used to analyze state standards in relation to these assessments. Users of the online system can also analyze the content of State assessments in math and science in relation to the international assessments. Our report includes an attached set of charts from the SEConline.org system provided in pdf format that shows both standards and assessment charts in relation to TIMSS and PISA content analysis.

The SEC content analysis methodology addresses the subject content of assessments and standards documents using two dimensions – Topics and Expectations (or cognitive demand). The analysis does not address other dimensions that users might be interested in analyzing an assessment, such as the type of design of items (forced choice vs. constructed response), the difficulty or rigor, or the quality of the item design and how it communicates to students. Several of these topics were addressed in a discussion with the content specialists at the conclusion of the two-day meeting of the teams. A summary of observations from the specialists about the TIMSS and PISA assessments is attached in the Appendix.

To review TIMSS and PISA Mathematics and Science content analysis data in relation to US national assessments, frameworks and state standards and assessments, through the SEC online system go to
http://seconline.wceruw.org/secWebHome.htm
click on **Content Analysis**; then, See,
"For access to content maps of Standards and Assessments analyzed thus far**," click here;**
then **Select Mathematics (Science), K-12, and Submit.**

In the left display chart, Use up or down arrow to select: 2008 TIMSS, grade 4; 2008 TIMSS, grade 8; or 2008 PISA.
In the right display chart, select any Standards or Assessment for comparison.
Click on **Update**

Data Charts Available for Review online

Mathematics

Grade 4 Ohio Math Indicators Standards by TIMSS grade 4

   Fine grain: Geometric concepts

Grade 4 NAEP vs. TIMSS grade 4

Grade 8 Minnesota Math standards by TIMSS grade 8

Grade 8 Montana Math standards by TIMSS grade 8

Grade 8 Oregon Math standards by TIMSS grade 8

Grade 8 NAEP vs. TIMSS grade 8

Grade 10 Idaho Math standards by PISA math

Grade 10 Ohio Math standards by PISA math

Grade 10 Ohio Math Test by PISA math

Grade 10 Rhode Island HS Math stands by PISA math

Grade 10 College Board HS Math by PISA math

Grade 11 Virginia Algebra standards vs. PISA math

Science

Grade 8 Wisconsin Science standards by TIMSS grade 8

    Fine grain: Ecology topic

Grade 10 State U Science by PISA science

Grade 4 Idaho Science standards by TIMSS grade 4

Grade 4 Indiana Science by TIMSS grade 4

Grade 4 Michigan Science by TIMSS grade 4

Grade 8 Illinois Science standards by TIMSS grade 8

Grade 8 North Carolina Science test by TIMSS grade 8

Grade 10 Ohio Science standards by PISA science

Grade 8 Oklahoma Science standards by PISA Science

Appendix

A) Notes from a Wrap-up Discussion on PISA and TIMSS, with Mathematics and Science Specialists, following completion of content analysis process, 11/12/08

Question to group: what additional supporting information do we need to help explain TIMSS and PISA for users of the SEC alignment analyses?

- Sample items would be the most helpful. Sample items are useful in understanding the intentions of the frameworks. For example, Singapore's framework doesn't use verbs and so exactly what is intended for mastery and measurement isn't entirely clear without seeing the items.

- It would be good if the website could allow side-by-side comparisons of more than two standards/assessment programs, or if there were a tool to isolate certain criteria and then see results for multiple programs.

- It would be useful to have some type of clear, global statements about the international assessments, particularly PISA. For example, PISA is not intended to measure school mathematics but rather literacy. Again, sample items help show what "literacy" means.

There was discussion and observations about PISA items in relation to national standards documents (e.g., ACHIEVE, AAAS). In general, both the math and science groups liked the PISA items and thought they tapped into something that is not assessed in NAEP, TIMSS, or many of the state assessments. They commented that the kinds of items in PISA are not necessarily what gets "pushed" in science assessments by states, but that these formats are relevant and important – they get closer to the idea of measuring how to think and apply than strictly measuring knowledge. The math groups did note that PISA has very little algebra or advanced algebra (considering the 15 year-old target), but that the level of reasoning and numeric literacy required is what gives the test validity and/or its difficulty.

This discussion was related to the idea that standards and the overall mindset nationally is still focused on a content orientation, so PISA-type items do not get a lot of emphasis in the U.S. context. Many participants noted that the reality is that state assessments have to embed their tests in content – it's harder for multiple reasons to do literacy – so there is a disjuncture between what the states are required to do with state assessments and where international assessments (PISA in particular) are going. However, educators do want to better address the PISA-type issues: inquiry, scientific habits of mind, etc.

One problem is treating these literacy/process skills as separate isolated units rather than integrating them across all content areas in science; for example, teaching reasoning and analysis is not something that teachers always know how to do. Another problem is that there is an external push (e.g., ACHIEVE) for content-driven standards and assessment and state education departments are being held accountable for content and not the

thinking skills side. It also was noted that advice and recommendations about state standards and assessments often comes from national organizations and consultants rather than from educators with state offices or within a state. Finally, there are financial considerations with building assessments that have lots of constructed response. The costs have been a factor in many states limiting the use of these types of items. Regarding TIMSS, there was some question about why the U.S. was not doing better because it appears very similar to NAEP and typical state assessments. One hypothesis was the low-stakes nature of the test.

B) Reliability Statistics for Content Analysis of TIMSS and PISA Assessments

The SEC methodology for conducting content analysis of assessment and other curriculum documents does not employ a consensus model, but rather encourages subject specialists to reach their own conclusions about the content being assessed by a given assessment document. The methodology accepts, and even expects that viewed from different perspectives, a given assessment item might be described differently by content experts.

In order for subject specialist analysts to gauge their own content descriptions, analysts work in teams of three to five members, and meet as a team to discuss each instrument or document being analyzed. These discussions focus on key assessment items (or passages of text in the case of academic content standards) selected by individual analysts as they complete the independent 'coding' phase of the analysis process. Analysts discuss the selected items in terms of the content descriptions, and the rationale for those descriptions for each item selected for discussion. After each analyst has shared his/her description and rationale with the team, and any ensuing discussion has occurred, each analyst may choose to change, add, delete, or keep as is, their original code. Oftentimes, though not always, analysts will reach a general agreement on how a particular item is best described, but the process does not require that agreement. Moreover, each rater can choose to utilize more than a single content description (up to 3 for assessments, 6 for standards, by convention). Because there is not a consistent number of descriptions provided for any one item across raters, calculating inter-rater reliability is not as straightforward as it might be for other data-sets. The method employed for judging inter-rater reliability, results of which are presented below, is arguably the most sensitive measure of agreement between raters, accounting for all agreements and disagreements in content descriptions.

Inter-rater Reliability Statistics				
	Mathematics		Science	
	Crs. Grain	Fine Grain	Crs. Grain	Fine Grain
Grade 4 TIMSS	0.78	0.59	0.65	0.69
Grade 8 TIMSS	0.73	0.53	0.62	0.69
Age 15 PISA	0.69	0.49	0.43	0.46

The results reported in the table are based upon comparisons of each analyst's description of a given instrument or item pool, against every other analyst's description of that same document or set of assessment items. Extent of agreement is measured using the alignment calculation typically used to compare alignment between standards and assessments, or between standards and practice. In this case however the alignment is measured between raters of the same document, and then the average of all pair-wise combinations of raters is used to generate the mean alignment index across raters. Since alignment can be calculated at both coarse grain (content areas) and fine grain (topic) levels of distinction, two alignment results are reported for each item pool: a coarse grain

result and a fine grain result. Both coarse and fine grain alignment results make distinctions by category of cognitive demand, so that even if two raters agreed on the topic being assessed, if they did not agree also on the category of cognitive demand being assessed, then their descriptions would not count as agreement, and thus would not count as 'aligned' content for purposes of calculating alignment.   It is this mean alignment number that is reported in the Content Analysis charts presented with the report.

Looking at the results in the table, two patterns are clearly discernible: Analysts of the mathematics assessments had relatively higher levels of inter-rater agreement or reliability, and analysts of the TIMSS assessments had higher levels of inter-rater agreement compared to the ratings of the PISA item pool, whether looking at mathematics or science

References

Blank, R.K., Porter, A. & J. Smithson, J. (2001) *New Tools for Analyzing Teaching, Curriculum and Standards in Mathematics & Science: Results from Survey of Enacted Curriculum Project*, Washington, DC: CCSSO.

Porter, A. (2002) "Measuring the Content of Instruction: Uses in Research and Practice," AERA Presidential Address, *Educational Researcher,* October.

SEC Websites
http://www.SECsurvey.org or
http://www.ccsso.org/Resources/Programs/Surveys_of_Enacted_Curriculum_(SEC).html

http://www.SEConline.org or
http://seconline.wceruw.org/secWebHome.htm

TIMSS and PISA information
http://nces.ed.gov/surveys/international/