Achieve, Inc.

Measuring Up 2004

A Report on Language Arts Literacy and Mathematics Standards and Assessments for

NEW JERSEY

ACHIEVE'S BENCHMARKING INITIATIVE

Achieve, Inc.

Created by the nation's governors and business leaders, Achieve, Inc., is a bipartisan, non-profit organization that helps states raise academic standards, improve assessments and strengthen accountability to prepare all young people for postsecondary education, work and citizenship. Achieve has helped nearly half the states benchmark their standards and tests against the best examples in this country and abroad and work in partnership to improve teaching and learning. Achieve serves as a significant national voice for quality in standards-based reform and regularly convenes governors, CEOs and other influential leaders at National Education Summits and other gatherings to sustain support for higher standards and achievement for all of America's schoolchildren.

Achieve helps states raise academic standards, measure performance against those standards, establish clear accountability for results and strengthen public confidence in our education system. To do this, we:

- help states **benchmark** their standards, assessments and accountability systems against the best in the country and the world;
- provide sustained **public leadership** and advocacy for the movement to raise standards and improve student performance;
- build **partnerships** that allow states to work together to improve teaching and learning and raise student achievement; and
- serve as a **national clearinghouse** on standards and school reform.

Michael Cohen, President

Matthew Gandal, Executive Vice President

Jean Slattery, Director of Benchmarking

Measuring Up 2004

A Report on Language Arts Literacy and Mathematics Standards and Assessments for

NEW JERSEY

ACHIEVE'S
BENCHMARKING INITIATIVE

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Achieve, Inc., is a bipartisan, non-profit organization created by the nation's governors and corporate leaders to help states raise their academic standards, improve their assessments and strengthen accountability to prepare all young people for postsecondary education, work and citizenship. A principal part of Achieve's mission is to provide state policymakers with an independent, expert review of the quality of their standards and tests.

Measuring Up 2004: A Report on Language Arts Literacy and Mathematics Standards and Assessments for New Jersey represents Achieve's third time working with the state. Four years ago, Achieve reviewed New Jersey's Core Curriculum Content Standards in language arts literacy and mathematics and analyzed the alignment of the state's tests at grades 4 and 8 and high school with those standards. In 2001, Achieve evaluated the state's science standards and assessments at grades 4 and 8, which resulted in many improvements to the science standards.

Since Achieve's first report in 2000, New Jersey has revised its Core Curriculum Content Standards in language arts literacy and mathematics. At the request of Governor James E. McGreevey and Commissioner of Education William L. Librera and with the encouragement and support of New Jersey United, a statewide business and education coalition, Achieve has reviewed the state's revised standards and the related tests. The goal is to help the state continue its progress in implementing a high-quality, aligned system of academic standards and assessments.

This report summarizes Achieve's findings and provides policymakers with answers to the following questions:

- How do New Jersey's revised standards in language arts literacy and mathematics compare with the earlier version of the standards? How do they compare with the best standards from other states and nations?
- How well do New Jersey's grade 4, grade 8 and high school tests assess the knowledge and skills found in the state's new standards? Are the tests sufficiently challenging?

RESULTS FOR NEW JERSEY

- **New Jersey's revised mathematics standards are a significant improvement over the previous version — they are as challenging as the best domestic and international standards.**

New Jersey has made significant progress since 2001 in raising the quality of its mathematics standards. The state has been successful in increasing the level of rigor of the standards by raising the overall intellectual demand, ensuring essential

knowledge and skills are included, and emphasizing core content. It has improved the progression of its standards through well-considered development of each content domain across grades 2–8. In the process, New Jersey increased the specificity and measurability of its standards and also clarified the use of technology. By reducing the number of standards from 16 to five, the state was able to highlight connections among the principal domains of mathematics and enhance the focus, coherence and utility of the document.

- **The revision of the language arts literacy standards resulted in a set of strong standards for the early grades, but the late-elementary, middle and high school standards still need improvement.**

New Jersey's standards in language arts literacy have improved but to a lesser degree than the math standards. Headway has been made in raising rigor and articulating progression, but this is generally confined to grades K–4, where New Jersey has given special attention to early literacy. The standards in grades 5–12 are not as clearly articulated as those in K–4 and, as a result, are not as challenging.

- **New Jersey's 4th-grade mathematics assessment is rigorous and well aligned with the standards, but the 8th- and 11th- grade tests are not as challenging as the standards imply they should be.**

Achieve reviewers were impressed with the quality of the state's 4th-grade math test. The test questions are well crafted and have an appropriate level of intellectual demand. The test also makes good use of its open-ended items to assess standards not readily assessed by multiple-choice items, namely, reasoning, problem solving and communicating mathematically.

The 8th- and 11th-grade tests did not fare as well in the Achieve review. Although the 8th-grade test shows a good emphasis on algebra, the relatively low level of demand of the test as a whole results from having too many "whole number" items and too many items that do not require students to demonstrate their mathematical knowledge in appropriately challenging ways. The same issues with low cognitive demand are found in the high school assessment. In fact, the test aligns better with the 8th-grade standards than the 12th-grade standards.

- **New Jersey's assessments in language arts literacy have some strengths, but they do not effectively measure the depth and breadth of the standards.**

Achieve found New Jersey's assessments in language arts to have three exemplary characteristics: they contain authentic passages; they include open-ended items that are designed to assess advanced thinking — a hallmark of quality tests; and they directly measure writing to determine how well students respond to the totality of a writing task — organizing and developing their response to match a specific audience and purpose.

Despite these strengths, however, the tests do not effectively measure the depth and breadth of the state standards. The most significant problem is on the reading portion of the tests, which contain lengthy reading passages but relatively few questions about those passages. Reducing the length and increasing the number of the passages would allow for a more thorough coverage of the standards and yield a more reliable indication of student performance in relation to the standards.

A related concern is the level of cognitive demand of New Jersey's tests. The reading passages on the tests are grade-appropriate, but none are likely to challenge high-end readers, except for the high school narrative selection. The reading questions also are not as challenging as they could be.

RECOMMENDATIONS FOR MOVING FORWARD

As New Jersey continues to move forward in its steady pursuit of a rigorous and aligned system of standards and assessments, Achieve recommends that the state consider making the following improvements:

- ✓ **Strengthen the language arts literacy standards in grades 5–12.**

New Jersey's language arts literacy standards in grades K–4 provide a model for improving the standards in grades 5–12. There is inconsistent development of key concepts in the standards from grade to grade and great variation in the grain size of the standards. Specifying the quality and complexity of texts students should read through the grades would help to clarify expectations in the standards, as would differentiating between the comprehension skills needed to read informational and literary texts.

- ✓ **Better define the mathematics and language arts literacy standards at the high school level.**

New Jersey has clearly defined student expectations for reading in grades K–4 and for math in grades 2–8. The standards for high school students, on the other hand, are not as well defined. In both subjects, standards are articulated for the 8th and 12th grades but not for the grades in between. This makes it more challenging for local districts and teachers to build curricula, and it makes it particularly difficult to determine what students are to be responsible for on the state's graduation exam administered in 11th grade. Achieve encourages the state to further delineate its standards in high school, either by creating grade-by-grade or course-by-course expectations or by breaking the 9–12 grade cluster into two-year spans.

- ✓ **Increase the level of cognitive demand on both the mathematics and language arts literacy assessments, particularly the grade 8 and high school tests.**

Neither the 8th- nor the 11th-grade math assessments are as challenging as they could be. Over time, Achieve recommends that New Jersey include more rigorous content and tap more advanced skills on both of these tests. This will help bring these tests into better alignment with the new state standards.

To increase the rigor of the language arts literacy assessments, New Jersey should increase the number of items that ask students to make more complex inferences and evaluate, extend and apply what they read. As the tests are currently written, too few items ask that students do little more than demonstrate basic comprehension. Reducing the length of the reading passages would allow New Jersey to include more passages of varying genres and also, more importantly, more items that test a broader range of language arts skills.

* * *

New Jersey has a history of undertaking periodic reviews of its standards and assessments and using the findings to bolster its system. This commitment to continuous improvement is admirable and important; it reflects an appreciation for the ever increasing demands and opportunities students face when they graduate from high school, as well as the need for the state to raise expectations over time to help ensure all students are prepared for these opportunities.

Achieve strongly encourages New Jersey to continue to upgrade its standards and tests, increasing rigor over time, to ensure that its graduates will be adequately prepared. This is not an easy task, nor is New Jersey the only state that needs to undertake it, but the state is well positioned to make it a priority. New Jersey has succeeded in building a long-term leadership coalition of the Department of Education, the governor's office, and the business and education communities that has helped set the direction for reform in the state. By building on the tangible improvements of the past and taking advantage of its assets, New Jersey can continue to raise expectations and achievement for all students.

ACHIEVE'S WORK WITH NEW JERSEY

Achieve was established after the 1996 National Education Summit by the nation's governors and business leaders to provide advice and assistance to state policy leaders on issues of academic standards, assessments and accountability. Under the auspices of Achieve's Benchmarking Initiative, 17 states have sought Achieve's external reviews of state education policy since 1998.

In 2000, Achieve completed a benchmarking review of New Jersey's Core Academic Standards in language arts literacy and mathematics, as well as an analysis of the alignment of related state tests at grades 4 and 8 and high school. In these reviews, Achieve found that the standards were overly broad, and we encouraged the state to make them clearer and more challenging. Achieve also provided advice for strengthening the assessments. A similar evaluation of the state's standards and assessments at grades 4 and 8 in science was conducted in 2001, and the state was able to make the most of Achieve's recommended changes to the science standards.

Since that time, the state has revised its Core Curriculum Content Standards in these subject areas and has drafted a set of standards for social studies to be submitted to the State Board for approval in March 2004. At the request of Governor James E. McGreevey and Commissioner of Education William L. Librera and with the encouragement and support of New Jersey United, a statewide business and education coalition, Achieve has agreed to review the state's updated standards and assessments in language arts literacy and mathematics and offer advice for continuous improvement aimed at helping New Jersey continue its commitment to excellence in as expeditious a manner as possible.

COMPONENT	DATE OF LAST REVISION
Content Standards: Language Arts Literacy, Mathematics, Science	2002
Benchmark Tests: Grades 4, 8, High School	Grade 4 (2003), Grade 8 and High School (in process)
Additional Tests (NCLB required)	Grade 3 (2003), Grades 5 and 6 (2005), Grade 7 (2006)
Test Specifications (benchmark tests)	(2001)
Sample Tests	Grades 4 and 8
Curriculum Frameworks	Language Arts Literacy (1998), Mathematics (1996)
CD Rom*	2001
State Report Card	2003
Performance-Based Assessment Pilots	2003

* The CD Rom, distributed to every teacher, includes the content standards, the frameworks, test specifications and a sample test.

At the heart of implementing an effective standards-based system is developing essential components, revising them on a regular cycle and keeping them in alignment. New Jersey

has established a system of instruction that is quite comprehensive in scope. The chart above provides an overview of the language arts and mathematics components in New Jersey's system and indicates when these components were last updated.

The last entry in the chart represents New Jersey's pioneering effort to develop performance-based assessments that can be wedded to its larger system of required tests. Although many thoughtful policymakers have advocated for multiple measures — appreciating that many complex performances, such as conducting research, cannot be assessed by large-scale, on-demand tests — few states have actually taken steps toward their implementation. New Jersey, however, has initiated a pilot study involving nine districts with the intent of making performance-based tasks part of its overall assessment system by school year 2007–08.

Currently, the state finds itself at a crossroads with a need to identify the most efficient and effective way of updating and aligning the components of its system so quality is not compromised and teachers and students receive maximal support. Achieve hopes the findings in this report will assist the state in its decision-making process.

In writing this report, Achieve synthesized the reviews of New Jersey's Core Curriculum Content Standards and the alignment of the assessments to the standards. These studies were conducted by teams of national experts with significant experience in analyzing academic standards and tests. The findings in this report represent consensus opinions of Achieve's experts, but final judgments and conclusions rest with Achieve. In addition to this summary report, Achieve has prepared a detailed technical report for the New Jersey Department of Education. Because the technical report contains references to secure test items, it is confidential.

Brief biographies of experts and consultants who participated in Achieve's standards benchmarking and assessment analysis for New Jersey can be found in the Appendix.

THE ACHIEVE BENCHMARKING METHODOLOGY

THE STANDARDS BENCHMARKING STUDY

Achieve compares a state's standards with state, national and international benchmark standards recognized for their quality and/or for producing high student achievement. To ensure that the most important aspects of standards, as described below, are addressed, Achieve prepares a set of guiding questions for our expert reviewers.

- **RIGOR**: **What is the level of intellectual demand of the standards?**

  Expectations for student learning should be sufficiently intellectually challenging to equip students with the knowledge and skills they will need to succeed at the next level. Rigorous standards closely represent and build on the essential core content of a discipline, providing a balanced perspective of key understandings and skills. Rigor is the most complex of the criteria used to evaluate standards, because it depends on the interplay of the content chosen for emphasis, how the intellectual demand of knowledge and skills evolves from grade to grade, and the precision with which expectations for learners are expressed — all as compared with state and international benchmarks.

- **FOCUS: Have judicious choices been made about what is most important for students to learn?**

  High-quality standards establish priorities about the concepts and skills that should be emphasized at each grade level. Choices should be based on the core knowledge and skills essential for students to advance to the next level of understanding. A sharpened focus also helps ensure that the amount of content to be learned in each grade level is manageable.

- **COHERENCE: Do the standards convey a unified vision of the discipline, and do they establish connections among the major areas of study?**

  Standards should be categorized and broken out into supporting strands that reflect the way the discipline is structured; they should not be an array of disconnected bits of knowledge and skills. Further, the standards should reveal significant relationships among the strands — their key concepts and the way they are linked to one another — and how the study of one complements the study of another.

- **PROGRESSION: Do knowledge and skills build clearly and sensibly on previous learning and increase in intellectual demand from grade to grade?**

  Development of coherent understanding in a subject area requires a carefully staged evolution of knowledge and skills, generally moving from the simple to the complex and from the concrete to the abstract. Standards must reflect this development. They must introduce content at the appropriate grade, grow more intellectually challenging from grade to grade, and they must delineate a

progression of knowledge and skills — rather than repeating concepts from year to year.

- **SPECIFICITY: Are the standards specific enough to convey the level of performance expected of students?**

  High-quality standards provide a sufficient amount of detail without being overly prescriptive and without becoming unmanageable for teachers. Overly broad standards leave too much open to interpretation, increasing the likelihood that students will be held to different levels of performance. Overly detailed standards encourage a checklist approach to teaching and learning that undermines students' overall understanding of the discipline. Furthermore, standards that maintain a relatively consistent level of precision ("grain-size") are easier to understand and use.

- **CLARITY: Are the standards clearly written and presented in a logical, easy-to-use format?**

  Clarity requires more than just plain and jargon-free prose. Grade-level expectations must be communicated in language that can gain widespread acceptance by teachers, parents, school boards and others who have a stake in schooling. A straightforward format, based on the structure of the discipline, facilitates user access and makes grade-level comparisons transparent.

- **MEASURABILITY: Does each standard describe the results of student learning in some observable or verifiable way?**

  Standards should focus on the results, rather than the processes, of teaching and learning. Objectives or indicators should make use of verbs calling for student performances that demonstrate knowledge and skills and avoid using those that refer to learning activities (such as examine, participate and explore) or to cognitive processes (such as appreciate, know, and learn) without articulating what it means for a student to recognize, know or learn a concept.

BENCHMARK STANDARDS

To ensure that the benchmark standards documents used as exemplars are indeed the best for this purpose, Achieve commissions expert reviews of a variety of sets of standards. Reviewers concluded that the English language arts standards from California (1997) and Massachusetts (2001) and the early literacy standards from North Carolina (1999), Texas (2001) and New Standards (1999) had the most value for benchmarking. In mathematics, Achieve chose standards from Indiana (2000) and Singapore (2001), as well as its own document *Foundations for Success* (2002), which details the mathematics that we believe all students should be expected to know before leaving 8th grade. Developed by a panel of leading research mathematicians and math educators, and based on a close analysis of the curricular expectations in the highest-performing European and Asian nations, *Foundations for Success* tells us where we need to aim if we want the next generation of

America's young people to be internationally competitive in their mathematical knowledge and understanding.

Selecting these benchmarks proved a difficult task because no single set of standards is perfect and judgments about the quality of standards are in some ways subjective. Still, we are confident that the choices used in our current work reflect some of the best thinking from around the country and that a careful comparison of a state's or district's standards with these benchmarks will yield helpful diagnostic information and policy suggestions for states and districts to consider.

THE ALIGNMENT STUDY

Alignment is a measure of the extent to which standards and assessments agree and the degree to which they work in conjunction to guide and support student learning. It is not a question that yields a "yes" or "no" response; rather, alignment is a considered judgment based on a number of factors that collectively determine the degree of match between a state's standards and the assessment used to gauge if students are meeting those standards. At its core, the Achieve analysis answers two key questions: Can everything on the assessments be found in the standards? In addition (and conversely), do the assessments do an effective job of measuring the knowledge and skills set forth in the standards?

METHODOLOGY

To determine how closely each New Jersey assessment was aligned to the related grade-level standards, Achieve convened two teams of content experts who followed a subject-specific, stepwise procedure, or *protocol*, that Achieve has used to evaluate numerous assessments in more than a dozen states.

In the first phase of the review process, a team of content experts evaluates each individual test item to determine (1) if it actually measures the indicator to which the test developer assigned it; (2) how well it matches the content and performance described in the related standard; (3) whether it is fairly constructed; and (4) how intellectually challenging it is. These are key issues. The information gained from a test is no better than the collection of items that make it up. If an item measures content and skills beyond what is contained in the standards, it is less likely that it will have been taught in classrooms. Similarly, an item that is flawed for such reasons as having no right answer, more than one right answer, a misleading graphic or implausible distracters will not give accurate information about students' performance. Tracking the level of cognitive demand that each individual item poses also is critical. If a test is truly standards-based, it should have a mix of basic and more challenging items that reflect the range of concepts and skills spelled out in the standards so differences in the performance of proficient and non-proficient students can be detected. In summary, Achieve's item-by-item analysis not only yields valuable information about critical aspects of alignment but also provides

quantitative data that contribute to the judgments made with respect to the overall balance and rigor of a test, as described below.

In the second phase of the alignment study, content experts take a more holistic view of the test in order to judge if the test is balanced overall and if it is appropriately rigorous for the grade level. Moving away from the item level, reviewers consider the test one standard at a time — such as literary response or geometry — and look at the collection or set of items that are meant to assess each standard.

To judge how balanced the set of items mapped to each standard is, experts ask, "Does this set of items succeed in measuring the breadth and depth of content and skills described in the standard?" Said another way, "To what extent does the set of items assess the key content and skills in the standard?" Because a single on-demand test cannot assess all the indicators that make up a state's standards, it is crucial to determine how well the items on a test sample the most essential indicators. Content experts also examine the reading passages as a set to ensure that both literary and informational text forms are fully represented and that the range of writing prompts reflects the variety of genres represented in the standards. In evaluating the rigor of a test, content experts follow the same general procedure they use when evaluating balance. They compare the overall intellectual demand encompassed by a set of items with the level of intellectual demand described in the related standard. Looking at each standard in turn, they ask, "Does doing well on the item set, which measures this standard, mean the student has mastered the challenging material contained in the standard?" Because experts rated each item earlier in the process as to its level of cognitive demand, they can determine if an item set has a span of difficulty appropriate for the grade level. Content experts also review the reading passages as a set to determine if they have a span of demand appropriate to the grade level tested, and they review writing prompts to ensure the genre, topic and characteristics of a response that will meet standards is clearly communicated in the directions to students.

At the close of the analysis, reviewers look across the standards, at the test as a whole, to determine how good a job the test does in measuring the knowledge and skills described by the standards and how rigorous the test is overall.

MAJOR FINDINGS: LANGUAGE ARTS LITERACY CORE CURRICULUM CONTENT STANDARDS

The New Jersey Core Curriculum Content Standards for language arts literacy are comprised of five main standards — Reading, Writing, Speaking, Listening, and Viewing and Media Literacy. Each standard is further broken out by strands (lettered) and learning expectations (numbered) — also called "cumulative progress indicators" — for individual grades (K–4) and grade-level clusters (5–6, 7–8, and 9–12). Achieve first reviewed the 1996 edition of New Jersey's language arts literacy content standards in 2000. Our second review provides us with an opportunity to look back, document the progress that has been made and suggest ways in which the state can continue to improve its standards.

ACHIEVE CRITERIA	LANGUAGE ARTS LITERACY STANDARDS	
	2000 Review	2003 Review
• **RIGOR**		
— Intellectual demand	⊙	◎
— Essential knowledge & skills	○	⊙
— Balance	○	⊙
• **FOCUS & COHERENCE**		
— Emphasis on core content	⊙	⊙
— Manageability	●	●
— Unified vision of discipline	●	⊙
• **PROGRESSION**		
— Placement of knowledge and skills	⊙	⊙
— Development of content & skills by strand	○	⊙
— Grade-by-grade articulation	○	⊙
• **SPECIFICITY**		
— Precision	○	⊙
— Sample text/recommended reading list	○	○
• **CLARITY**		
— Language	○	⊙
— Organization & format	⊙	⊙
• **MEASURABILITY**	⊙	⊙

Key:
○ Needs Significant Improvement
⊙ Needs Some Improvement
◎ Approaching Benchmark Quality
● Meets Benchmark Standards

COMPARISON OF NEW JERSEY'S 2002 STANDARDS FOR LANGUAGE ARTS LITERACY WITH ITS 1996 STANDARDS

When New Jersey's language arts literacy standards of 1996 and its current standards of 2002 are evaluated against Achieve's defined criteria for benchmark standards, it is apparent that the state has made progress in strengthening its Core Curriculum Content Standards. However, there are a number of key areas where improvements are still needed if the standards are to compare favorably with the best. The chart above

summarizes core elements of Achieve's criteria along with New Jersey's success to date in improving its standards.

As the chart indicates, New Jersey has made significant strides in addressing two of the most important criteria that distinguish exemplary standards — rigor and progression. The revised language arts literacy standards are more demanding for the grade levels in which they are placed than previously, and they contain fewer gaps in essential content and skills than did the state's 1996 standards.

New Jersey has developed standards for additional grades and has broken out grade-level expectations for grades K–4, delineating a sequence of knowledge and skills that can readily be traced across the grades. As a result, the development of content and skills across strands is more evident than in the previous edition of the standards. Indeed, the new K–4 standards provide a model that the state should use as a guide for creating grade-level standards through grade 12.

The most serious flaw in the standards continues to be a lack of specificity. Specificity is a pivotal criterion — when standards are imprecisely worded and are not grounded by sample text or recommended reading lists, clarity can be diminished, grade-to-grade progression obscured and the level of rigor reduced. We strongly recommend the state address this issue.

COMPARISON OF NEW JERSEY'S 2003 STANDARDS WITH THE BENCHMARK STANDARDS

STRENGTHS OF THE CONTENT STANDARDS

- **New Jersey has developed grade-by-grade standards in K–4 that provide solid guidance for teachers, especially in the area of early reading.**

The most significant area of improvement in the state's new language arts literacy standards is in the early grades. New Jersey has developed grade-level indicators for kindergarten though 4th grade for all five areas of language arts. For the most part, the K–4 standards communicate a progression of knowledge and skills that lay the foundation for a solid language arts curriculum and provide clear and specific guidance for teachers at each of these grade levels. Particularly in the area of early reading, the K–4 standards describe with a good deal of clarity how skills are to be developed. New Jersey should extend this framework and develop a full complement of standards to K–12.

- **New Jersey's language arts literacy standards are comprehensive in scope; they provide a framework for addressing most of the essential content of the discipline.**

By housing the standards within five categories (reading, writing, speaking, listening and media study), the New Jersey document provides an effective overarching, conceptual

framework. Of special note is New Jersey's emphasis on listening, speaking, and media literacy — essential skills that are often somewhat skimmed over in many state standards. New Jersey presents them as separate standards, thereby affording them a place of prominence rather than simply paying lip service to them. As media have so visibly changed what and how people read and write, students' ability to evaluate the credibility of what they see and hear and to recognize media strategies is increasingly important. Our reviewers agreed that at this point in time, it is wise to have these strands stand on their own; although they cautioned the state not to imply that these skills be isolated in terms of instruction.

AREAS FOR IMPROVEMENT

- **The level of rigor of New Jersey's language arts literacy standards is difficult to judge because the standards do not sufficiently specify the level of performance expected of students.**

In language arts literacy, two factors control the level of cognitive demand in a standard: one is the intellectual complexity of the *task* — what students are asked to do with text; the other is the quality and complexity of the *text* students are asked to read. Therefore, to make the expectation clear at a grade level, it is essential to identify not only the skill called for but also the quality and complexity of text to which the skill is applied. When the level of intellectual demand at each grade is grounded this way, the progression of skills across the grades is more transparent.

The New Jersey standards lack sufficient specificity in describing both tasks and texts. Take, for example, the following indicator: Reading Grades 7–8 E (2): *Use increasingly complex text guides, maps, charts, and graphs to assist with reading comprehension.* This expectation begs the questions: What are the characteristics and types of organizers that students should be working with? And are students expected to create these organizers to help organize their understanding and knowledge or simply to read them to answer questions?

Consider another example from the descriptive statement preceding Standard 3.1 Reading: "Students should read grade-appropriate or more challenging classic and contemporary literature and informational readings both self-selected and assigned." Again, there is no indication of what "grade-appropriate" means or what students are expected to do with — or learn from — their reading.

Striving for an appropriate level of specificity in standards is worth the effort because the precision with which standards are stated has a major impact on how teachers, parents and students interpret and implement them. As with all criteria, there are trade-offs. When standards are too specific, they read as if they were part of a checklist, and the overall significance of a concept or skill may well be lost. When they are overly general, however, they are open to differences of interpretation; teachers, students and parents are left with insufficient guidance as to what is most important for students to learn. Since students receive instruction in a subject from a dozen or more teachers in the course of

their K–12 education, widely varying interpretations may result in students experiencing very different treatments of essential content, as well as inevitable gaps in coverage. The greater clarity that states provide, the more likely it is that schools will hold students to consistent expectations so that over time they can successfully close the achievement gap. In addition, without the clarity provided by examples of text, it is quite difficult to gauge the rigor of New Jersey's standards. For example, it is hard to grasp the cognitive demand of indicators such as the grade 5/6 indicator, "Recognize characterization, setting, plot, theme, and point of view in fiction," or the grade 7/8 indicator, "Analyze ideas and themes found in texts," without first knowing which texts students will be working with.

Achieve's benchmark states have come up with different strategies for communicating the level of text students are expected to comprehend. Massachusetts offers a series of sample grade-level reading passages in their standards, while New Standards provides a sample reading list. Indiana responds to the need for precision by including examples in the indicators themselves:

> **Indiana Grade 5 Standard** *"Narrative Analysis of Grade-Level-Appropriate Text,"* **Indicator 5.3.4:** Understand that theme refers to the central idea or meaning of a selection and recognize themes, whether they are implied or stated directly.
>
> ***Example****: Describe the themes in a fictional story, such as* A Wrinkle in Time *by Madeleine L'Engle, in which the themes of courage and perseverance are explored as the children in the story go on a dangerous mission in search of their scientist father.*

Adopting any one of these strategies will make the expected level of student performance far more concrete.

- **The New Jersey standards do not give sufficient attention to some essential content, especially literature and literary analysis.**

The New Jersey standards embed expectations about students' knowledge of literature into the Reading strand (and to a lesser extent the Writing strand). For the most part, literature is addressed in 3.1 Reading G. Comprehension Skills and Response to Text. Although one can make a reasonable rationale for this embedding, the consequence is that this central area of the reading and language arts domain becomes "one-among-many" in a list of comprehension goals and gets little systematic attention. Furthermore, because of the comprehension context, the expectations neglect content knowledge goals, such as the study of American or other literary traditions, familiarity with the works of particular authors or periods, understanding of author's craft, knowledge of genres, and facility with literary elements and analysis. Teachers may very well come away from this document without a strong sense of what students are expected to know and be able to do in this area across the grades — or worse, concluding that this area is simply not important for students to study.

The benchmark documents all do a more thorough job of specifying expectations related to literary traditions, literary elements and literary analysis. In addition, as stated above, the benchmark documents specify either directly or through appended lists some of the texts with which students should be familiar, whereas New Jersey assumes that students will comprehend more complex and varied genres as they move across the grade levels but does not specify the quality or complexity of texts.

Reviewers identified three other areas where expected student performances should be described more completely:

1. Grammar, usage and spelling are not as fully delineated in the New Jersey standards as they are in the benchmark standards. New Jersey's expectations make no mention of the structure of English language or where and how students should acquire the ability to write with standard English usage.

2. The viewing and media standard is less developed than New Jersey's other standards in language arts. While New Jersey's intent in culling out this standard as a separate area of study is laudable, the strand is sparsely articulated. The state should develop a systematic set of progress indicators, leading to a much fuller specification of expectations regarding media. Otherwise, providing instruction in the viewing and media standard could be relegated to an afterthought, as it is not entirely clear what teachers are expected to teach students along the way.

3. Vocabulary development is given short shrift in the New Jersey standards. Vocabulary study is a widely accepted and critical aspect of reading comprehension, and the state would be wise to be more specific in its articulation of how students should develop a broad vocabulary.

- **The language arts literacy standards do not delineate clear lines of progression for each strand across successive grade levels.**

A thoughtfully developed progression of knowledge and skills is a hallmark of exemplary standards. Standards must introduce key concepts at the appropriate grade and show how they grow more intellectually challenging from grade to grade, rather than repeating the same skills from year to year. Constructing language arts standards is especially challenging because for the most part a similar set of knowledge and skills are present from grade to grade but spiral up in demand in terms of the task presented and, in reading, the text to which the task applies. Avoiding repetition, yet ensuring that lines of development are clear and coherent, is a difficult line to walk.

Although New Jersey has made strides since its 1996 standards, there are still two factors that undermine progression in the current version of the standards. First, there is inconsistent treatment of key concepts through the grades; some are well developed, others are not. Second, the cumulative nature of the standards leaves too much of the development open to interpretation.

For the most part, New Jersey has made defensible decisions about where in the standards core knowledge and skills are introduced, especially in the early grades. However, the grade-level placement of content at grades 4–12 is inconsistent — knowledge and skills often are introduced and then disappear from the standards only to reappear later, as in the following example:

> **Grade 4: 3.1 Reading G.7:** *Identify and summarize central ideas in informational texts.*
> **Grade 6: 3.1 Reading G:** No indicator
> **Grade 8: 3.1 Reading G:** No indicator
> **Grade 12: 3.1 Reading G.1:** *Identify, describe, evaluate, and synthesize the central ideas in informational texts.*

In other places, there are inexplicable gaps in content. In the following example, poetry is introduced but is not developed further than grade 8, making it difficult for high school teachers to have a clear sense of what they should teach students about poetry.

> **Grade 4: 3.1 Reading G.11:** *Identify the structures in poetry.*
> **Grade 6: 3.1 Reading G.12:** *Identify and respond to the elements of sound and structure in poetry.*
> **Grade 8: 3.1 Reading G.13:** *Read critically and analyze poetic forms (e.g., ballad, sonnet, couplet).*
> **Grade 12: 3.1 Reading G:** No indicator

Finally, an increase in intellectual demand is not always clear, as in the following example:

> **Grade 4: 3.1 Reading G.1:** *Discuss underlying themes across cultures in various texts.*
> **Grade 6: 3.1 Reading G:** No indicator
> **Grade 8: 3.1 Reading G.5:** *Analyze ideas and themes found in texts.*
> **Grade 8: 3.1 Reading G.12:** *Identify and analyze recurring themes across literary works.*
> **Grade 12: 3.1 Reading G:** No indicator

The second factor that contributes to lack of progression is New Jersey's having taken the approach of adopting cumulative progress indicators. At each grade level (or grade span), the progress indicators are preceded by a statement that says, "Building upon knowledge and skills gained in preceding grades, by the end of Grade "__," students will …." While this may be intended to prevent repetition across grades, the consequence is that some important indicators fade in and out without a readily discernable reason, making it difficult for teachers to know where to focus instruction. In many cases, it is assumed that teachers will know how concepts should build from grade span to grade span, whereas Achieve's reviewers felt strongly that such development should be made explicit.

In large part, this issue will be ameliorated by having grade-by-grade standards in grades 5–8 developed with the same thoughtfulness and care that New Jersey's K–4 reading standards exhibit. To the state's credit, it has already moved in this direction by breaking down its cluster standards for grades 5/6 and 7/8 and drafting standards for each individual grade 5–8. Achieve encourages the state to extend the standards into the high school grades as well so that student expectations at each grade level are explicit. At a minimum, the state should make clear its expectations for its exit exam test.

- **The clarity and measurability of New Jersey's standards in language arts literacy are not yet on par with Achieve's benchmark documents.**

For standards to be accessible to educators and the public, they must be phrased in straightforward language and maintain a consistent level of generality throughout — not vacillate between broad and narrow statements. Although New Jersey's standards generally are jargon-free, there are some instances in which the standards are opaque. Statements such as "students should be helped to understand the recursive nature and shifting perspectives of the writing process" (standard 3.2 descriptive statement) do not communicate clearly. What are the shifting perspectives of the writing process?

Even seemingly straightforward expectations on second look are not entirely clear. Consider Grade 4 Reading A. Concepts about Print/Text (1): *Develop knowledge about various print formats, including newspapers, magazines, and reference resources.* "Knowledge" in this expectation is undefined, although an experienced teacher might readily expand it to a variety of features of newspapers and magazines, as well as reference resources considered appropriate at grade 4. But some teachers may read the indicator and come away with the impression that students are simply expected to know what newspapers and magazines are.

Achieving a consistent "grain size" in the cumulative progress indicators is another area that requires attention. To illustrate, the standards state that 12th-grade students should "Identify, describe, evaluate, and synthesize the central ideas in informational texts" in one standard, and in the very next students are told that they should "understand the study of literature and theories of literary criticism." The first is quite specific, the second almost impossibly general (exactly how would one demonstrate that one understands the "study of literature"?). The dramatically different grain size of these two standards exacerbates the problem of giving teachers practical guidance about what to emphasize.

Measurable standards describe performances that are observable or verifiable in some relatively straightforward way and that are focused on what students should learn, rather than on how they should learn it. New Jersey expectations, however, contain a number of indicators that focus on the kind of experiences students should be having (learning standards), rather than on what they should be learning (content standards). Examples are evident throughout the standards, especially in reading and writing. Consider Reading Grade 4 G (13): *Read regularly in materials appropriate for their independent reading*

level. This statement is more of a curricular goal than a performance expectation. Another example appears in Writing Grade 12 A (1): *Engage in the full writing process by writing daily and for sustained amounts of time*. This indicator is primarily a teaching suggestion since it does not specify what successful performance would look like.

- **Structural flaws in the standards reduce their coherence.**

The overall organization of the Core Curriculum Content Standards into five defining standards is effective at the macro level. However, the substrands do not lend themselves to a careful cross-grade tracing of indicators for several reasons. First of all, the numbering system of the indicators is inconsistent and seems to "bounce around." Indicators having to do with the same content may be listed first in one grade and 10th in another. This finding generally holds true for all the standards; one example from reading will serve to illustrate the general concern. "Authors' point of view" has a different number across the grade level, making it hard to follow.

> **Grade 4: 3.1 Reading G.6:** *Recognize an author's point of view.*
> **Grade 6: 3.1 Reading G.13:** *Respond critically to an author's ideas, views and beliefs***.**
> **Grade 8: 3.1 Reading G.13:** *Compare several authors' perspectives of a historical character, setting, or event.*
> **Grade 12: 3.1 Reading G:** No indicator at grade 12.

In addition, indicators within a single strand do not always follow one another logically, nor are they necessarily conceptually linked to one another, making the set feel incoherent. For example, at grade 3, indicator G.2 calls for students to "distinguish cause/effect, fact/opinion, main idea/supporting details in interpreting texts," while G.14 calls for students to "use information and reasoning to examine bases of hypotheses and opinions." The absence of a consistent framework across grades makes it difficult to trace expectations across the grades.

More important, the strands [A. Concepts about Print/Text; B. Phonological Awareness; C. Decoding and Word Recognition; D. Fluency; E. Reading Strategies (before, during and after reading); F. Vocabulary and Concept Development; G. Comprehension Skills and Response to Text; H. Inquiry and Research] in each section do not function well as a guide for the reader in understanding the structure of the domain.

New Jersey has designated E, Reading Strategies, as its own strand, separate from strand G, Comprehension Skills and Response to Text. Although strategies are typically inappropriate for large-scale assessment, their inclusion in a set of expectations for the state's classrooms presents an important statement, and many of these skills and habits can be assessed locally at the classroom level.

What is problematic, however, is that reading strategies are not linked conceptually in the standards to reading comprehension. It makes little sense to separate these out when they are essential parts of a larger whole.

Reviewers also were concerned with the New Jersey standards' failure to make a distinction between reading informational texts and reading literary texts; rather, the expectations for both are listed (in no apparent order) under Strand G, Comprehension Skills and Response to Text. Comprehension skills are often used in subtly different ways when readers read different genres, and the standards do not make clear how these different sets of skills should be developed. Neither teachers nor students are helped by the assumption that reading is reading and that all comprehension strategies are created equal. In addition, when informational and narrative skills are not clearly detailed, nonfiction materials may get short shrift.

There is a somewhat different structural problem in writing. In Strand A, Writing as a Process, reviewers noted the lack of evolution in the strategies students are expected to use in generating and organizing ideas for writing as compared with the strategies they are to use in planning and producing writing for a specific purpose and audience. The lack of careful development makes it particularly difficult for teachers to understand and plan instruction so that students build an increasing array of strategies for the various stages of the writing process and learn to apply them appropriately as they craft their writing.

RECOMMENDATIONS FOR IMPROVEMENT

- ✓ **Provide more definition to the high school standards and clarify what is required to graduate from high school.**

New Jersey made good use of external experts, as well as state educators, to ensure that its standards in early literacy would reflect exemplary practice and provide strong guidance to teachers at each grade level. Achieve recommends that New Jersey use the same approach to complete its continuum of standards by providing more delineation of high school expectations. One approach is to develop standards that span two years, grades 9–10 and 11–12, as California and Massachusetts do. As it currently stands, New Jersey's end-of-grade-12 standards represent the cumulative knowledge of four years of coursework, obscuring the progression of knowledge and skills expected, and making it hard to determine what students will be responsible for on the graduation exam given in the 11th grade.

- ✓ **Develop a K–12 matrix for the standards in language arts literacy so that the progression of knowledge and skills contained in each strand can be readily traced from one grade to the next.**

To improve the next version of the New Jersey standards, Achieve recommends that the state create a cross-grade matrix that traces each strand of content indicators through the

grades, indicating what new knowledge and abilities are expected at each grade and also where development reaches a level of automaticity (as the present document indicates for phonological awareness by grade 4). Several states, including Massachusetts and Ohio, have included matrices in their standards documents to delineate the progression across the grades. Maryland has used a similar approach in laying out its Voluntary State Curriculum.

A K–12 matrix could serve as an interim step in bringing the components of New Jersey's instructional system into fuller alignment. Achieve sees a number of advantages to this approach since a matrix, by its nature, directs attention to sequencing and specificity.

Having a mechanism to track and adjust indicators would help ensure that:

1. core knowledge and skills are situated in the optimum grade with all prerequisites in place;
2. no significant gaps in core content appear in strands across the grades;
3. content evolves in cognitive complexity from one grade to the next;
4. language is precise enough for teachers to understand the level of performance expected of students;
5. standards are organized as tightly as possible so redundancies are eliminated, priorities are ordered and categories have a consistent grain size; and
6. opportunities for integration across and within strands are made more visible.

This type of matrix could serve a number of broader, complementary functions, as well, by providing supporting detail for the standards and by serving as a guide for the state's next round of test development. It also could be used as a tool for formulating test specifications and developing sample tests.

✓ **Improve the specificity of the standards in language arts literacy.**

New Jersey should provide more direction regarding the tasks students should be able to complete and the level of text students are expected to comprehend. New Jersey could amplify reading indicators to include a description of the kind of texts students should read, as Indiana does, or it could offer a series of sample grade-level reading passages, as Massachusetts does, and/or a sample reading list, as New Standards does. Any of these approaches would help communicate a shared understanding of what an abstract skill, such as comprehension, means at a given grade level and also could bring needed clarity to standards for vocabulary development.

- ✓ **Strengthen the rigor of the standards by including expectations related to literary traditions, literary elements and literary analysis and clarifying grade-level expectations for standard English usage.**

The New Jersey standards embed expectations about students' knowledge of literature into the Reading strand (and to a lesser extent the Writing strand). Consequently, this central area of the reading and language arts domain receives little systematic attention. The expectations neglect content knowledge goals, such as the study of American or other literary traditions, familiarity with the works of particular authors or periods, understanding of author's craft, knowledge of genres, and facility with literary elements and analysis. In addition, standards for grammar, usage, spelling and vocabulary development are not as fully explicated in the New Jersey standards as they are in the benchmark standards, and they deserve more attention. We encourage the state to strengthen these areas in the standards.

MAJOR FINDINGS: LANGUAGE ARTS LITERACY
ALIGNMENT OF ASSESSMENTS TO STANDARDS

Achieve carried out a detailed study of the alignment of New Jersey's assessments in language arts literacy to the state's Core Curriculum Content Standards. Achieve reviewed three assessments:

New Jersey Assessment of Skills and Knowledge (NJ ASK 4) Form A (Spring 2003)
Grade Eight Proficiency Assessment (GEPA) Form S (Spring 2003)
High School Proficiency Assessment (HSPA) Form B (Spring 2003)

It is important to note that of these three assessments, only that for grade 4 is based on the 2002 edition of the New Jersey standards. The state is making the transition to its new standards; consequently, the assessments at grade 8 and high school are based on the previous edition of the standards (1996). Achieve's review of these two tests relative to the 2002 standards will help the state to identify gaps in alignment and ensure that the new assessments are more strongly aligned to the revised standards.

STRUCTURE OF THE ASSESSMENTS

NJ ASK 4 Form A (Spring 2003)

The grade 4 assessment consists of five sections, administered over two days of testing. (One section is a field test, which does not contribute to the students' scores and was not reviewed by Achieve.) On the sections evaluated by Achieve, students read two passages — one narrative and one procedural — and responded to 14 related comprehension items (11 multiple choice and three open-ended). Students also responded to two writing prompts — one a narrative response to a picture prompt, the other a narrative response to a poem. Overall, students responded to a total of 23 items, including field test items. Students received scores on their performance on a total of 16 items, including two writing prompts, three open-ended items and 11 multiple-choice items.

GEPA Form S (Spring 2003)

The grade 8 assessment consists of six sections, administered over two days of testing. (One section is a field test; it does not contribute to the students' scores and was not reviewed by Achieve.) On the sections evaluated by Achieve, students read two passages — one a narrative selection, the other an essay — and responded to 24 related comprehension items (20 multiple choice and four open-ended). Students also responded to two writing prompts — a narrative response to a picture prompt, a persuasive letter and also to a revise-edit task. Overall, students responded to a total of 39 items, including 12 field test items. Students received scores on their performance on a total of 27 items,

including 20 multiple-choice comprehension items, four open-ended items and three writing tasks.

HSPA Form B (Spring 2003)

The high school assessment consists of six sections that are administered over two days of testing. (Two sections are field tests; these do not contribute to the students' scores and were not evaluated by Achieve.) On the sections evaluated by Achieve, students read two passages — one a narrative selection, the other an essay — and responded to 24 related comprehension items (20 multiple choice and 4 open-ended). Students also responded to three writing prompts — a narrative prompt, a persuasive prompt and an editing task that is a field test. Overall, students responded to a total of 39 items, including 12 field test items. Students received scores on their performance on a total of 26 items, including 20 multiple-choice comprehension items, four open-ended comprehension items and two writing tasks.

STRENGTHS OF THE ASSESSMENTS

- **On all three assessments, New Jersey makes use of full-length, authentic passages that resemble materials that students read in and out of class.**

One factor that contributes to the quality and rigor of a test is the level and complexity of the texts students are asked to read. Drawing assessment selections from "classical" children's texts or from worthwhile, widely read contemporary literary or informational text, as opposed to using "test text" — passages constructed expressly for the purpose of testing — is an exemplary practice. First, if the goal is to determine whether students can read and comprehend the kind of materials they are accountable for in the classroom, students should be tested on text similar to what they typically encounter and which requires an equivalent level of thought and depth of processing. Second, using texts that are similar to classroom materials makes it possible for test developers to construct the same kind of questions good teachers ask in rich classroom discussions. When the text and related questions in a testing situation are authentic, they send the very message New Jersey would hope for — the best way to prepare for state tests is to read and analyze quality literature and informational text.

- **New Jersey's assessments include open-ended items.**

The inclusion of open-ended items is one hallmark of quality assessments. Open-ended items are essential because of their potential to assess standards that demand advanced, analytical thinking, such as those that require students to make inferences, synthesize and evaluate. The closer assessments come to the real demands of postsecondary education and the workplace, the more reliable they will be in determining whether students are prepared for their next career steps.

- **New Jersey's assessments measure writing directly.**

The state's decision to include on-demand writing is a good one. Asking students to compose a piece in response to a well-structured directive (prompt) is a reasonable way to evaluate — on a large-scale, on-demand test — how well students can respond to the totality of a writing task — identifying the audience and purpose, organizing the piece, and attending to grammar, syntax and punctuation. Getting a realistic picture of a student's writing proficiency is a benefit of requiring direct writing in and of itself, and it also has the important effect of encouraging teachers to dedicate classroom time to helping students develop strong composition skills.

AREAS FOR IMPROVEMENT

- **The alignment of New Jersey's language arts tests and standards should be strengthened.**

As stated in the beginning of this review, only the grade 4 test is based directly on the 2002 edition of the New Jersey standards. The state is making the transition to its new standards; consequently, the assessments at grade 8 and high school are based on the previous edition of the standards (1996). Nonetheless, the items on all three New Jersey's assessments pass the first test of alignment: nearly every item on the tests assesses knowledge and skills found in the standards.

However, all of the assessments — including that for grade 4 — are less successful in meeting the second test of alignment: effectively sampling the breadth and depth of the standards. In fact, some content and skills are overassessed, while other important areas are not assessed at all. In grade 8, for example, two to three items within a passage set target the same indicator, at the expense of leaving other critical indicators unassessed.

Moreover, in cases where a progress indicator includes two or more cognitive skills, reviewers found New Jersey's items tend to measure the least demanding component of the indicator. This tendency negatively affects both the balance of the test and its level of challenge. The state will want to pay careful attention to coverage of the entire indicator by developing different items to assess each important skill and by verifying that items are doing more than measuring the lower of the cognitive skills.

It should be kept in mind that, to some extent, alignment of tests to standards is dependent on the structure of the standards. Due to the cumulative nature of New Jersey's standards, as well as to some of the problems with progression and clarity mentioned previously, reviewers identified a number of items (especially in the HSPA) that are aligned to standards at lower grades. In several instances, these items are reasonable questions, worth asking of students at the particular grade level, even though they draw on a previous grade's skills. In other cases, however, the items are at too low a level of cognitive demand for the grade level assessed, and they diminish the rigor of the tests.

- **New Jersey's assessments include two passages and contain relatively few items that measure reading comprehension.**

New Jersey has chosen to use comparatively few reading selections (that is, two scored reading passages), which are longer than is characteristic of most state tests. (Note: Students respond to three reading passages, but one is a field test and does not contribute to their scores.)

New Jersey's decision to limit its assessments to two scored passages makes passage selection an extremely critical element of test design because fully one-half of students' performance in reading comprehension is based on their understanding of a single passage. This exacerbates the problem of "passage effects," which are fluctuations in student performance that are the result of the content, topic or characteristics of the selection rather than the students' abilities. This problem could be mitigated by adding passages to each assessment.

Although the authenticity of the passages is a positive attribute of New Jersey's tests, their length may make them more of an endurance test for elementary and middle school students than the state may have intended. A student may end up spending a disproportionate amount of time reading rather than analyzing text. In comparison, a review of three other state and national tests of reading comprehension showed that they used four and five shorter selections as opposed to New Jersey's two, the longest of these other states' passages being approximately one-half to two-thirds the length of the text selections on the New Jersey test. Moreover, the current draft of the NAEP Frameworks in reading contains the following guidelines as to passage length that the state will want to consider: grade 4 (200–800 words); grade 8 (400—1,000 words); grade 12 (500—1,500) words. In contrast, Achieve estimates that the narrative passages on the New Jersey tests it analyzed for this report are of the following lengths: grade 4 (2,000 words); grade 8 (1,800 words); grade 12 (2,750 words).

Regardless of passage length, however, New Jersey may wish to consider increasing the number of items that measure reading comprehension. At grade 4, there are only 14 items (11 multiple choice and three open-ended); at grades 8 and 11, there are 24 items (20 multiple choice and four open-ended). In contrast, many reading tests currently in use include 35 to 50 items. Students in other states tend to be reading more passages and responding to more items. Additional passages and items allow for more thorough coverage of the standards and yield a more reliable indication of students' skills and performance related to the standards.

- **New Jersey's assessments of reading comprehension are uneven in terms of cognitive demand.**

The level of difficulty of a reading comprehension test stems from two main sources: the difficulty and complexity of the text selections and the focus and complexity of the items related to the selections — what students are asked to do in relation to the text. At each

grade level, students read three passages and respond to related items. However, only two passages and associated items count toward a student's score; the third is a field test. While reviewers took issue with aspects of the selections, they agreed that all the selections were grade-appropriate. However, many states using three or more selections on tests include a passage that students reading at grade level would find challenging. Only on the HSPA does New Jersey include a passage that would challenge students who are just meeting standards.

Even though none of the passages on New Jersey's assessments at grades 4 and 8 is especially challenging, the passages are sufficiently robust to support items that tap advanced thinking skills, such as student's abilities to interpret, analyze, synthesize and evaluate — performances called for by New Jersey's standards. The overall rigor of the state's assessments of reading comprehension is lower than the standards suggest because of the interplay of the following factors: First, the test questions do not always take full advantage of the opportunities the passages offer to probe student's thinking. Also, items tend to be pitched at lower levels of cognitive demand, seldom requiring students to do more than demonstrate low-level skills. Moreover, items on all three tests failed to assess a number of reading indicators that demand in-depth thinking. Reviewers also noted that two of 14 total comprehension items on the grade 4 test could be answered on the basis of sidebar features without students having to read the related text, thereby diminishing the level of rigor. (Typically, charts and tables extend the text, and items that rely on them also rely on an understanding of the selection.)

In addition, the level of rigor would have been raised had the state taken full advantage of its open-ended items to assess knowledge and skills delineated by the standards but not readily assessed with items in a multiple-choice format. For example, on the grade 4 assessment, one of the open-ended items could have been used to ask students to make inferences across two selections or apply or extend what they read.

Finally, the number of items that measure standards at lower grades is another example of the way in which the level of rigor of the tests is reduced. This is due in part to the cumulative nature of New Jersey's standards and in part to the fact that important knowledge and skills are not evenly developed across the grades as described by the cumulative progress indicators. Including some items that map to standards at a lower grade than the grade level of a given test is not a problem where a state has cumulative standards, provided such items are not disproportionate in number and are worthy of being asked at any grade level**.** In New Jersey's case, reviewers noted that in some instances items mapped to lower grade levels were in fact appropriate for assessment, but in others the items were too low a level of cognitive demand for the grade in question. The inclusion of the latter class of items diminishes the rigor of the tests.

- **The writing prompts do not always provide enough guidance to students.**

The writing prompts present students with the opportunity to address a number of the writing standards in their responses. However, the prompts are not always clear and do

not always provide students with sufficient guidance for the task to be completed. At times, they may hinder rather than help students demonstrate their ability to meet the standard, simply because the expectations are not precise. Directions, prompts and writers' checklists should work in concert, helping students focus their writing on the demonstration of their achievement of specific standards. Unfortunately, the state's writer's checklists are often too general and do not always help students demonstrate that they meet the standards.

As a final note, reviewers noted that New Jersey's choices of genres for prompts do not always reflect the appropriate goals, particularly on the 8th grade and high school tests. In particular, they suggested replacing the personal letter prompt at 8th grade with an informational writing prompt to prepare students more effectively for high school demands. They also suggested replacing the narrative prompt at high school with an informational prompt, as students are far more likely to encounter the need for this kind of writing in their lives after high school.

RECOMMENDATIONS FOR IMPROVING THE LANGUAGE ARTS LITERACY ASSESSMENTS

Reading

- ✓ **Raise the level of demand of all three assessments, particularly the HSPA.**

On all three tests, too few items demand more than drawing simple inferences. All of the passages would have supported more cognitively challenging items that asked students to engage in critical thinking and reflection about what they read. If the state wants to challenge all students, it will have to add questions with a higher level of demand. Moving forward, the state would be wise to require test developers to take fuller advantage of the selected texts, to assess the important and challenging content and tap into higher-level cognitive skills. The state should take care that items do not overassess a given standard to the exclusion of other equally or more significant standards and that they are not systematically targeting the easiest component in a compound standard — a practice that erodes rigor.

To raise the level of rigor on HSPA, the state should increase the number of items that ask students to make broader inferences and evaluate, extend and apply what they read — skills that they will need to be successful in postsecondary education and in an information-based economy.

- ✓ **Reduce the length of the passages on the grade 4 and 8 assessments and increase the number of selections on all three assessments without increasing testing time.**

Increasing the number of texts and range of genres assessed will reduce the potential risk of passage effects suppressing students' performance. At the same time, a broader range of passages and genre can create the opportunity to ask a wider variety of questions that

will allow for more thorough coverage of the standards. (Ideally, informational texts should make up approximately half of the test.) Ensure that the test developers of all three tests analyze each text in terms of what opportunities it presents relative to the New Jersey Core Curriculum Content Standards. Such "mining" of passages should identify vocabulary, literary devices, text structures, and any other topics that reflect the standards and are too important to be left unassessed.

- ✓ **Add items to each assessment, without increasing testing time.**

Reducing the length of texts would allow students time to respond to more items. It would be reasonable to ask eight to 12 items associated with each of the recommended four passages. These changes would allow for more thorough coverage of the standards and would yield a more reliable indication of students' abilities related to the standards.

Writing

- ✓ **Rework the writing prompts to make absolutely clear what genre is being assessed and what the expectations are for a strong piece of writing.**

Language in a prompt should be clear, parsimonious and consistent. After students read a prompt, they should have a clear sense of at least three parameters — the genre, the topic and the characteristics of a piece of writing that will meet the standards. A good addition to the prompts would be the inclusion of a statement that calls attention to the need for an opening and closing. Achieve also recommends that the state emphasize the kinds of writing students will be confronted with in postsecondary education and the workplace.

Test Blueprint

- ✓ **Create a transparent test blueprint that links the assessments to the standards. Consider having the test blueprint specify which standards should be assessed using the open-ended items.**

Achieve advises New Jersey to develop a test blueprint that is clear, specific and readily accessible. The blueprint should specify the indicators that should be assessed at each grade level, guiding test developers to design items and prompts for these very explicit expectations. In some cases, open-ended items on the present tests do not take full advantage of the opportunity to assess students' knowledge in areas that require advanced, analytical thinking. A well-crafted blueprint could solve this problem by specifying which indicators need to be assessed by open-ended items and which are assessed readily by multiple-choice items.

MAJOR FINDINGS: MATHEMATICS
CORE CURRICULUM CONTENT STANDARDS

The New Jersey Core Curriculum Content Standards for mathematics are comprised of five standards — Number and Numerical Operations; Geometry and Measurement; Patterns and Algebra; Data Analysis, Probability and Discrete Mathematics; and Mathematical Processes. Each standard is further broken out by strands (lettered) and learning expectations (numbered) (also called "cumulative progress indicators") for individual grades 2, 3, 4, 5, 6, 7, 8 and 12. Achieve first reviewed the 1996 edition of New Jersey's mathematics content standards in 2000. Our second review provides us with an opportunity both to acknowledge substantial progress and to suggest ways in which the state can continue to improve its standards.

COMPARISON OF NEW JERSEY'S 2002 STANDARDS FOR MATHEMATICS WITH ITS 1996 STANDARDS

Since Achieve conducted its initial review of New Jersey's Core Curriculum Content Standards for mathematics in 2000, New Jersey has made a significant number of revisions, which has resulted in a much stronger standards document as the chart below attests:

ACHIEVE CRITERIA	NEW JERSEY MATHEMATICS STANDARDS	
	2000 Review	2003 Review
• **RIGOR**		
— Intellectual demand	⊙	●
— Essential knowledge & skills	⊙	◎
— Balance	⊙	⊙
• **FOCUS & COHERENCE**		
— Emphasis on core content	⊙	◎
— Manageability	⊙	⊙
— Unified vision of discipline	⊙	◎
— Connections	○	⊙
• **PROGRESSION**		
— Placement of content & skills	○	●
— Development of content & skills	○	◎
— Grade-by-grade expectations	○	⊙
• **SPECIFICITY**		
— Precision	○	⊙
— Sample problems	○	○
• **CLARITY**		
— Language	⊙	⊙
— Organization & format	⊙	●
— Use of technology	○	◎
• **MEASURABILITY**	⊙	⊙

Key:
○ Needs Significant Improvement
⊙ Needs Some Improvement
◎ Approaching Benchmark Quality
● Meets Benchmark Standards

As is evident from the chart, New Jersey has made substantial progress in advancing the quality of its standards for mathematics. A comparison of the revised edition of the standards with its predecessor demonstrates positive growth in intellectual demand, inclusion of essential knowledge and skills, emphasis on core content, making connections among the principal domains of mathematics, specificity, measurability, organization, and the use of technology. In reducing the number of content standards in mathematics from 16 to five, New Jersey also improved the focus, coherence and utility of the standards — especially for elementary school teachers, who are primarily generalists. Even more impressive is the improvement in progression. The state has generally succeeded in depicting the evolution of knowledge and skills on a strand-by-strand basis across grades 2–8.

Developing standards for grades K and 1 and grades 9, 10 and 11 is an important next step and will complete the K–12 continuum. Adding sample problems is also critical: These will improve both the clarity and specificity of the standards and make the state's expectations for rigor transparent.

COMPARING NEW JERSEY'S 2003 STANDARDS WITH THE BENCHMARK STANDARDS

All standards documents, including Achieve's benchmark standards, are based on hard choices. In the end, states have to be concerned that they have maximized the time devoted to learning important core knowledge and that the approach they have taken to developing standards across grades is internally consistent. When we compare New Jersey's revised standards with the benchmark documents, for the most part they compare favorably, with the possible exception of grade 12, where the lack of intermediary standards at grades 9–11 and low level of precision make it difficult to judge rigor. Comparisons of New Jersey's standards at grades 4, 8 and 12 with the benchmark documents reveal some key differences that New Jersey will want to take note of when it undertakes its next revision of the standards.

GRADE 4 SUMMARY BY STRAND

At grade 4, New Jersey is on par with both Indiana and Singapore with respect to number sense, numerical operations and estimation with whole numbers; and New Jersey also is generally on par with Indiana in fractions and decimals, although Indiana also includes writing mixed numbers and converting them to improper fractions. Singapore expects a bit more in operations with fractions (specifically addition of fractions and multiplication of a fraction by a whole number).

New Jersey includes appropriate exploration of contexts that give rise to negative numbers (temperatures below 0º and debts) — a topic not addressed by either Indiana or Singapore. The match between New Jersey's and Indiana's standards for Patterns and Algebra is strong, and New Jersey is somewhat ahead of Singapore in the attention it gives to patterns, functions and modeling.

The most striking differences between New Jersey and the benchmark documents occur in New Jersey's treatment of Geometry and Measurement and Data Analysis, Probability and Discrete Mathematics. In geometry, New Jersey has higher expectations for student proficiency regarding 3D shapes — including cube, rectangular prism, sphere, cone, cylinder and pyramid — and transformations — tessellations, slides, flips and turns. Similarly, New Jersey is significantly ahead of Singapore and Indiana in expecting students to represent probability as a fraction, explore tree diagrams and study vertex edge graphs. Students also must investigate mean, median and mode, although learning the definitions of these terms does not seem to be required — an apparent oversight that should be addressed.

GRADE 8 SUMMARY BY STRAND

The standards for Number and Numerical Operations as described by New Jersey, Singapore, Indiana and *Foundations for Success* are roughly equivalent. All emphasize the crucial distinctions between rational and irrational numbers.

As was true of grade 4, New Jersey's overall expectations for Geometry and Measurement, as well as for Data Analysis, Probability and Discrete Mathematics, in grade 8 continue to exceed those of Singapore and Indiana. The margin is considerable relative to Singapore and smaller relative to Indiana. All three benchmark documents (Indiana, Singapore and *Foundations for Success*) expect students to have an understanding of the Pythagorean Theorem and to apply it in different situations. This expectation is higher than that set forth in the New Jersey standards relative to the Pythagorean Theorem. However, none of the benchmark standards develop understanding of fractals and self-similarity, as does New Jersey.

New Jersey and the benchmark documents treat linear patterns, linear functions and modeling with linear functions similarly, *except* for the concept of slope, which is discussed below. However, New Jersey gives much more emphasis to non-linear patterns and recursive patterns prior to grade 8 than do Singapore and Indiana. In its treatment of non-linear patterns and functions in grade 8, New Jersey emphasizes geometric sequences and exponential functions. In contrast, Indiana and *Foundations for Success* pay more attention to quadratic functions. This difference in emphasis has consequences for the development of algebraic skills.

The New Jersey standards are not as rigorous as those of the benchmark documents with respect to algebraic procedures; the state expects less in the types of manipulations it requires and the complexity of the equations it considers. Singapore is clearly ahead in calling for procedures such as expansion of binomials and factoring the difference of two squares, and *Foundations for Success* has even greater expectations for grade 8 students. New Jersey postpones the above algebraic manipulations until grades 9–12. In addition, both Indiana and Singapore are much clearer as to the linear equations (and linear inequalities) students must solve. For example, in grade 7, they specify two-step linear equations and simple literal equations: "Solve the equation $4x - 7 = 12$ and check your

answer …" and "Solve the formula $C = 2\pi r$ for r." In sum, the benchmark documents either demand more than New Jersey regarding algebraic procedures or at least are more precise in laying out their expectations.

GRADE 12 SUMMARY BY STRAND

New Jersey's standards for Number and Numerical Operations, Geometry and Measurement, and Patterns and Algebra cover at least two-thirds of the expectations of the Indiana Algebra I-Geometry-Algebra II course sequence. One can only make a rough judgment because the New Jersey standards are not as sharply drawn and specific as those of Indiana and are not broken down course by course. In contrast to the rest of the New Jersey standards, where it appeared that procedural knowledge was emphasized at the expense of conceptual knowledge, in algebra the New Jersey standards tend to stress key concepts rather than procedures.

New Jersey standards omit some major topics that Indiana includes in its Algebra I course: (1) completing the square to solve quadratic equations; (2) deriving the quadratic formula by completing the square; (3) solving a 2□2 linear system (that is, a pair of linear equations in two variables); and (4) simplifying algebraic fractions and solving algebraic proportions (for example, $(x + 5)/4 = (3x + 5)/7$).

In comparison with Indiana's Algebra II course, New Jersey standards omit the following topics: (1) complex numbers, (2) solving polynomial equations with complex numbers, (3) conic sections, (4) operations with algebraic fractions, and (5) logarithmic functions.

Indiana's expectations for geometry generally exceed those of New Jersey's in the following areas: (1) the use of coordinate geometry to deduce geometric properties such as congruence, similarity and symmetry of polygons; (2) attention to geometric constructions with straight edge and compass; and (3) more insistence on students' being able to prove *specific* facts and demonstrate *specific* forms of reasoning. Both Indiana and *Foundations for Success* expect students to prove the Pythagorean Theorem in grade 8. Indiana further requires that students: "state, use, and examine the validity of the converse, inverse, and contra-positive of 'if-then' statements," and "use and compare a variety of ways to present deductive proofs, such as flow charts, paragraphs, and two-column and indirect proofs." The New Jersey standards do not make formal mathematical reasoning requirements so explicit.

As a result of these grade-by-grade comparisons, reviewers identified overall strengths and areas for improvement that Achieve believes will be helpful to the state.

STRENGTHS OF THE CONTENT STANDARDS

- **New Jersey has adopted an ambitious, rigorous set of standards that will steer mathematics education in a somewhat different direction than that taken by most states.**

New Jersey's revised standards go further than those of any other state Achieve has reviewed in the areas of data analysis, probability and discrete mathematics. They expect students to learn elements of contemporary applied mathematics such as statistics, combinatorics, graphs and algorithms at levels that are unprecedented in state K–12 standards.

The approach taken by New Jersey has considerable merit. It introduces significant mathematics of interest to many students and prepares students for advanced study in an increasing number of fields and for employment in such areas as computer science, operations research and management science. Equally important, it develops the critical mathematical skills of reasoning and abstraction. Moreover, taken together, the expectations in New Jersey's Data Analysis, Probability, and Discrete Mathematics standards give students the mathematical background for decision making in economic, political and social contexts — tools for informed citizenship.

New Jersey is unequivocal in its intent and its vision: The introduction to the Core Curriculum Content Standards for Mathematics makes it clear the standards are intended for *all* students and meant to confront today's reality of an increasing number of occupations requiring knowledge and skills in new areas — data analysis, problem-solving, pattern recognition, statistics and probability. Thus the fact that these areas receive more attention in New Jersey's standards than in most states is not accidental; it is purposeful and reflects the state's vision for contemporary mathematical standards.

Pioneering is not without its perils, however, and Achieve offers some cautions later on in our report that New Jersey will want to consider.

The cumulative progress indicators develop a thoughtful progression of knowledge and skills from grade 2 to grade 8, with key content introduced at appropriate points in the progression of the standards.

Progression of knowledge and skills in New Jersey's standards for mathematics across grades 2–8 is strong; the sequencing of topics generally reflects a carefully staged evolution of knowledge and skills based on the logical structure of mathematics as a discipline. (Since the expectations for grades 9–12 are not broken down by grade or by course, it is not possible to determine whether they are sequenced appropriately.)

Two examples underscore the power of articulating standards with such care. The first is the way in which the New Jersey standards develop the skill of writing formulas for the *nth* term of arithmetic and geometric sequences called for at grade 12. The study of arithmetic and geometric sequences is under way by grade 4, where students must be able to "recognize and continue various natural number patterns." By grade 6, students must "write formal iterative formulas" for patterns such as geometric sequences. In grade 8, students must "relate such number patterns to iterative geometric patterns and be able to write a formula for the general term" of certain special examples. This culminates in grade 12, when students must be able to "write a formula for the *nth* term of any

geometric sequence, and use that formula to find a formula for the sum of the terms in that sequence." Strand D, Vertex-Edge Graphs and Algorithms, in the data analysis standard, provides a second illustration of how concepts should be incrementally developed. Beginning with just the elements of a graph (vertices and edges) in grades 2 and 3, it advances to fairly elaborate problems and structures by grades 8 and 12.

New Jersey generally handles two other aspects of progression well: (1) most content and skills are introduced at appropriate points (as comparisons with the benchmark documents at grades 4 and 8 demonstrated), although some duplication exists from grade to grade; and (2) essential prerequisites are usually included in the standards and introduced in a timely manner (the few exceptions are detailed in Achieve's technical report).

- **New Jersey's standards in mathematics are clear in defining the topics to be learned and are mathematically accurate.**

New Jersey standards are written in straightforward, jargon-free prose and, with a few exceptions that can be readily remedied, are mathematically correct. In comparison with the benchmark documents, the New Jersey standards are clearer than those of Singapore (since they include more explanation) but not as clear as those of Indiana (which include both examples and definitions).

Although New Jersey's standards in mathematics can still benefit from more attention to clarity, the state has come a long way in improving this aspect of its standards since Achieve's 2000 review. Consider the following comparison of grade 4 standards for measuring perimeter and area as contained in the New Jersey 2002 edition contrasted with its previous 1996 edition.

NEW JERSEY 1996	NEW JERSEY 2002
4.5.4. Use a variety of tools to measure mathematical and physical objects in the world around them. 4.7.9. Understand the variety of ways in which geometric shapes and objects can be measured.	4.3.D.2 Select and use appropriate standard units of measure and measurement tools to solve real life problems ▪ Area—square inch, square centimeter 4.3.E.1 Determine the area of simple two-dimensional shapes on a square grid 4.3.E.2 Determine the perimeter of simple shapes by measuring all of the sides.

New Jersey also has succeeded in not making the standards more confusing by the unnecessary use of jargon. That said, the lack of definitions and examples in the state's standards tend to make the document fully accessible only to those who are very mathematically literate. All of the standards, other than Number and Numerical Operations, contain terminology that requires definition or illustration. As a case in point,

while the Patterns and Algebra standard defines arithmetic sequence and geometric sequence, it fails to illustrate other commonly used terms such as linear function, linear equation, quadratic equation, polynomial, binomial and monomial. It would make sense for New Jersey to add a glossary and/or examples in text, both of which Indiana K–8 standards provide.

Clarity with respect to the use of calculators is an important criterion in evaluating mathematics standards, and the New Jersey standards do a good job in this area. As student testing becomes increasingly high-stakes, and as society depends more and more on automated numerical calculations, clarity about school expectations for calculation becomes increasingly important. Expectations for numerical calculation by hand (on paper), with calculators (or computers) and through mental methods (both exact and approximate) should be forthright. In addition to having a standard dedicated to technology, New Jersey has generally specified the expectations for which the use of technology is appropriate and/or expected.

Finally, New Jersey's indicators are mathematically accurate, unlike a number of other state standards Achieve has reviewed. (The few instances where revisions are called for are identified in Achieve's technical report.)

AREAS FOR IMPROVEMENT

- **New Jersey's lack of explicit course or grade-level standards for grades 9, 10 and 11 and clear expectations for its grade 11 exit assessment (HSPA) are problematic.**

The lack of grade-by-grade standards in grades 9, 10 and 11, coupled with the lack of definition around expectations for the high school assessment, is a cause for concern. Achieve's experience has been that without clear guidance as to core content and the level of rigor expected, teachers may well end up offering very different levels of preparation to students across the state. Other states have brought definition to the high school program, while remaining sensitive to issues of local control. Maryland has published state expectations for basic algebra, data and geometry to provide guidance for preparing students for the state's high school tests. Massachusetts has published high school standards for grades 9/10 and 11/12 and also course-by-course standards. Districts are free to use either as a guideline, allowing them, for example, to offer a program of integrated mathematics rather than the traditional algebra, geometry sequence.

All the mathematics expectations listed for the end of grade 12 will not be completed by the time students take New Jersey's exit exam in grade 11. Therefore, it is a matter of some urgency for the state to prioritize these expectations and identify those expectations that are essential for the 11th-grade assessment.

- **The totality of New Jersey's standards may not be manageable for all students and their teachers.**

As mentioned earlier, New Jersey is breaking new ground by expecting students to reach high levels of proficiency in data analysis, probability and discrete mathematics by the time they graduate. Achieve's reviewers cautioned that the standards may be too ambitious given the amount of time devoted to math instruction and the current state of preparation of the teaching force.

Achieve recommends that New Jersey establish a complete continuum of K–12 standards, by adding standards in K, 1, 9, 10 and 11 and by combing through the current standards to ensure that knowledge and skills are placed at the optimum grades and progress without gaps in significant content. New Jersey may wish to consider developing a matrix similar to the one Achieve recommended that the state develop in language arts literacy.

Based on comparison to the Indiana standards, the New Jersey standards seem to require more than three years of mathematics after grade 8 for completion. Reviewers found that New Jersey's grade 12 standards for number, Geometry and Algebra strands embrace a good deal (approximately 70 percent) of Indiana's standards for Algebra I, geometry, and Algebra II. In fact, there are expectations in New Jersey's algebra standard (for example, summing an infinite geometric series in grade 12) that are only articulated in Indiana's precalculus course. These additional expectations in the algebra standard combined with those in the data analysis standard appear to require at least a year of mathematics, embracing much of Indiana's discrete mathematics and probability-statistics courses.

There is a tension between establishing a sufficiently rigorous mathematics education that ensures quantitative literacy for all students and one that provides the necessary requirements for students to pursue continuing education and careers in fields in which a very high level of mathematics proficiency is required.

At this point in time, most states require students to meet articulated standards of Algebra I and geometry, and sometimes Algebra II. Since New Jersey is endeavoring to break new ground with standards that go beyond what most states now expect in data analysis, probability and discrete mathematics, the state will want to ensure that preparatory work in geometry and algebra is sufficiently strong for students to be successful in traditional college mathematics and also to put supporting structures in place.

Perhaps the most challenging aspect of these standards will be their implementation in the classroom. It is widely reported that many teachers are unprepared to do justice to the challenging mathematics standards states are adopting, even when their focus is on the traditional curriculum of algebra and geometry. The fact that a good many of the topics in the Data, Probability and Discrete Mathematics section of the New Jersey standards are likely to be unfamiliar to the majority of teachers, especially veteran teachers, exacerbates the challenge of strengthening teachers' knowledge base. For New Jersey to be successful in modernizing its mathematics program of study, it will have to put an aligned, aggressive program of professional development in place. To its credit, New Jersey has made a concerted effort to advance the use of technology in its districts and

will want to take advantage of this capacity to provide increased support for teachers and students.

- **The cumulative progress indicators are not always precise enough to convey the intended nature or level of work expected.**

For the most part, the statements in the New Jersey standards do not spell out the level of understanding expected from students. Lack of precision in standards undermines rigor. When the intended nature and level of work expected of students is ambiguous, it is difficult for teachers to know where to peg the level of rigor in their instruction and for test developers to construct assessment items that accurately measure students' progress toward the standards.

Indiana's standards tend to have a smaller grain size than do New Jersey's and, as a result, greater specificity. For example, consider the differing approach of the two states with respect to a core concept — slope.

NEW JERSEY 2002	INDIANA 2000
4.3.B Functions & Relationships 1. Graph functions, and understand and describe their general behavior • Equations involving two variables • Rates of change (informal notion of slope)	Algebra & Functions 7.3.6 Define slope as vertical rate of change per unit of horizontal change and recognize that a straight line has constant slope or rate of change *Example: Examine a table of values and make a conjecture about whether the table represents a linear function* 7.3.7 Find the slope of a line from its graph *Example: Draw the graph of* y = 2x – 1. *Choose two points on the graph and divide the change in* y-*value by the change in* x-*value. Repeat this for other pairs of points on the graph. What do you notice?* 7.3.8 Draw the graph of a line given the slope and one point on the line, or two points on the line *Example: Draw the graph of the equation with slope of 3 and passing through the point with coordinates (0,–2)* 8.3.5 Identify and graph linear functions and identify lines with positive and negative slope *Example: Draw the graphs of* y = 2x – 1, y = 3x – 1, y = –3x – 1. *Find the slope of each graph. What do you notice?* 8.3.6 Find the slope of a linear equation given the equation and write the equation of a line given the slope and any point on the line. *Example: Write the equation of the line with slope 2 and* y-*intercept –4.*

Compared to Indiana, New Jersey's expectations regarding the slope of a straight line are somewhat vague and obscure. New Jersey's expectations up to grade 8 regarding the slope of a straight line are extremely general, and while New Jersey's grade 8 standard refers to an "informal" notion of slope, the formal notion is missing. It is only required after grade 8 and then only *implicitly* in the grade 12 algebra standard. In contrast, Indiana begins its explicit treatment of slope in grade 7 and develops it further in grade 8.

The following are additional examples where greater attention to precision is warranted.

1. The reference to the "four basic arithmetic operations" (Number and Numerical Operations standard 4.1.B.1) in grade 4 should be clarified. It is not clear whether the operations are meant to apply only to whole numbers in light of the related boldfaced comment.

2. The complexity of the linear equations that students should be able to solve in grade 8 should be spelled out with a sharp distinction drawn between the linear equations solved algebraically (Patterns and Algebra standard 4.3.D.2) in grade 8 and those solved under the same standard in grade 12.

3. There is some confusion in grade 8 (Number and Numerical Operations standard 4.1.A.1): What does "construct meanings" mean? Does it mean to invent a situation that is represented by a particular number? Does it mean to translate into another representation? Does it mean to develop analogies that bring large or small numbers to human scale? Clarification is called for here.

4. In grade 12 (Geometry and Measurement standard 4.2.4): Which conjectures or theorems? Just as the standard in Number that names the kinds of numbers (two-digit, … rational, … real) so should the standard on Geometry be clear about a core of triangle and circle theorems that are expected.

5. In grade 12 (Patterns and Algebra standard 4.3.B.1): Which "relations and functions" are being referred to? Basic functions should be named to make clear the scope and depth of the algebra expectation — and are they expected to be used in 4.3.C.3?

6. Some statements of the standards related to probability at grade 12 are indistinguishable from grade 8. For example: grade 8 (standard 4.4.B.6) reads, "Play and analyze probability-based games, and discuss the concepts of fairness and expected value." And, grade 12 (standard 4.4.B.1) reads, "calculate the expected value of a probability-based game, given the probabilities and payoffs of the various outcomes, and determine whether the game is fair." The subtle difference of calculating the expected value needs more clarification.

- **New Jersey standards omit some significant content.**

New Jersey has developed a comprehensive, challenging set of standards. Consequently, dealing with missing content cannot be an additive process. Rather, we advise New Jersey to consider making judicious substitutions, eliminating redundancies and linking topics across standards.

The standards at grade 4 are quite comprehensive — perhaps overly so. As it now stands, the totality of the grade 4 content may not be manageable. New Jersey may wish to even out the content load by postponing topics that could be addressed later without weakening students' foundational knowledge.

The missing content at grade 8 includes proportionality — a key concept that is not explicitly dealt with at that grade level. As noted previously, slope deserves a more thorough treatment. Direct variation is introduced rather late, in grade 12. By delaying slope and direct variation, the New Jersey standards miss a marvelous opportunity in grade 8 to relate the concepts of scale factor and constant of proportionality (that are so central in the application of similarity relationships) to direct variation and slope. Such integration is a major recommendation of *Foundations for Success* and would bring greater focus and coherence to grade 8's standards.

To make room for these concepts, New Jersey might wish to consider delaying the study of content not found in the benchmarks, such as self-similarity, fractals and vertex-edge graph algorithms. Also, without limitations set, the grade 8 objectives in Data Analysis, Probability and Discrete Mathematics may be overly ambitious, especially the strand dealing with discrete mathematics. For example, in grade 8 (standard 4.4.C.3), which techniques — and in which contexts — are students expected to know and understand? The expectation, "apply techniques of listing, counting, and reasoning in a variety of different contexts," covers a lot of territory.

There are three areas that New Jersey may wish to revisit in its secondary standards. While Number generally receives less attention in secondary grades than in K–6, some topics — an introduction to complex numbers, weighted averages, index numbers and the subtleties of round-off — deserve further consideration, in part, because of their importance in civic life. The state also may want to give some more attention to Euclidian geometry, notably: classic straightedge and compass constructions; the role of Euclidean geometry in the emergence of deduction as a worldwide standard for rigorous thinking; and an explicit mention of the parallel postulate and its role in the emergence of non-Euclidian geometry. The Algebra strand would be enhanced by making its historical connections with geometry explicit. The algebraic solution of 2×2 linear systems of equations appears to be missing, unless it is implicit in grade 12 — only the graphical technique is explicitly mentioned. In any case, solution of 2×2 linear systems of equations should be addressed prior to grade 12. The binomial theorem, already implicit in the New Jersey standards, should be made explicit for the sake of completeness and

coherence. It is inconsistent to omit the theorem and yet introduce Pascal's triangle in grade 6, introduce combinations as early as grade 8 and hint at the theorem in grade 12.

Even within the strand of Data Analysis some topics may be considered for substitution. These include: two-way tables and Simpson's disaggregation paradox, false positives, distinction between correlation and causation, the concept of "random" and its role in making inferences, observational versus controlled studies, and using spreadsheets effectively and with understanding.

Reviewers also suggested the state might wish to consider making the following trade: Understanding the ubiquity of exponential models in real world phenomena, such as compound interest, seemed more important to reviewers than summing finite and infinite geometric series as called for in grade 12. Matrices are another example of content that could be pruned in favor of more essential content.

New Jersey also should consider streamlining the standards from grade level to grade level, taking an especially hard look at grades 4 and 8. Even though an attempt has been made (as stated in the introduction of the standards document) to avoid redundancy, there are many topics that are repeated from grade level to grade level. This not only makes it hard to know when a student is expected to have facility with a topic, but the repetition also makes the content required at each grade level difficult to manage. Furthermore, the repetition of content limits opportunities to include critical missing content, the absence of which diminishes the overall balance and rigor of the content standards.

The following examples illustrate the concern. The statement: "Demonstrate an understanding of place value concepts" appears in grades 2, 3 and 4 without clear illustrations of how the expectations about place value concepts should progress from grade level to grade level. Grade 5–6 and 7–8 include exactly the same indicator, "apply techniques of systematic listing, counting and reasoning in a variety of contexts to the state's standards." There is a good deal of overlap in the expectations regarding estimation in grades 2–6. Also, rational numbers, percents and exponents are listed as part of the expectations for the number standard in both grades 7 and 8, but again, it is not clear how the expectations differ. Clarifying the expectations at each grade level will improve the standards' coherence and efficiency, help teachers understand what their students are expected to know by the end of the school year, and also will flag grade levels where the totality of the content is overwhelming.

Making explicit connections among the four content standards so that they support each other, rather than merely contributing to a lengthy checklist of independent topics, is another tack to make the standards more coherent and instruction more efficient and more effective in deepening student understanding.

Following are two examples: Most problems in data analysis involve numbers, estimation, graphs and sometimes equations. So if units in number and algebra intentionally employ examples from data analysis, students can develop their

understanding of three standards simultaneously. Many aspects of algebra have historical and mathematical roots in geometry, so examples from one standard can reinforce topics in the other.

- **The standards do not make clear how the content standards are to be integrated with the mathematical processes standard.**

Mathematical competence requires a balanced mix of procedural skills, conceptual understanding and the capacity to solve practical problems. The difficulty in judging how well New Jersey executes this requirement stems from the fact that the state has a separate Problem Solving standard. It is not always clear how this standard is to be integrated with the other content-based standards and, therefore, how teachers are expected to assess whether students are meeting the problem-solving standard. In general, there are many more statements in the standards that seem to emphasize procedural knowledge over conceptual understanding. Better balance could be achieved by using illustrations that highlight conceptual understanding.

As one looks across the document, the balance among concepts, procedures and problem solving is usually attained in geometry (especially, units of measurement) and data analysis; however, practical problem solving is either not stressed enough or is missing at key points in the number and algebra standards. For example, by grade 8, concepts and procedures involving percentage are well established. But grade 8 indicators do not spell out the full range of practical percentage problems students should master, namely, sales tax, tips, interest, discounts and compound interest. And although the grade 8 algebra standard mentions translation of a verbal sentence into an equation or inequality, there is no emphasis on translating problem contexts into equations thereafter, despite the further development of concepts and procedures for solving equations in the grade 12 standards. New Jersey standards have no analogs of the Indiana standards "Solve word problems that involve linear equations …" (Algebra I) and "Solve word problems using quadratic equations" (Algebra II), both of which are followed by examples.

New Jersey's choice of wording sometimes leaves the impression that the indicators are overly focused on procedures or recalling definitions, rather than on understanding concepts. As an example, in the grade 8 Number Sense strand, the indicator, "Recognize that repeating decimals correspond to fractions and determine their fractional equivalents," is likely to be interpreted by teachers as asking for a memorized procedure. Explaining that procedure, on the other hand, requires logical reasoning, as well as understanding the heart of base-ten numbers and place values. The state will want to be careful about the way it articulates its indicators so as not to imply it is asking for students to recall when in fact it wants them to be able to comprehend the mathematics involved in the procedure. A strong aspect of Indiana's examples is that they often scaffold tasks with the final task in the set expressing the highest level of understanding expected.

RECOMMENDATIONS FOR IMPROVING THE MATHEMATICS STANDARDS

✓ ***Develop grade-level or course-level standards for additional grades (K, 1, 9, 10 and 11) and clarify what is required to graduate from high school.***

To promote equity and excellence, New Jersey should complete its continuum in mathematics. Foundation skills in grades K and 1 are important. Without standards for these grades, mathematics education could be shortchanged in these early years. As it stands now, the grade 12 standards are built on the cumulative knowledge of four years of solid coursework, but New Jersey's exit exam is administered in approximately the third quarter of the grade 11 school year. To promote equity and excellence, New Jersey should clarify expectations up through the time the graduation exam is administered in addition to maintaining its standards in the 12th grade.

✓ ***Improve specificity.***

New Jersey should add illustrations or sample problems to convey the intended cognitive level of its expectations. There are a number of compelling reasons to do so:

1. The inclusion of illustrations or sample problems in mathematics standards really helps teachers, students and the larger education community to grasp the level of proficiency expected. Currently, there are many statements that are too close to each other for the reader to discern the difference from grade level to grade level.

2. Pressure is on for schools not only to raise the level of mathematics proficiency for all of its students but also to close achievement gaps between subgroups of students who have been historically underserved by the education system and the larger majority. Teachers who are new to the profession, or underprepared in their understanding of mathematics and/or how to teach the discipline, are far better supported by explicit standards. The more explicit standards are, the easier it is for teachers to figure out exactly what they themselves need to know and the repertoire of approaches they need to help struggling students succeed.

3. When a state chooses, as New Jersey has, to cast mathematical processes as a separate standard, it is critical to provide concrete examples of exactly what Problem Solving, Communication, Connections, Reasoning, Representations and Technology mean at each grade level.

4. It is important to help students see mathematics as a holistic, coherent discipline. This is optimally done through supplying concrete examples that show significant relationships among the strands, for example, geometry and algebra, and how knowledge of two seemingly disparate concepts can be brought to bear on a problem.

- ✓ ***Ensure that the standards focus on essential knowledge and skills: Some significant content is missing or insufficiently developed.***

New Jersey's standards are comprehensive, so attending to missing content is more a process of painstaking substitution, reducing redundancy, and making natural and powerful links among the standards than addition. If revision is carefully carried out, New Jersey students will be prepared for both the world of work and the demands of college-level mathematics.

Even when New Jersey's standards are delineated grade by grade, the differences in the expectations from one grade to the next are not always precise. The state should concentrate on refining the standards to clarify what is new at each grade level and to avoid statements that are too much alike.

MAJOR FINDINGS: MATHEMATICS ALIGNMENT OF ASSESSMENTS TO STANDARDS

Achieve carried out a detailed study of New Jersey's assessments and their alignment to New Jersey's revised Core Curriculum Content Standards in Mathematics (July 2, 2002). Achieve reviewed three tests:

New Jersey Assessment of Skills and Knowledge (NJ ASK 4) Form A (Spring 2003)
Grade Eight Proficiency Assessment (GEPA) Form S (Spring 2003)
High School Proficiency Assessment (HSPA) Form B (Spring 2003)

It is important to note that of these three assessments, only that for grade 4 is based on the 2002 edition of the New Jersey standards. The state is making the transition to the new standards; consequently, the assessments at grade 8 and high school are based on the previous edition of the standards (1996). Achieve's review of these two tests as compared to the 2002 edition will help the state to identify gaps in alignment and develop a fully aligned system of standards and assessments. Each of the assessments was analyzed using Achieve's Assessment-to-Standards Protocol for mathematics.

To help evaluate the cognitive demand of individual items, and ultimately the rigor of an assessment overall, Achieve has constructed a six-point scale for mathematics. Each level is briefly described below.

- Level 1 items require that the students recall information such as a fact, definition, term, symbol, notation, common equivalents or properties of objects.
- Level 2 items call for the use of a practiced routine to achieve an accurate result.
- Level 3 items require a level of mental effort that goes beyond a rote response.
- Level 4 items require that the student formulate and implement a solution.
- Level 5 items require that the student reason beyond the given information to consider other instances and/or the limitations of a mathematical statement.
- Level 6 items do not take place in an on-demand setting. The complexity and kind of reasoning expected here is higher, and the student's reasoning is informed by use of a broader range of resources.

STRUCTURE OF THE ASSESSMENTS

GRADE 4 ASSESSMENT

The New Jersey Assessment of Skills and Knowledge consists of six sections (one section is composed of field test items) that are administered over two days of testing. The first day the test includes sections 1 to 4 and takes about one hour and 35 minutes.

The second day it includes sections 5 and 6 and takes one hour and 10 minutes. These times include testing time, directions and breaks. In any given test section, students may go back and review their work on that section only.

Overall, the test includes 32 multiple-choice and 6 open-ended items. The field-test section includes six multiple-choice and two open-ended items.

The students are not allowed to use a calculator to complete the first four items (section 1 of day one) in the test, but they are allowed to use a calculator in the rest of the test. In addition, students get a set of mathematics punch-outs that includes the following: 1 hexagon, 2 trapezoids, 3 rhombi, 6 triangles and 1 ruler.

GRADE 8 AND HIGH SCHOOL ASSESSMENTS

The Grade Eight Performance Assessment and the High School Proficiency Assessment consist of three 30-minute segments. (Students also take a fourth segment composed of field test items.) The exam and field test items are given to students on a single day; a two-and-a-half-hour session (including a 30-minute break) is recommended.

Each segment of the exam presents 10 multiple-choice items followed by two open-ended items. Students record their answers in a separate answer booklet. Students may review their answers only while they work on a segment.

Students may use calculators, including scientific and graphing calculators, on all parts of the exam. They are also provided a 6-inch/15-cm punch-out ruler, as well as a formula sheet with commonly used measurement formulas.

In contrast to New Jersey's assessments in language arts literacy, which, for the most part, have similar structures, strengths and areas in need of improvement at the different grade levels, the state's assessments in mathematics have substantial differences. For this reason, it makes sense to report Achieve's findings with respect to the newly developed grade 4 test (NJ ASK 4), based on the state's revised standards, separately from the findings for New Jersey's tests at grade 8 and high school, which are still based on the earlier edition of the standards.

STRENGTHS OF THE ASSESSMENTS

NJ ASK 4

- **NJ ASK 4 is strongly aligned to New Jersey's revised Core Curriculum Content Standards in Mathematics.**

Since NJ ASK 4 is based on the July 2, 2002, edition of the standards, it is not surprising that it is more strongly aligned to the revised standards than are either the grade 8 test (GEPA) or the high school test (HSPA) — both of which are based on the 1996 edition of

the standards. The grade 4 test can serve as a model for bringing not only GEPA and HSPA into alignment but also the additional tests New Jersey intends to add in response to NCLB. Items on NJ ASK 4 are strongly aligned to the 4th-grade-level standards with 100 percent of the items being well aligned to the content described by the standards and 95 percent of the items being well aligned to the performances called for by the standards (the remaining 5 percent are partially aligned). Moreover, NJ ASK 4 is a well-crafted assessment: only two items were found to have technical flaws that could yield misleading information about student performance, and both flaws are easily remedied. Achieve has reviewed more than 25 large-scale mathematics assessments, and these data are among the best we have found. NJ ASK 4 in mathematics has other excellent characteristics.

Number and Operations: The items students are asked to solve without the use of a calculator are exactly the kind of items targeting computation that one would hope to see on a large-scale assessment such as this.

Patterns and Algebra: The three items that ask students to extend patterns are complementary; they do not ask students to do the same things over and over again as assessments too often do.

Geometry and Measurement: Students must have a good grasp of the concepts to answer the items correctly. Geometry items on state assessments often focus on measuring procedural skills only. New Jersey's item set is a good mix of items that measure skills, conceptual understanding and application. This is evident particularly in the way the open-ended questions are used to assess different aspects from those assessed through the multiple-choice items.

Data Analysis, Probability and Discrete Mathematics: The set of items measuring this standard is rigorous for the grade level of the students. The expectations set forth in this standard are not the most common set of expectations found in other sets of standards. In addition, the items assessing this standard show to have the best balance of assessing skills, conceptual understanding and application.

- **The overall rigor and balance of NJ ASK 4 is generally on target.**

Reviewers determined that, in terms of the intellectual demand they place on students, 5 percent of the items were Level 1 (requiring students to recall information); 90 percent were Level 2 (requiring students to carry out a practiced routine, such as measure, calculate or compare two values, to obtain an accurate result); and 5 percent were Level 3 (requiring students to engage in thinking beyond a rote response, such as organize and display data, re-express and use a formula, interpret information in a table or graph, explain a concept, or carry out multiple steps). Moreover, the rigor of each set of items is appropriate and, consequently, so is the level of the grade 4 assessment as a whole.

By making a few judicious item substitutions, as detailed in Achieve's accompanying technical report, the state can improve the coverage of some important content and bring the assessment squarely into balance.

- **New Jersey generally makes good use of open-ended items on its grade 4 test.**

The inclusion of open-ended items enhances the challenge of assessments. The power of items with an open-ended format lies in their giving students the opportunity to demonstrate a deeper, more complex understanding of content than can other formats. New Jersey is sending a strong signal to its students, teachers and public that it values students' ability to reason, problem solve and communicate mathematically. Open-ended items are used particularly well in the Geometry strand.

GRADE EIGHT PROFICIENCY ASSESSMENT (GEPA)

- **Although it was not designed to measure New Jersey's revised standards, the state's grade 8 test (GEPA) generally aligns with the content and performance expectations of New Jersey's revised standards in mathematics.**

Reviewers found that 94 percent of the items are well aligned to the content described by the standards, while 88 percent of the items also are well aligned to the performances called for by the standards; the remaining 12 percent are partially aligned. In addition, only one item was found to have a technical flaw and should be replaced.

- **The sets of items on GEPA assessing New Jersey's four content standards have a number of positive aspects that the state can build on to improve its test.**

Following is a standard-by-standard summary highlighting major findings regarding GEPA:

Number and Operations: While the greater proportion of questions were based on whole numbers, some items included positive and negative values and some went beyond whole numbers to include operations with fractions and decimals. In addition, several questions were asked about proportionality and percent — two key ideas at this grade level.

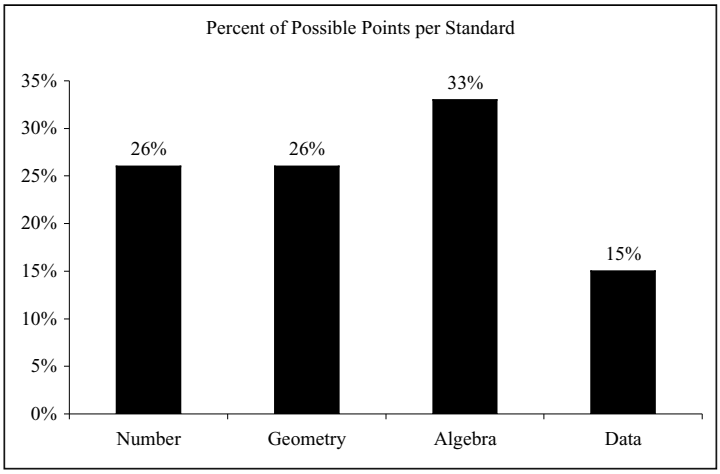
Patterns and Algebra: The set of items in this section assesses the ideas that are core for these standards. These items are appropriate for assessing the progress indicators here.

Geometry and Measurement: The level of cognitive demand of the items in the Geometry section was the best of all the sections. One item seems to be more challenging than many of the items included in the Geometry section in the HSPA.

Data Analysis, Probability and Discrete Mathematics: Of the set of items that measure this standard, the most appropriate items are those concerned with probability because

they ask students questions that are appropriate for this grade level and in a way that requires students to have an understanding of probability concepts.

One positive aspect of the test, evident from the chart, is the fact that the point distribution for the current test preserves an emphasis on algebra. That emphasis is an appropriate emphasis on a grade 8 test and should be retained when the test is fully calibrated to the revised standards. It is important to ensure that the items within each standard assess a balance of concepts.



AREAS FOR IMPROVEMENT

- **The test items on the GEPA tend to present students with too low a level of intellectual demand for grade 8.**

The relatively low level of demand stems in part from having too many items based on whole numbers, using numbers that are too simple for a grade 8 assessment and not requiring students to apply their knowledge in sufficiently challenging ways. Even though many items were scored a "3" for level of cognitive demand, the demand for the most part stemmed from students being required to apply multiple steps to get to a solution. The items did not ask for deeper knowledge or complex thinking. Some of those more cognitive-demanding aspects not assessed include: re-expressing and using a formula; explaining a concept; organizing, displaying, and interpreting data; using the correspondence between two mathematical representations to answer a question; or explaining why a procedure works.

By being more strategic in the way it constructs and selects open-ended items, the state can raise the rigor of its test. To illustrate, some open-ended items ask students to "explain," when an explanation is unnecessary or is provided in the statement of the problem; thus, in these instances, items in a multiple-choice format would have sufficed.

Also, two open-ended items assess the same aspect of percents, when one would have been sufficient.

New Jersey also might want to consider adding gridded-response items to its mathematics assessments as a way to ratchet up rigor in a cost-effective way. This item format demands more thinking on the part of students since gridded-response items do not provide students with a choice of answers — they have to come up with the answer themselves and record the numerical answer directly.

- **The GEPA tends to overassess some aspects of New Jersey's standards and underassess others.**

Reviewers found that particular aspects of New Jersey's standards are assessed multiple times, while other critical aspects — especially looking toward high school requirements — are not assessed at all. Examples of cumulative progress indicators being hit too many times are 8.4.1.B.4 *Solve problems involving proportions and percents* (4 items); 8.4.3.A.1 *Recognize, describe, extend, and create patterns involving whole numbers, rational numbers, and integers* (6 items); and 8.4.3.D.4 *Create, evaluate, and simplify algebraic expressions involving variables* (4 items).

In contrast, Algebra and Patterns (standard 3) includes 10 progress indicators, of which only three are assessed on this test (8.4.3.A.1, 8.4.3.B.1 and 8.4.3.D.4). The set of items assessing this standard include 11 items of which six assess 8.4.3.A.1 (stated above) and four items assess 8.4.3.D.49 (stated above). Overassessment comes at the expense of having no items or too few assessing other indicators that are key for this grade level (for example, 8.4.3.C Modeling, and 8.4.3.B Functions and Relationships). On the other hand**,** no items were devoted to assessing squares, cubes, irrational numbers, percents greater than 100, the Pythagorean Theorem and measures of central tendency.

- **New Jersey's high school test (HSPA) is not a good match with New Jersey's end-of-grade-12 standards; many items map better to the state's grade 8 standards.**

When reviewers compared the alignment of the current high school test to New Jersey's revised edition of its standards, they found that a significant number of items more closely assess the standards for grade 8 than those for grade 12. Because the state has not yet broken out standards for high school either by course or by grades 9, 10 and 11, reviewers also tried to flag items that would most likely map to standards in grades 9–11 were they to be broken out. In doing so, they found that only a few of the items were likely to assess content falling somewhere between grades 9 and 11; in the end, more than 45 percent of HSPA items most closely measure 8th-grade standards.

- **The items on the HSPA tend to present students with a low level of demand.**

The overall problem in judging the level of challenge and balance on the HSPA is that too few items map to New Jersey's end-of-grade-12 standards. For example, in both Patterns and Algebra and Data Analysis, Probability and Discrete Mathematics, there are not enough items mapped to either set to make a reliable judgment. (Only three items map to the Patterns strand and just one to the Data Analysis strand.) If the items included in this test are the ones intended to assess the knowledge and abilities of students in the third quarter of grade 11 these items would represent too low an expectation, given their emphasis on recall and routine procedural application.

While it is commendable that each section of the test includes open-ended items, the state could have been more strategic in the construction and selection of those items. Consider two open-ended items about the probability of sums of the dots on two dice. These two items (1) overly guide students to a solution path (one item even provides the table asked for in the second item) and (2) assess very similar ideas about probability and ask probability questions more appropriate for grade 8 (and that are even asked in grade 4 as an investigation type of activity). This example points to the need to use open-ended items as effectively as possible.

The probability statements in the standards are indistinguishable from those in grade 8, providing no sense of progression. Therefore, the items assessing probability in this test could be grade 8 items.

APPENDIX: BIOGRAPHIES

ACHIEVE'S BENCHMARKING STAFF

The following Achieve staff and senior consultants led the analysis and report development for New Jersey.

MATHEW GANDAL
EXECUTIVE VICE PRESIDENT, ACHIEVE

Mathew Gandal joined Achieve in 1997, shortly after governors and business leaders created the organization. He opened the organization's Washington, D.C., office and helped build its programs and services.

As executive vice president, Mr. Gandal has senior responsibility for overseeing Achieve's major initiatives. He supervises Achieve's work with states and helps shape the organization's national agenda. Mr. Gandal played a lead role in organizing the 1999 and 2001 National Education Summits attended by governors, corporate CEOs and education leaders from across the country.

Mr. Gandal has extensive experience reviewing academic standards and education policies in the United States and abroad. He has written dozens of reports and articles on the topic. He has also served on a variety of national and international panels and has helped advise academic standards commissions and legislative bodies in numerous states.

Before joining Achieve, Mr. Gandal was assistant director for Educational Issues at the American Federation of Teachers (AFT), where he oversaw the national organization's work on education standards, testing and accountability. Mr. Gandal helped the AFT launch a variety of programs and publications designed to support standards-based reform efforts in states and school districts. He was the author and chief architect of *Making Standards Matter*, an annual AFT report evaluating the quality of the academic standards, assessments and accountability policies in the fifty states. He also wrote a series of reports entitled *Defining World Class Standards*, which compared student standards and achievement in the United States with that of other industrialized nations.

Prior to his role with AFT, Mr. Gandal served as assistant director of the Educational Excellence Network, an organization founded by Checker Finn and Diane Ravitch. In addition to work on domestic policy issues, Mr. Gandal was responsible for directing a series of projects aimed at helping emerging democracies around the world build democratic education systems.

Mr. Gandal is a proud graduate of the public school system in the state of Maryland. He earned a B.A. in philosophy from Trinity College in Hartford, Connecticut. He lives in Maryland with his wife and three children.

JEAN SLATTERY
DIRECTOR, BENCHMARKING INITIATIVE, ACHIEVE

Jean Slattery has been with Achieve since 1999 and currently serves as director for the Benchmarking Initiative. She was supervising director of curriculum development and support in Rochester, N.Y., from 1989 to 1997, with responsibility for overseeing the work of all subject-area directors in the K–12 instructional program. Her earlier responsibilities as a district-level administrator included serving as director of the middle school (1987–89) and junior high (1985–87) programs. During that period, she initiated Teachers as Partners, a peer-coaching staff development program funded by the Ford and Matsushita (Panasonic) Foundations.

Dr. Slattery served as a peer consultant on standards and assessment for the U.S. Department of Education. She also has served as a consultant to the Washington, D.C., school district; San Diego Unified School District; a Washington state consortium of rural schools; and the Arkansas and Illinois Departments of Education. Dr. Slattery has also worked for the Council for Basic Education on projects involving the Flint Community School District, the Nevada Education Department and the Cleveland Municipal School District.

Dr. Slattery received a bachelor's degree in chemistry from Albertus Magnus College, a master's degree in science education from Yale University and a doctorate in science curriculum from the University of Rochester.

JOANNE THIBAULT ERESH
SENIOR ASSOCIATE, ENGLISH LANGUAGE ARTS, ACHIEVE

JoAnne Thibault Eresh is a senior associate at Achieve, where she leads the English language arts aspects of the Standards-to-Standards Benchmarking and Assessment-to-Standards alignment reviews. She taught writing at the university level and English at private and public high schools in St. Louis, Mo., and in Fitchburg, Mass. She began her work in curriculum design and performance assessment in 1979 under Superintendent Richard C. Wallace, Jr., and from 1981 to 1994 was director of the Division of Writing and Speaking for the Pittsburgh Public Schools. During that time she directed The Pittsburgh Discussion Model Project, funded by the Rockefeller Foundation and part of the CHART network, and she later directed the imaginative writing part of the ARTS Propel Project, a joint project with Harvard's Project Zero and the Educational Testing Service. She was the Pittsburgh district coordinator for the New Standards Project and wrote the teachers' guides for the New Standards ELA Portfolios. In 1995 she was one of the original resident fellows at the Institute for Learning at the University of Pittsburgh's Learning Research and Development Center, as well as coordinating the New Standards Linking Projects. From 1997 to March 2001, she was the coordinator of staff

development in Community District Two in New York City where she was responsible for the hiring, training, and coordination of that district's staff development group.

KAYE FORGIONE
SENIOR ASSOCIATE, MATHEMATICS, ACHIEVE

Kaye Forgione began consulting work with Achieve in 2000 and joined Achieve as senior associate for mathematics in March 2001. Her primary responsibilities are managing Achieve's Standards and Benchmarking Initiatives involving mathematics. Prior to joining Achieve, Dr. Forgione served as assistant director of the Systemic Research Collaborative for Mathematics, Science and Technology Education (SYRCE) project at the University of Texas at Austin funded by the National Science Foundation. Her responsibilities at the University of Texas also included management and design responsibilities for UTeach, a collaborative project of the College of Education and the College of Natural Sciences to train and support the next generation of mathematics and science teachers in Texas.

Before her work at the University of Texas, Dr. Forgione was director of academic standards programs at the Council for Basic Education, a nonprofit education organization located in Washington, D.C. Prior to joining the Council for Basic Education in 1997, Dr. Forgione worked in the K–12 arena in a variety of roles, including several leadership positions with the Delaware Department of Education. Dr. Forgione began her education career as a high school mathematics teacher. She taught mathematics at the secondary and college levels as part of adult continuing education programs.

Dr. Forgione received a bachelor's degree in mathematics and education from the University of Delaware, a master's degree in systems management from the University of Southern California, and a doctorate in educational leadership from the University of Delaware.

MARA CLARK
RESEARCH ASSOCIATE, BENCHMARKING INITIATIVE, ACHIEVE

Mara Clark is the research associate for Achieve's Benchmarking Initiative and assists in the coordination of Achieve's state benchmarking work and the production of the initiative's publications. She also contributes to the English language arts Standards-to-Standards Benchmarking and Assessment-to-Standards alignment reviews. Before joining Achieve in this capacity, she was with the American Diploma Project (ADP), a joint partnership of Achieve, The Education Trust and the Thomas B. Fordham Foundation. While with the American Diploma Project, she worked closely with postsecondary faculty, high school teachers and business representatives from across the

country on the development of ADP's Benchmarks for College and Workplace Readiness.

Ms. Clark holds a B.A. in English from the University of Dallas in Irving, Texas.

CONSULTANTS AND EXPERT REVIEWERS

Achieve relied on the expertise of nationally respected experts in academic content, standards, curriculum and assessment design to inform and conduct the standards benchmarking and alignment of assessments to standards.

ENGLISH LANGUAGE ARTS

ARTHUR N. APPLEBEE

Arthur N. Applebee is leading professor in the School of Education, University at Albany, State University of New York, and (with Judith Langer) is director of the federally sponsored National Research Center on English Learning and Achievement. The Center has an active research and development agenda across subject areas in elementary and secondary instruction, including teacher education and effective uses of technology.

During his varied career, Professor Applebee has worked in institutional settings with children with severe learning problems, in public schools, as a staff member of the National Council of Teachers of English and in research and professional education. He joined the faculty at the University at Albany from Stanford University in 1987, as part of a SUNY-wide Graduate Research Initiative in Writing and Literacy.

With degrees from Yale, Harvard and the University of London, Professor Applebee focuses his studies on how children and adults learn the many specialized forms of language required for success in school subjects, life and work. In particular, his research has reframed the ways in which both scholars and practitioners think about critical issues in curriculum and instruction in reading, writing and the English language arts. Since the early 1970s, he has also worked with the National Assessment of Educational Progress, helping to design, implement, interpret and report a continuing series of evaluations of the educational attainment of U.S. students. His first book (1974) became a classic in its field, and the seven other books, 14 National Assessment monographs, and 67 reports, articles, and book chapters that have followed have been equally influential.

An internationally recognized expert, Professor Applebee consults at the national, state and district levels on effective approaches to curriculum, instruction, assessment and professional change. He is a former editor of *Research in the Teaching of English*, a past president of the National Conference on Research in Language and Literacy and a

recipient of the David H. Russell Award for Distinguished Research in the Teaching of English, from the National Council of Teachers of English.

EUNICE ANN GREER (LEAD REVIEWER)

Eunice Ann Greer is a principal research analyst at the American Institutes for Research in Washington, D.C. Her work is focused on assessment design and development and the alignment and implementation of standards-based systems of instruction and assessment. Dr. Greer was an associate superintendent for the Illinois State Board of Education, where she directed the Illinois Reads Statewide Reading Initiative. Prior to that, she was the division administrator for standards and assessment for the Illinois State Board of Education. Dr. Greer was instrumental in Illinois's successful application for a $37 million Reading Excellence Act Grant from the Department of Education. Under her leadership, Illinois was the first state to receive the Five Star Award for Exemplary Statewide Reading Initiatives from the International Reading Association. Dr. Greer also has worked as an assistant professor in the department of curriculum and instruction at the University of Illinois at Urbana-Champaign, as the director of research for an urban middle school reform project at the Harvard Graduate School of Education and as a literacy assessment coordinator for the University of Illinois's Center for the Study of Reading.

CAROL JAGO

Carol Jago teaches English at Santa Monica High School in Santa Monica, Calif., and directs the California Reading and Literature Project at UCLA. She also edits the California Association of Teachers of English (CATE) quarterly journal, *California English*. Carol has written a weekly education column for the *Los Angeles Times*, and her essays have appeared in *English Journal*, *Language Arts*, *NEA Today*, *The Christian Science Monitor*, as well as in other newspapers across the nation. She has served as director of the National Council of Teachers of English Commission on Literature and currently is a member of NCTE's Secondary Section. NCTE has published her books *Nikki Giovanni in the Classroom*, *Alice Walker in the Classroom* and *Sandra Cisneros in the Classroom*. Her other books for teachers, *With Rigor for All: Teaching the Classics to Contemporary Students*, *Beyond Standards: Excellence in the High School English Classroom*, and *Cohesive Writing: Why Concept Is Not Enough*, are published by Heinemann.

JAMES MARSHALL

James Marshall is professor of English and English education at the University of Iowa, where he currently serves as associate dean for teacher education. A high school English teacher for six years, Marshall is especially interested in the relationships between writing and literary understanding and the ways that literature is discussed in English classrooms. His books include *Ways of Knowing: Research and Practice in the Teaching of Writing*, with James Davis; *Teaching Literature in the Secondary School*, with Richard Beach; and *The Language of Interpretation: Patterns of Discourse in Discussions of Literature*, with Peter Smagorinsky and Michael Smith. He has served as chair of the

National Council of Teachers of English Standing Committee for Research, as chair of the NCTE Research Assembly, as a trustee of NCTE's Research Foundation, as executive secretary for the Standards Project in English Language Arts and as chair of the English Language Arts Assessment for the National Board of Professional Teaching Standards. He has won the James N. Murray Award for Teaching, the Outstanding Teaching Award, the Regents' Award for Faculty Excellence and the Hancher-Finkbine Medallion for Faculty Leadership at the University of Iowa.

LAURA MCGIFFERT

Before joining Achieve in 1998, **Laura McGiffert** was a high school English language arts teacher for five years in Colorado, where she was involved in a districtwide effort to refine and align local standards and assessments. In 1995, she was awarded "Best First Year Teacher" in Eagle County School District at the high school level. She also taught writing and literature at Colorado Mountain College. While at Achieve, Ms. McGiffert has brought these experiences to bear as a member of the English language arts team for Achieve's Benchmarking Initiative and has worked on the reviews for 10 states.

Ms. McGiffert is currently the director of the Mathematics Achievement Partnership (MAP), a multistate collaboration to improve mathematics performance in the middle grades. As the principal manager of this project, Ms. McGiffert assumed primary responsibility for the development of *Foundations for Success: Mathematics for the Middle Grades*, which represents the core knowledge and skills that students should learn to be prepared for high school and beyond. These expectations now are being back-mapped to kindergarten to indicate a clear progression of important knowledge and skills through the grades. To this end, Ms. McGiffert oversees Achieve's Mathematics Advisory Panel, an expert panel of mathematicians, mathematics educators, curriculum specialists, and state and local math supervisors representing a broad spectrum of perspectives about mathematics education. She also is responsible for coordinating state activities, including professional development pilot projects and meetings of state mathematics education leaders.

A native Washingtonian, Ms. McGiffert holds a master's degree in education policy from the Georgetown University Public Policy Institute, a master's degree in secondary education from the University of Colorado at Boulder and a bachelor's degree in English and American literature from Harvard University.

LEROY MILLER

Leroy Miller received his B.A. in English from Fairleigh Dickinson University and his M.A.T. in secondary English from Miami University. Mr. Miller is in his 28th year of teaching 11th-grade English at Sidney High School, in Ohio, where he also serves as English department chair for a nine-member department. Mr. Miller also is a teacher of advanced placement and honors American literature and an adviser for the Academic Competition Team. In addition, he served as a commissioner on the Governor's

Commission on Student Success in 2000 and is a member of Ohio's Academic Content Standards Writing Team for English language arts.

MATHEMATICS

HAROLD ASTURIAS (LEAD REVIEWER)

Harold Asturias is the deputy director of the Mathematics and Science Professional Development at the University of California Office of the President. He provides oversight to the Mathematics Professional Development Institutes (MPDI) and the California Subject Matter Projects (CSMP). Both statewide projects join K–12 teachers with University faculty to improve teacher content knowledge. Previously, he served as the director of the New Standards Portfolio Assessment Project and the Mathematics Unit for New Standards. In that capacity, he led the development team of experts whose efforts, involving many states and more than a thousand teachers, resulted in the successful production of two assessment systems: the New Standards Portfolio and the Reference Examination. Mr. Asturias was a member of the writing group for NCTM's Assessment Standards for School Mathematics. He has extensive experience providing professional development in the areas of standards and assessment in mathematics for teachers in large urban districts (Chicago, Los Angeles, New York City) and small rural districts. Over the past three years, he has focused in the area of designing and implementing professional development for K–12 California mathematics teachers who teach English language learners.

PAM BECK

Pam Beck taught for 10 years in central California public schools before joining the Balanced Assessment team at the University of California at Berkeley (a project funded by the National Science Foundation). This team produced assessment tasks for students from elementary to high school level that were published by Dale Seymour under the title *Balanced Assessment for the Mathematics Curriculum*. Since 1994, she has worked at the University of California, developing mathematics curriculum and assessment.

During this time, Ms. Beck directed the development of a standards-based mathematics examination (the *New Standards Reference Examination*) given at the elementary, middle and high school levels. She helped develop the *New Standards Performance Standards*. She worked as part of the team that wrote *Core Assignments in Mathematics*, published by the National Center on Education and the Economy. During this same period, Ms. Beck provided professional development to numerous and varied districts (including Hanford, Calif.; Los Angeles Unified; and New York City). She currently directs an NSF-funded project to develop a Web-based task bank. This task bank's purpose is to provide teachers and others with a wide variety of mathematics tasks useful for classroom assessment and indexed for optimum usefulness.

JOAN I. HELLER

Joan I. Heller is an educational psychologist who specializes in analyzing school-based learning and thinking. Dr. Heller earned her Ph.D. in 1979 from the University of Pittsburgh where she did doctoral and postdoctoral research at the Learning Research and Development Center. For the next decade, Dr. Heller directed several research projects at the University of California, Berkeley. On these projects, she designed and assessed instructional materials and conducted research evaluating teacher and student learning in response to innovative classroom structures and activities.

In 1992, Dr. Heller joined the Center for Performance Assessment at Educational Testing Service where she conducted research and development on projects related to assessment. One such project was a collaboration between ETS and the California Department of Education aimed at developing a statewide portfolio assessment system in language arts, mathematics, science and history/social science. She also facilitated teams of teachers in developing and trying out curriculum-embedded performance assessment tasks in these subject areas.

Dr. Heller currently directs Heller Research Associates, the research division of Gordon & Heller, Inc. Evaluation studies performed by Dr. Heller and her associates combine quantitative and qualitative approaches to evaluate the impact of teaching, assessment and learning experiences on teachers, classrooms and students. These studies have evaluated the effectiveness of publicly and privately funded projects in improving teaching and learning of a variety of academic subjects, including mathematics, science and the visual and performing arts.

DONALD R. KING

Donald R. King received his Ph.D. in mathematics from the Massachusetts Institute of Technology. Professor King is an associate professor of mathematics at Northeastern University. Previously, he was a visiting assistant professor at Salem State College, a visiting assistant professor at the University of California at San Diego and a high school mathematics teacher in Boston, Mass. Professor King is a member of the Mathematical Association of America, American Mathematical Society and the National Association of Mathematicians. He also is active in professional and community service: he was a parent member of the Mathematics Focus Group for Boston Public Schools in 1997; director from 1993 to 1994 of NUMATH, Northeastern University's program to foster minority mathematical achievement and talent in high school; an adviser to algebra-in-middle-schools projects from 1990 to 1992; a review panelist for three years for Ford Foundation postdoctoral fellowships for minorities; and an adviser to Massachusetts's pre-engineering program for minorities from 1988 to 1991. Professor King recently gave a speech at the American Mathematical Society's Special Session on Teaching Mathematics in the New Millennium: "Changing School Outcomes: Raising Standards and Promoting Equity." Professor King has advised Achieve on the quality of standards and assessments in a number of states including Minnesota, Oklahoma, New Jersey and Texas.

FABIO A. MILNER

Fabio A. Milner is a professor of mathematics at Purdue University in West Lafayette, Ind., where he has been on the faculty since 1983. He was an associate professor of mathematics at the University of Rome II in Rome, Italy, from 1987 to 1992. In addition, he was a consultant for UNESCO at the "International Meeting on Training and Pedagogical Improvement in the Thematic Areas" in La Plata, Argentina, and for the Indiana Department of Education in the development of end-of-course high school evaluations.

Professor Milner spent many months as a visiting professor in various institutions in Argentina, China, France and Italy, for periods ranging from two weeks to several months. He is the author or co-author of more than 60 scientific papers and two books, and he has delivered presentations to audiences in more than 20 countries. He has directed and codirected ten doctoral dissertations and his former students' positions range from researchers in large companies to full professors in research institutes.

He received a doctorate in mathematics from the University of Chicago, a master's degree from the same institution and a licenciado degree from the University of Buenos Aires, Argentina.

LYNN A. STEEN

Lynn A. Steen received his Ph.D. in mathematics from the Massachusetts Institute of Technology. Professor Steen has been a member of the St. Olaf College faculty since 1965 and currently is the director of Institutional Research and Planning. He is leader of the quantitative literacy initiative of the National Council on Education and the Disciplines (NCED) at the Woodrow Wilson Foundation and a member of the Mathematics Achievement Partnership (MAP), a project of Achieve, Inc., to create a common set of expectations for mathematics for grades K–8. Professor Steen is former president of the Mathematical Association of America and former chair of the Council of Scientific Society Presidents. He has reviewed mathematics standards for Achieve from Texas, Oklahoma and Massachusetts.

Board of Directors, Achieve, Inc.

Co-Chairs
Arthur F. Ryan, Chairman and CEO,
Prudential Financial

Governor Bob Taft
State of Ohio

Co-Vice Chairs
Governor Gary Locke
State of Washington

Kerry Killinger, Chairman, President and CEO,
Washington Mutual, Inc.

Board Members
Craig R. Barrett, CEO,
Intel Corporation

Governor James E. McGreevey
State of New Jersey

Governor Bill Owens
State of Colorado

Governor Edward G. Rendell
Commonwealth of Pennsylvania

Governor Mike Rounds
State of South Dakota

Edward B. Rust, Jr., Chairman and CEO,
State Farm Insurance

President
Michael Cohen

Chairman Emeritus
Louis V. Gerstner, Jr.

Achieve, Inc.
www.achieve.org

1775 Eye Street, NW
Suite 410
Washington, DC 20006
Phone: (202) 419-1540
Fax: (202) 828-0911