

MEASURING UP:

**A STANDARDS AND ASSESSMENT
BENCHMARKING REPORT FOR**

NEW JERSEY

Prepared by Achieve, Inc. for

Commissioner David Hespe
New Jersey Department of Education

Mr. Arthur F. Ryan
Chairman and CEO, Prudential

Governor Christine Todd Whitman
State of New Jersey

November 2000

TABLE OF CONTENTS

Executive Summary.....	4
Introduction.....	8
The Achieve Benchmarking Methodology.....	12
Results for New Jersey.....	18
Major Findings: Standards.....	19
Major Findings: Assessments.....	32
Recommendations for Moving Forward.....	40
Appendix	44

EXECUTIVE SUMMARY

Achieve, Inc. was created in 1996 by governors and business leaders to serve as a clearinghouse and resource center on education standards, assessment and accountability. As part of its mission, Achieve provides states with candid feedback on the quality of their academic standards and assessments. At the request of the Department of Education and Prudential, Achieve conducted such an evaluation for the state of New Jersey over the summer and fall of 2000.

This report presents the results of Achieve's in-depth evaluation of the quality, rigor and alignment of New Jersey's standards and assessments in language arts literacy and mathematics. In particular, it provides New Jersey policymakers with answers to the following questions:

- How do New Jersey's education standards compare with those of high-performing states and nations? Are the expectations for New Jersey's schools and students high enough?
- How well do the New Jersey assessments measure the knowledge and skills laid out in the standards?

By becoming one of the first states to seek out Achieve's benchmarking of state standards and tests, New Jersey has demonstrated its commitment to establishing a world-class education system. The state understands that continuous improvement of standards and tests is necessary to keep pace with other states and nations. As New Jersey begins the next round of revision of its standards, and a subsequent review of its tests, the state has the opportunity to reflect on what has worked well and what can be strengthened. This report should prove helpful in this regard, as our evaluation points to a number of strengths in the New Jersey system and a number of areas that deserve special attention if the standards and assessments are to continue driving instruction and achievement in the right direction.

RESULTS FOR NEW JERSEY

- ❑ **Achieve's benchmarking evaluation found that New Jersey's language arts literacy and mathematics standards contain some important strengths.** On the whole, the standards are straightforward and jargon-free. The content that is spelled out is reasonable and manageable and represents most of what one would expect students to know by the end of grades 4, 8 and 12 in language arts literacy and mathematics. And, the test specifications and curriculum frameworks documents for language arts literacy and mathematics provide more detail about what content the standards expect.
- ❑ **However, Achieve's review also found a number of areas that must be addressed if schools and parents are to use the standards to dramatically raise student achievement.** The clarity and specificity of the New Jersey standards fall well below that of exemplary standards from other states and nations, providing insufficient guidance and assistance for teachers and parents. In some standards, student learning develops with sufficient rigor and complexity from grade level to grade level, but others lack such progression. In addition, some standards are either too demanding or not rigorous enough for the grade level in which

they appear. And, the language arts literacy standards omit important content, particularly in early literacy.

- ❑ **Achieve’s review of New Jersey’s assessments in language arts literacy and mathematics found that they stand out among the 9 states Achieve has reviewed.** With two key exceptions, the tests are challenging and measure important knowledge and skills in language arts literacy and mathematics. Reviewers found consistently that the level of challenge of the test items in most cases is appropriate for the grade level tested, and that the items tap both students’ basic knowledge and skills as well as higher level skills such as reasoning. However, the grade 11 mathematics assessment, currently being field-tested, is substantially less rigorous than what the standards imply is expected of all students.
- ❑ **New Jersey’s language arts literacy and mathematics standards do not clearly communicate what students should know and do to succeed on the state assessments.** As noted above, the state assessments are of high quality. But due to the standards’ vagueness, the tests sometimes focus on some standards to the exclusion of others. The language arts literacy tests overall are very strong, yet they are related to only a few of the state’s standards; while the math tests particularly for grades 8 and 11 emphasize lower level content, omitting important concepts in algebra and geometry.

RECOMMENDATIONS FOR MOVING FORWARD

We understand that New Jersey will initiate a revision of the standards in early 2001 and that the state assessments are already being phased in, starting with the core subjects in elementary and middle school. Given this timing, and in the spirit of continuous improvement, we offer the following suggestions for strengthening the state’s standards and assessments:

- ✓ **New Jersey should build on the strengths of the state assessments and revise the standards to provide explicit guidance to educators and the public about what all students are expected to learn.** The state should provide more clarity and specificity to New Jersey teachers and citizens by developing a new set of standards based on the expectations embodied in the assessments and in exemplary standards from the U.S. and abroad. The standards should be made more useful to all classroom teachers by providing detail and guidance for two grades at a time — as the curriculum frameworks do — rather than every four years, and they should illustrate what the state expects of all students using sample math problems, sample reading passages to show the complexity of reading materials children can read as they grow older and sample student writing.
- ✓ **With relatively minor revisions, the assessments can be made stronger to ensure that they provide solid information about student progress toward the standards.** After creating a new set of standards, New Jersey should turn its attention to developing the next set of standards-based tests. The state should focus on strengthening the high school mathematics assessment to ensure that it reflects the challenging expectations the state has for high school students and continue to address issues surrounding the scoring of open-ended items throughout the language arts literacy tests.

-
- ✓ **To reinforce the state’s goal of high standards for all students, New Jersey should consider launching an early reading initiative and joining the Mathematics Achievement Partnership, a multistate consortium focused on raising mathematics achievement to world-class levels.** The state should emphasize a balanced program in early literacy and mathematics, paying attention to both the basics and more advanced skills. Focusing attention and resources on early literacy and middle grades mathematics will send the right signals to educators and the public — that the state is committed to raising the academic achievement of all students.

* * *

To date, New Jersey has made great progress in its education reform efforts. By taking the steps we outline in this report to improve the Core Curriculum Content Standards and alignment of the state assessments at grades 4 and 8 and the field test at grade 11, we believe that policymakers can address the concerns being raised by educators, parents and the business community, while continuing their commitment to high standards for all. We at Achieve are grateful for the opportunity to conduct this analysis and hope that the information provided in this report will be helpful to New Jersey as the state continues to work toward a higher-performing and more accountable education system.

INTRODUCTION

RAISING STANDARDS IN AMERICA'S SCHOOLS

Since the release of *A Nation at Risk* in 1983, schools, states and national policymakers have been concerned with improving the level of academic achievement of all students. At that time, most school systems awarded diplomas based on Carnegie units, which generally represented “seat time,” as opposed to a demonstration of knowledge and skills. Under the Carnegie or credit-based system, postsecondary institutions and employers had difficulty determining what students had learned, since course content varied from school to school. Additionally, it was increasingly clear that a large percentage of our nation’s students, particularly minorities and the poor, were being dramatically underserved by their education system, as they were not provided with the same rich curriculum and learning opportunities as their wealthier counterparts in the suburbs.

In an attempt to raise the level of student learning across the board and create a more publicly accountable education system, states, districts and national organizations began the process of drafting content standards to define what students should know and be able to do. In New Jersey and across the country, standards are now the driving force in efforts to improve equity and excellence in education by holding *all* students to common, high expectations.

Since the early 1990s, 49 states have developed academic standards for their students and 48 are putting in place assessments to measure those standards. By stating clearly the knowledge and skills students are expected to gain as a result of their schooling, reformers hope that students will better understand what is expected of them, schools will improve their programs to help students achieve those expectations, and low-performing districts and schools in particular will be challenged to raise the level of teaching and learning. Because states have made substantial investments in the new standards and tests — and many states are beginning to hold students and schools accountable for performance — policymakers and the public want to know how their standards compare to what other states and countries expect. Achieve was created precisely to address this issue.

Born out of the 1996 National Education Summit, Achieve helps states ensure that they have in place standards that compare favorably with the academic expectations of other states and high-performing nations and assessments that accurately measure student achievement against those standards. An independent, bipartisan, nonprofit organization overseen by a Board of Directors composed of governors and corporate CEOs, Achieve serves as a clearinghouse and resource center on education standards, testing and accountability, working primarily with states to support their work in these areas.

RAISING STANDARDS: THE NATIONAL CONTEXT

We have come a long way in the effort to improve schools for all students through the development and implementation of standards. Many states have had a chance to develop standards, begin implementation, incorporate feedback, and revise standards, assessments and

policies as necessary. While in earlier years developing clear, rigorous, measurable standards was a challenge for many states, we now have a better picture of what strong academic standards look like. They are clear and specific enough to guide curriculum planning and test development without infringing on local control; they set rigorous yet reasonable expectations for all students and raise the bar higher than it is currently set for many students; they integrate content knowledge with important skills and processes; and they are widely read and understood by parents, educators, business people and policymakers.

Now states are grappling with other important issues. In particular, they are beginning to refine the relationship between the state and local schools. Of course, each state has a different balance, but states and districts are learning that respecting local control while ensuring high standards for all students means being clear and specific about the knowledge and skills students are expected to learn, while leaving the “how” of education up to schools and districts.

This innovation is at the heart of standards-based reform: The previous state role of monitoring compliance to rules and regulations has shifted to one of setting expectations for results and supporting schools and districts in meeting those expectations. In the standards-based system, schools now can focus on how to organize themselves to achieve results, including integrating standards into local curriculum, improving instructional strategies and programs, making decisions about textbooks and other resources, providing extra help to students in danger of not meeting the standards, interpreting and using data from state and local assessments, and organizing the school day and year to meet the needs of all students.

BENCHMARKING TO THE BEST

To help states like New Jersey in their efforts to continuously improve expectations for all students, Achieve provides *standards and assessment benchmarking*. Through benchmarking, Achieve compares a state’s academic expectations against the best available models from the U.S. and the world. States like New Jersey that have sought benchmarking services from Achieve are committed to raising standards for student performance and holding schools accountable. They want their citizens to know that the standards they have set compare favorably with the expectations other states and nations have for their students. They also want to understand whether the tests they use to assess student progress against the standards truly measure what they expect all students to know and be able to do. Lastly, they want objective, credible, concrete recommendations for ways to improve their standards and assessments.

Benchmarking is a highly respected practice in the business world. It is an activity that looks outward to find best practices and high performance and then measures actual business operations against those goals. Benchmarking in education follows the same principle. It is appropriate at a time when state education reforms are focused on raising student and school performance, as states want and need an external yardstick to gauge their efforts.

By benchmarking academic standards and assessments, Achieve hopes to help states answer the following questions:

- How do our education standards compare with those of other high-performing states and nations? Are the expectations for our students and schools high enough?
- How well do our assessments measure the knowledge and skills laid out in the standards?

Achieve is involved in benchmarking for another important reason: States have traditionally had limited access to high-quality, trustworthy information about education standards. This is due partly to the fact that the standards movement in education is relatively young. But it is also a result of the disparate nature of much of the work that has been done to date. While the standards reviews and “report cards” issued by other organizations have helped to focus national attention on the quality of standards, their judgments have often been in conflict, and their tone has not always been constructive. States increasingly are looking for independent, credible advice on these issues.

Achieve’s benchmarking efforts are not designed to grade or rank states. Instead, we have created a service that is diagnostic in nature — one that yields detailed, reliable information that we hope states will find useful. In addition, our focus on assessments as well as standards and the alignment between standards and assessments allows us to truly determine what the state expects all its students to know and be able to do and whether the standards are a strong enough foundation for the state’s efforts to improve education performance.

THE ACHIEVE BENCHMARKING METHODOLOGY

Establishing and refining standards and assessments is an iterative process; so too is designing procedures to judge their quality. The foundations for Achieve's two-step approach to judging the quality of standards and their alignment with assessments were laid when Michigan Governor John Engler and North Carolina Governor James Hunt agreed to participate in a pilot study in 1998. Since the pilot study we have systematically refined the benchmarking methodology and will continue to do so as our experience base grows.

PHASE ONE: DEVELOPMENT

For assistance in designing and carrying out the initial benchmarking process, Achieve contracted with two nationally recognized leaders in standards and assessment, the Council for Basic Education (CBE) and the Learning Research and Development Center (LRDC) at the University of Pittsburgh.

CBE designed the original procedure for benchmarking state standards to state, national and international "benchmark" standards recognized for their quality and/or for producing high student achievement. Using a scoring rubric developed for Achieve, CBE compared the content and skills set forth in the Michigan and North Carolina standards to several benchmark documents, and further analyzed these state standards on the basis of their clarity, specificity and measurability.

The process for examining the extent to which the Michigan and North Carolina state assessments measure the standards was developed by LRDC. Skilled judges with expertise in academic content and assessment design applied a multistep procedure, or protocol, to determine the degree to which those states' assessments align with the standards. This protocol yields valuable information about the strengths and weaknesses of the assessments.

PHASE TWO: REFINEMENT

Drawing upon the work of CBE and LRDC in the Michigan and North Carolina pilot experiences, the findings and research from the Third International Mathematics and Science Study (TIMSS), as well as the work of other researchers and analysts, Achieve refined its benchmarking process in several critical areas. Following is an overview of the kinds of issues that standards benchmarking addresses and the methodology that Achieve reviewers use in carrying out the process.

Achieve's comparison of state standards to benchmark standards is designed to answer the following questions:

- Do the standards define a comprehensive, yet manageable academic core for all students? Are there key concepts or skills missing in the standards?

-
- Do the standards convey both the breadth and depth of knowledge and skills expected of students? Have choices been made about what is most important for students to learn, and when?
 - Are the standards as rigorous as they should be? Do the expectations described by the state standards compare favorably to those of the benchmark standards?
 - Do the standards define both what students should know (i.e., content knowledge) and what they should be able to do with that knowledge? Or is one overemphasized at the expense of the other?
 - Are the standards clear and explicit, conveyed in a way that educators and parents can understand and use to improve student achievement?

To ensure that the benchmark standards documents used as exemplars are indeed the best for this purpose, Achieve commissioned expert reviews of a variety of sets of standards. Reviewers concluded that California’s and Massachusetts’ standards in English, those of North Carolina, Texas and New Standards in early literacy, and those of Arizona and Japan in mathematics had the most value for benchmarking.¹

Selecting these benchmarks proved to be a difficult task because no one set of standards is perfect, and judgments about the quality of standards are 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 the state standards to these benchmarks will yield helpful diagnostic information and policy suggestions for states to consider.

For example, in the area of language arts, California distinguishes those standards concerned with informational text from those concerned with literature. This makes it easy for teachers to help students grasp the different strategies used in reading, understanding and writing the two types of text. Massachusetts’ standards are organized in grade spans common to a number of states, including New Jersey, but Massachusetts articulates its expectations for student learning more precisely than do most states. And in mathematics, Arizona’s standards provide a level of specificity and detail that is helpful for teachers — especially for elementary generalists — while Japan’s standards provide an economical, yet focused and rigorous treatment of the discipline.

Once the benchmarks were selected, Achieve consultants constructed “side-by side” charts for both English and mathematics which reorganized the standards by key topics, such as “writing strategies,” or “geometry: symmetry, congruence and similarity.” A state’s standards were then mapped alongside the benchmark standards so that differences in their respective treatment or sequencing of topics could be readily discernible.

¹ Achieve created “benchmark profiles” for each of these documents that provide contextual information about the standards and summarize their strengths and weaknesses. The profiles are available upon request.

In the future, Achieve will continue to examine documents as potential benchmarks. We may decide to use part of a document as an exemplar in one strand of a content area and simultaneously use a different part of another set of standards as an exemplar in another strand.

Achieve also refined the methodology used in judging the degree of alignment between state assessments and standards to ensure that the most important elements were fully captured. Achieve’s assessment-to-standards comparison has been developed to address the following questions:

- Does each assessment measure only content and skills reflected in the standards? In other words, can everything on the test be found in the state standards?
- Does each assessment fairly and effectively sample the important knowledge and skills in the standards? In other words, to what extent does each assessment measure the key content and skills for a grade level?
- Overall, is each assessment sufficiently challenging?

Alignment is not a “yes or no” question; it is the degree to which standards and assessments are in agreement and serve in conjunction with one another to guide and support student learning. Consequently, responding to the above questions requires a systematic procedure to probe the different factors that contribute to alignment. Achieve’s process or protocol for determining the alignment of assessments to standards is based upon five criteria. Some criteria apply to individual items and result in individual ratings, while others apply to the entire set of items related to a given standard, such as “measurement” or “literary analysis,” and result in written judgments. Application of the protocol provides rich information about alignment of tests and standards that is typically unavailable to states. The fundamental criteria for Achieve’s alignment process are as follows:

- **Confirmation or construction of test blueprint.** Reviewers check to see that each item corresponds to at least one standard. If no test blueprint is provided or the blueprint supplied by the test developer does not stand up to scrutiny (i.e., a significant number of items mapped to one standard or objective are found to be more closely related to a different one), then reviewers construct a new test blueprint. Because test blueprints are typically the basis for state score reports it is important that they be as accurate as possible.
- **Content centrality.** This criterion examines the quality of the match between the content of each test item and the content of the related standard. Reviewers determine how closely the content of the item aligns with that of the related standard, and then assign the item to one of four categories based on degree of alignment.
- **Performance centrality.** Each item places a certain type of cognitive demand on a student (e.g., the student is asked to “identify” or “analyze.”) If an item simply requires a student to “identify,” and the corresponding standard requires a student to “analyze,” then there is a mismatch between the two performances. Reviewers assign each item to one of four categories based on the degree of alignment.

-
- **Challenge.** Two factors are considered in judging challenge. Reviewers determine if the *source* of the challenge in each item comes from content in the standards that students must know to answer the item correctly, or for some extraneous factor such as the language of the item or technical problems in the way the item is constructed. Reviewers determine the *level* of challenge by comparing the kinds of performances required by a set of items to those called for by the related standard, and also judge whether the set of test items has a span of difficulty appropriate for students at the grade level tested.
 - **Balance and range.** Range compares the extent to which the knowledge and skills delineated in the standards receive the same emphasis on the assessment, and if that emphasis is appropriate. Range examines the degree to which the assessments sample the knowledge and skills described in the standards, since it is very difficult for one assessment to measure the full complement of knowledge and skills required by state standards. Evaluating balance and range provides both qualitative and quantitative information about the choices states or test developers have made.

The alignment analysis is a process of managing expert judgment. There is no mathematical formula for matching a test to standards. Rather, the process relies on experienced, knowledgeable educators who bring their experience and knowledge to bear in applying the criteria for judging alignment.

The judges selected by Achieve to analyze the alignment between assessments and standards are a deliberate mix of classroom teachers, curriculum specialists and subject-matter experts, with extensive expertise in content and assessment design. They often have experience in large-scale assessments and/or standards development and represent a diversity of viewpoints. This diversity of backgrounds has proven invaluable in arriving at considered judgments (for example, determining the appropriateness of a test item for a particular grade level). Teams of reviewers, facilitated by an experienced group leader, meet in small groups to discuss their judgments. These judgments are further structured by the tools developed for the process, including more detailed rubrics for judging the various dimensions of alignment outlined above.

Determining how well a particular assessment captures the content and performances expected by the standards requires that reviewers not only gauge the quality of the match between the assessment and the standards, but also provide a critique of particular strengths and weaknesses of the assessment. An added benefit of the process is that it inevitably sheds light on the clarity, comprehensiveness and rigor of the standards themselves.

In addition to both refining the process by which benchmark standards were selected and applied and clarifying the elements of the protocol, Achieve also has broadened the impact of the benchmarking evaluations by adding a significant professional development component, whereby state officials were trained in how to apply the benchmarking protocol.

WORKING WITH NEW JERSEY

The Achieve benchmarking process examined New Jersey’s academic standards against our benchmark standards and then compared New Jersey’s assessments against the state standards. As the first step, Achieve consultants constructed “side-by-side” charts comparing the content and skills found in the New Jersey standards to those of California and Massachusetts in English, North Carolina and Texas in early literacy, and Arizona and Japan in mathematics. Achieve then asked several national experts in standards and assessment to review New Jersey’s standards and the comparisons to the benchmark standards and react to a set of questions about the standards. These experts hold diverse opinions about standards, curriculum and assessment matters, and each has considerable experience in writing and researching standards and assessments.

To gauge the alignment of New Jersey’s assessments to the standards, Achieve conducted an intensive comparison of its assessments and standards. For New Jersey, Achieve gathered educators with deep experience in teaching, curriculum, standards and testing to carefully study the proficiency assessments and apply the Achieve protocol.

Achieve staff and consultants analyzed the various reviews, highlighting the strengths and weaknesses of the New Jersey standards as identified by the subject area experts, and summarizing the results of the alignment study as reported by the teams of experienced reviewers. The findings described in this report represent, to the extent possible, consensus opinions of Achieve’s consultants and experts.

(Lists of Achieve’s experts and consultants who participated in the standards benchmarking and assessment analysis for New Jersey can be found in the Appendix.)

RESULTS FOR NEW JERSEY

New Jersey's Core Curriculum Content Standards, adopted in 1996, not only define the expectations for student learning in core subjects but also define educational effectiveness in the state. According to the Comprehensive Plan for Educational Improvement and Financing, a 1996 law, a thorough and efficient education is one in which children meet the outcomes defined in the standards. The state has developed statewide "proficiency assessments" in grades 4 and 8, and is field-testing an assessment in grade 11, to determine the extent to which students have met the standards. The high school test, which will be administered in 2002, eventually will be used to determine whether students graduate from high school.

The Core Curriculum Content Standards are broad statements of student learning. Accompanying Descriptive Statements help explain the overall intent of the standards and provide additional insight as to what students might be expected to know and do as a result of the standards. Each standard is followed by a set of Progress Indicators that are meant to delineate further the content knowledge and skills students are expected to acquire by grades 4, 8 and 12.

To help districts and schools develop curricula to enable students to meet the standards, the state department of education has developed curriculum frameworks in language arts literacy and mathematics. The frameworks are distributed to principals throughout the state and are available on the department's Web site. The state department has also published a Directory of Test Specifications and Sample Items for each test. These documents, which describe the knowledge and skills in the content standards that are measured on the tests, are distributed to curriculum specialists and test directors throughout the state, but are not available on the Web site because of copyright limitations.

The focus of Achieve's analysis is the Core Curriculum Content Standards in language arts literacy and mathematics and the alignment of the state assessments to those standards. Nonetheless, Achieve appreciates that the New Jersey standards are part of a comprehensive system that also includes curriculum frameworks, test specifications and sample tests (at grades 4 and 8). The system is further supported by the state's well-developed Web site, which provides helpful information to teachers, parents, students and the general public and direct access to all relevant documents, except the test specifications. Indeed, New Jersey is to be commended for aggressively pursuing opportunities to make technology readily accessible to teachers and students and for its ongoing commitment to making technology a powerful and flexible vehicle for advancing the level of instruction and student achievement.

MAJOR FINDINGS — LANGUAGE ARTS LITERACY AND MATHEMATICS STANDARDS

In its standards review, Achieve commissioned experts from a range of perspectives to analyze the content standards, curriculum frameworks and test specifications for language arts literacy and mathematics and compare them to "benchmark standards" in each subject. Achieve's analysis of standards was not intended simply to render a judgment about the quality of a state's standards documents. States that ask Achieve to benchmark their standards against the "best in class" standards from the U.S. and abroad do so because they lack an external yardstick — a way of determining whether the expectations they have set for student performance are as high as they can be and whether the standards provide an appropriate level of guidance to teachers. We have sought to answer these questions for states by convening experts to select the best standards documents in each subject that they can find and evaluate how the state's standards compare to those benchmark documents.

Achieve's standards are very high; our recommendations are aimed not just at how New Jersey can improve the standards, but at how to make them as good as the best standards in the world.

While Achieve's expert reviewers found that the language arts literacy and mathematics standards are not clear, specific or consistently rigorous enough to support higher achievement, (particularly in language arts literacy), they also concluded that the math and literacy statewide assessments embody challenging, meaningful expectations that should help all schools focus their efforts on improved teaching and learning. Moreover, the detailed guidance already provided by state assessment rubrics, curriculum frameworks and test specifications can be used to great advantage in revising and fortifying the standards.

The following discussion summarizes and explains the most important findings from the benchmarking study. More detailed findings are included in the accompanying technical report.

Strengths of the Content Standards

In general, Achieve's expert review of the New Jersey standards in language arts literacy and mathematics found some significant strengths:

1. On the whole, the New Jersey standards are straightforward and jargon-free.

The language is readable and understandable to a wide range of audiences, including students, parents, teachers and others. And for the most part the standards are written in measurable terms, although there are a few "process" standards that are difficult to measure.

2. The amount of content in the New Jersey standards is reasonable and manageable, representing most of what one would expect by grades 4, 8 and 12 in language arts and mathematics.

In contrast to the standards in some states, which are too voluminous for students to possibly master, the New Jersey content standards in language arts literacy and mathematics represent a relatively parsimonious set of expectations that students could reasonably attain. New Jersey has made choices about what is most important for all students to know and be able to do in language arts literacy and mathematics. The mathematics standards commendably refer to the majority of essential topics in the disciplines; they present the subject as a unified whole, rather than a set of “knowledge bits,” thus nicely illustrating the connections among the important concepts in math. However, there is some important content that is insufficiently addressed or missing in both subjects, and in some areas it is unclear if the essential content is included. The state has also made an attempt to show which standards students should have mastered in a particular grade level and which they should continually address at increasingly complex levels. But this attempt was not always successful, as we will illustrate below.

Areas for Improvement

Despite its strengths, our review found several areas that should be addressed if New Jersey schools and parents are to use the standards to substantially raise student achievement. In particular, Achieve’s experts found that the standards tend to be overly general and do not show a clear progression of knowledge and skills across the grades. Moreover, the standards in some areas are not as rigorous as the benchmark standards, and in some cases the progress indicators are not appropriate for particular grade levels. Our experts also noted specific instances in which the standards should be developed more fully.

1. The clarity and specificity of the New Jersey standards fall well below that of the benchmark standards, leaving too much room for varying interpretations of the standards’ content and rigor.

Developing standards that effectively articulate a considered, balanced and explicit progression of knowledge and skills that students are expected to attain from kindergarten through grade 12 is a major challenge that all states face. New Jersey’s challenge is particularly acute, since the state faces, on the one hand, a constitutional guarantee that “all children will receive a thorough and efficient” education and, on the other, a legislated mandate for local control over curriculum and instruction. New Jersey’s solution to this dilemma appears to represent a compromise — one that many states in the early stages of the standards movement attempted: Adopt a general set of standards at the state level that can accommodate multiple programs of study.

Unfortunately, this is a compromise that for the most part works better in theory than in practice. Teachers are left with vague exhortations that provide little sense of what they should do to develop students’ knowledge and skills. The state has attempted to provide guidance to teachers by publishing curriculum frameworks, where each progress indicator is enhanced with multiple activities meant to illustrate how it can be addressed in the classroom. But the frameworks represent the opposite extreme from the standards: They are voluminous and overly detailed and they often fail to give teachers — or students and parents — a clear sense of what every student in New Jersey should know or be able to do.

While most of the progress indicators in the New Jersey standards are not difficult to understand on their face, their vagueness makes them difficult to interpret with precision. Lack of clarity is caused mainly by the lack of specification. More specificity is needed throughout to avoid needless repetition and to ensure comprehensive coverage of important content in New Jersey's schools. As things stand, many of the standards and indicators force the teacher, student and assessment developer to guess about the parameters.

A comparison with the benchmark documents illustrates the lack of clarity and specificity in the New Jersey content standards. Consider the following comparison of standards for measuring perimeter, area and volume from New Jersey, Japan and Arizona.

New Jersey	Japan	Arizona
<p>4.5.4. Use a variety of tools to measure mathematical and physical objects in the world around them.</p> <p>4.7.9. Understand the variety of ways in which geometric shapes and objects can be measured.</p>	<p>4.2 Help children understand the concept of area and become able to measure the area of simple geometrical figures, and become able to measure the size of an angle.</p> <p>4.2.B.1. Enable children to understand the concept of area and measure the area in simple cases.</p> <p>4.2.B.1.a. Understand the meaning of unit and measurement in area.</p> <p>4.2.B.1.b. Know about the units (square centimeter, square meter, square kilometer, acre and hectare) to be used in measuring area.</p> <p>4.2.B.1.c. Know how to measure the area of squares and rectangles.</p>	<p>5.4M-E4 Represent and solve problems relating to size, shape, area and volume using geometric models</p> <p>5.4M-E4 PO1. Solve problems using given formulas for simple area and perimeter</p> <p>5.4M-E4 PO2 Identify a variety of shapes having the same perimeter and area</p> <p>5.5M-E3 PO Differentiate between perimeter and area of quadrilaterals using concrete and illustrative models</p> <p>5.5M-E4 PO1.A Develop a procedure or formula to calculate area and perimeter of simple polygons</p> <p>5.5M-E4 PO2.A Use given formulas to find area and perimeter of simple polygons</p> <p>5.5M-E5 Describe how a change in the linear dimension of an object affects its perimeter, area and volume</p> <p>5.5M-E5 PO1 Describe the effect on perimeter and area when one dimension of an object is altered</p>

In the New Jersey standards, the language leaves it unclear what exactly a teacher should do to enable students to demonstrate a mastery of the standard. Does the standard mean that students should know that plane figures, like squares and triangles, have both perimeter and area (i.e., two different kinds of measures)? Or does it mean that students need to know that plane figures, like squares and triangles, can have their perimeters measured with appropriate measuring devices or have such dimensions calculated indirectly via the use of formulas? In contrast to the much more specific language from Japan and Arizona, New Jersey's lack of precision is bound to be problematic for teachers, curriculum developers, test developers and parents.

Likewise, because the language arts literacy standards often are written at a very general level, it is difficult to tell what mastery would look like. For example, grade 4 language arts literacy indicator 3.4.6. states, “Read literally, inferentially, and critically,” and 3.4.8 simply requires students to “read with comprehension.” When standards are this vague they give very little guidance to teachers or test developers about what kinds of “literal” or “critical” reading “comprehension” grade 4 students can do or what distinguishes “inferential” reading in grade 8 from grade 4. Can students in grade 4 recall details from a story, or do they understand the story’s main idea? Can students in grade 8 understand enough information from an instruction manual to follow directions? Can they connect what they read in a story to something they learned in class? None of this is clear from the standards as they currently stand. In contrast, Massachusetts and California spell out what basic and advanced reading comprehension of literature and of informational texts looks like for students in grades 4 and 8.

Or consider New Jersey’s writing standards. The only writing standards for grade 12 are indicators 3.3.18 and 3.3.19: “Write for real audiences and purposes, such as job applications, business letters, college applications, and memoranda” and “Write a research paper that synthesizes and cites data.” Other than delineating some possible classroom and assessment tasks, these indicators say little about the characteristics of the tasks that would be assessed. Thus they provide little guidance to test developers about how to structure the tasks to elicit such characteristics, or to teachers about how to structure instruction to enable students to reach the standard. By contrast, writing standards in Massachusetts make explicit the features of writing that ought to be developed — and thus what should be assessed. Consider the following excerpts for grades 11–12:

Massachusetts Learning Standard 19: Students will write compositions with a clear focus, logically related ideas to develop it, and adequate detail.

11-12: Write coherent compositions with a clear focus, adequate detail, and well-developed paragraphs, and evaluate the effectiveness of the strategies they used to generate and organize their ideas.

Massachusetts Learning Standard 20: Students will select and use appropriate genres, modes of reasoning, and speaking styles when writing for different audiences and rhetorical purposes.

11-12: Use effective rhetorical strategies and demonstrate understanding of the elements of discourse (purpose, speaker, audience, form) when completing expressive, persuasive, informational, or literary writing assignments.

Massachusetts Learning Standard 21: Students will demonstrate improvement in organization, content, paragraph development, level of detail, style, tone, and word choice (diction) in their compositions after revising them.

11-12: Revise their writing to improve style, word choice, sentence variety, and subtlety of meaning after rethinking how well they have addressed questions of purpose, audience, and genre.

The New Jersey standards’ lack of specificity also makes it hard to tell whether they are as rigorous and demanding as the benchmarks. For example, New Jersey standard 3.3, “Students

will write in clear, concise, organized language that varies in content and form for different audiences and purposes,” does not make clear, as the benchmark documents do, that students will be accountable for knowing the structure of the English language, sentence structure, punctuation, capitalization or spelling. Similarly, in mathematics, reviewers were uncertain how much New Jersey expects in terms of solving algebraic and trigonometric equations, the standard arithmetic and algebraic algorithms and their meaning, and formal and informal proofs.

Finally, many of the New Jersey progress indicators are difficult to assess because they are often too vague. In mathematics, reviewers noted that New Jersey frequently uses verbs such as “explore,” “investigate” and “reflect on” in its indicators — verbs that do not adequately describe student behaviors that can be effectively measured. Standard 4.7 (spatial sense), for example, overuses “explore,” which speaks more to the learning environment than to student results. Other verbs such as “understand” or “recognize” are open to wide interpretation. As a case in point, exactly what does a student need to do to “demonstrate,” as called for in 4.11.19, an understanding of “the role of functions as a unifying concept in mathematics?”

It is clear that New Jersey does not mean for the standards to serve as a curriculum, but rather as a catalyst for curriculum development. However, there is a middle ground. A level of detail similar to that of the benchmark standards would help New Jersey’s standards achieve this purpose. More detail and some examples of student work that meets the standards would make the New Jersey standards more useful to teachers as they develop curricula and prepare students for tests and would help parents more clearly understand exactly what their children should know and be able to do. Without more detail in the standards, it will largely fall to the state assessments to communicate exactly what the state expects of all students.

2. While in some standards, student learning develops with sufficient rigor and complexity from grade level to grade level, other standards lack such progression.

The standards do not develop sufficiently over time in large part because of the way they are organized. Some progress indicators are included at two or three grade levels, with a clear indication that there are increasing levels of difficulty at each. Others are not mentioned in later grade levels, implying that they should have been mastered at the earlier level or that they should be continued. In practice, this principle does not seem to be followed consistently, nor is it an effective organizing tool to communicate how students’ learning develops over time.

This finding is particularly acute throughout the language arts literacy standards, and represents a significant issue for the state to address. Our experts found that, while the expectations in the listening and speaking strands become increasingly rigorous as students move through the grades, the expectations for reading, writing and literature lack a logical progression of knowledge and skills. There is little indication of increasing intellectual demand as students move through the grades. By the end of high school, one would expect to see substantial rigor beyond what is required in middle school and elementary school; however, this is often not the case in the New Jersey standards.

The cumulative nature of the New Jersey standards tends to hinder the articulation of knowledge and skills from grade level to grade level. Take for example the difference in the way New Jersey and California address the writing process:

	CALIFORNIA	NEW JERSEY
Grade 4	Grade 4 1.10(W) Edit and revise selected drafts to improve coherence and progression by adding, deleting, consolidating, and rearranging text.	3.3.10 Revise content, organization and other aspects of writing, using self, peer, and teacher collaborative feedback (the shared responses of others). 3.3.11 Edit writing for developmentally appropriate syntax, spelling, grammar, usage, and punctuation.
Grade 8	1.6(W) Revise writing for word choice; appropriate organization; consistent point of view; and transitions between paragraphs, passages, and ideas.	
Grade 12	1.9 Revise text to highlight the individual voice, improve sentence variety and style, and enhance subtlety of meaning and tone in ways that are consistent with the purpose, audience, and genre.	

The fact that New Jersey does not address the above concepts and skills again in grades 8 and 12 could leave one to assume that the state does not expect students to achieve a higher level of sophistication with these concepts and skills. Even if one assumes that a higher level of sophistication is expected, it is hard to know exactly what that level should look like at grades 8 and 12. Which content and skills should be emphasized? What do developmentally appropriate syntax, spelling, grammar, usage and punctuation look like at grade 12 as opposed to grade 4? How should a student in grade 12 revise for content and organization, as opposed to a student in grade 4? These questions remain unanswered.

In mathematics, there are similar problems with uneven and varying progression. For example, one standard where increasing levels of difficulty is expected is the grade 8 progress indicator 4.7.16: “Develop, understand and apply a variety of strategies for determining perimeter, area, surface area, angle measurement and volume.” This indicator is expected to be continued through grade 12, but does that mean that the number of methods should increase, that the methods should become more standard or perhaps that the methods should increase in accuracy and/or precision? Because the expected continued growth is not specified, it may lead to very different interpretations. Moreover, perimeter should be mastered before grade 8, much less grade 12.

Our reviewers also found that the organization of the math standards diminishes their manageability and progression, compared with those of the benchmark states and nations. In looking across Japan’s standards, for example, reviewers find it is quite clear that elementary students are expected to learn the basic arithmetic and operational skills with whole numbers, fractions and decimals. Their middle-level standards concentrate on development of algebraic

and geometric skills. By the time Japanese students reach high school, they are expected to focus on higher-level knowledge and skills, e.g., more advanced algebra (including quadratics), geometry (including trigonometry) and probability.

New Jersey, however, has decided to earmark some progress indicators in grade 4 for which students are expected to demonstrate continued progress as they move to grades 8 and 12. It appears that grade 8 students are still expected to spend time “extending” their understanding and use of such topics as arithmetic operations, number sense, estimation, measurement, and perimeter, area and volume. When coupled with the new concepts and skills in such areas as algebra, geometry, probability and statistics that students at this level are also expected to master, the stage is set for a curriculum that is too packed to allow for adequate depth. This is an important issue. One of the key findings from TIMSS was that countries whose standards are more focused, having traded breadth for depth, tend to demonstrate higher student performance on rigorous mathematics assessments.

New Jersey’s struggle with depth may also stem from its decision to cull out 16 standards in math, as opposed to Arizona’s six standards and Japan’s five. Consolidating some of the standards could help sharpen their focus. And while having numerous standards is not at odds with presenting a holistic view of mathematics, consolidation could have the added effect of further illuminating the underlying structure of the discipline and how its branches of study are interrelated. Interestingly, New Jersey’s test specifications are organized into four or five “clusters,” each with a set of related “macros.” The state might want to consider also adapting this scheme to its standards.

If New Jersey is to retain the cumulative approach it currently takes in both language arts literacy and math, it would help to add components to the standards to illustrate their meaning at successive grade levels. In language arts literacy, increasing intellectual demand often comes from the texts students are using. Sample reading passages or suggested reading lists illustrate the kinds of texts students should be working with at each grade level, helping to show a clear progression of expected knowledge and skills. Similarly, samples of student writing could illustrate the level of sophistication expected by the writing progress indicators for grades 4, 8 and 12. In mathematics, the standards must specify what kind of growth is expected along what dimension (multiple steps, more complexity and/or greater abstraction). Adding more specificity about the content students should learn and including sample math problems would help show the level of rigor students are asked to demonstrate.

3. While some progress indicators are appropriately rigorous for the grade level in which they appear, others are either too demanding or not rigorous enough.

Each progress indicator should be appropriately demanding for the grade level in question. The knowledge and skills demanded in a grade 8 indicator, for example, should be reasonable for a student in that grade and should also match the level of demand seen in the other indicators for that grade. Overall, the developmental appropriateness of the New Jersey standards is uneven.

In mathematics, in some cases the level of performance expected of students is difficult to discern because the standards are written at a very general level. For example, mathematics

indicator 4.12.7 states that grade 4 students need to be able to “Make predictions that are based on intuitive, experimental, and theoretical probabilities.” Without further definition, this is a lot to ask of grade 4 students; on the other hand, it can be interpreted much less rigorously.

Similarly, some language arts literacy progress indicators are unreasonably challenging for the grade level in which they are placed, also in part because of their lack of precision or specificity. For example, progress indicator 3.4.18 expects grade 8 students to “Begin to identify common aspects of human existence.” In grade 12, students are asked to “Understand the relationship between contemporary writing and past literary traditions” (3.4.26); “Understand the study of literature and theories of literary criticism” (3.4.29); and “Understand the range of literary forms and content that elicit aesthetic response” (3.4.32). Even for grade 12 these expectations are extremely sophisticated and, without further clarification about exactly what high school students are expected to learn, they could encompass, as one expert wrote, PhD-level study in English literature.

The standards’ vagueness also makes it difficult to know whether the performance indicators in the earlier grades adequately prepare students for more in-depth study later on. That depends largely on how a teacher interprets the demands. For example, one standard at grade 4 reads, “Use the conventions of spoken English, such as grammar and appropriate forms of address.” Do they mean all conventions? All grammar? Other examples include: “Edit writing for developmentally appropriate syntax, spelling, grammar, usage, and punctuation” and “Use print concepts in developmentally appropriate ways.” Just what is developmentally appropriate? The definition often varies by teacher, by school, by district, even by region. Standards should provide enough information for teachers, students, assessors and others to come to a meeting of the mind about just what is developmentally appropriate.

At the same time, the level of generality may mask a lack of rigor when the standards are compared with the clear, rigorous expectations of the benchmark documents. Japan unambiguously states that students are to “master the four basic operations with whole numbers and effectively apply them to consider phenomena; use addition and subtraction of decimals and fractions.” Students are expected to carry out whole number addition, subtraction, multiplication and division and the addition and subtraction of decimals and fractions with a common denominator. In contrast, New Jersey states that students should “develop meaning for the four basic operations by modeling and discussing a variety of problems” and that they should “use a variety of mental computation and estimation techniques.” The emphasis that New Jersey gives to estimation in the early grades may be at the expense of building fluency in computational and operational skills.

Sometimes the level of expectation is clearly too low for a particular grade level, especially when compared with the benchmark standards. For example, some language arts literacy grade 4 progress indicators are more appropriate for earlier grades, namely:

- 3.4.13: Identify elements of a story, such as characters, setting, and sequence of events.
- 3.4.2: Listen and respond to whole texts.

Similar expectations appear as early as kindergarten and first grade in the California standards:

Kindergarten, 3.3: Identify characters, settings, and important events.

Grade 1, 3.1: Identify and describe the elements of plot, setting, and character(s) in a story, as well as the story's beginning, middle, and ending.

Kindergarten, 2.5: Ask and answer questions about essential elements of a text.

Grade 1, 2.2: Respond to who, what, when, where, and how questions.

Likewise, it seems rather simplistic for the New Jersey grade 8 standards to include progress indicator 3.4.17: “Read more than one work by a single author” and progress indicator 3.4.23: “Understand the role of characters, setting and events.” These are things that could be done in the elementary grades, yet they are not specifically addressed until grade 8. And, the grade 12 indicator “Understand the effect of literary devices, such as alliteration and figurative language, on the reader’s emotions and interpretation” (3.4.31) builds only slightly on the grade 8 indicator 3.4.21 “Understand the concepts of figurative language, symbolism, allusion, connotation, and denotation.”

In addition to finding inconsistencies in the level of demand across grade levels, our reviewers also found inconsistencies between standards and between grade levels in the New Jersey standards. For example, the reading strand does not explicitly state that students will *understand* figurative language until the expectations for the end of grade 8 (progress indicator 3.4.24: “Understand the concepts of figurative language, symbolism, allusion, connotation, and denotation”). However, the writing strand explicitly states that students will be *using* figurative language by grade 4 (progress indicator 3.3.9: “Use figurative language, such as simile, metaphor, and analogies to expand meaning”).

These issues — the overlap and lack of clarifying specificity among similar indicators at different grade levels, the omission in middle and high school of more sophisticated concepts (such as foreshadowing and irony in 3.4.31), and the uneven progression and rigor through the grades — are a significant weakness of the language arts literacy standards.

4. There are scant expectations for early literacy; teachers and parents have little information about how to help all children learn to read. The language arts literacy standards also inadequately describe or omit other important content.

The absence of early literacy expectations can be attributed to the way the New Jersey standards are organized. We note that, in order to maximize local control of curriculum, New Jersey has made a conscious decision to present expectations in grade spans, rather than grade by grade. New Jersey’s earliest set of expectations describes what students should know and be able to do by the end of grade 4. Since the most intensive early reading instruction occurs in the grades leading up to grade 4, the grade 4 expectations assume that students have already mastered such early literacy essentials as concepts about print, phonemic awareness, decoding and word recognition. Since virtually no mention is made of the early stages of literacy development, no systematic phonics study is specified. Instead, using context clues seems to be the major strategy that is advocated by the grade 4 expectations. This is clearly an important strategy, but not a

sufficient one, according to the nationally-heralded report from the National Research Council, *Preventing Reading Difficulties in Young Children*, which calls for a balanced approach to teaching reading that combines intensive instruction in phonics with significant exposure to rich literary experiences.

New Jersey clearly states that curriculum is locally controlled, and that the general nature of its standards is due to the fact that school districts should be able to tailor them to their particular needs. Even with this in mind, the absence of direction for early literacy instruction leaves a hole in this standards document. How will a local school district, teacher or parent know what is recommended for pre-K–3 students? How will New Jersey know if its youngest children are making good progress towards the grade 4 standards? The grade 4 performance level is generally rigorous enough to provide students a strong foundation for increasingly demanding work in the later grades, yet it assumes that students are in possession of certain basic skills, namely the early literacy skills that are never mentioned in the standards document. Without these foundation skills, the progress indicators for grades 4, 8 and 12 could be nearly impossible to achieve.

Achieve’s benchmark standards for early literacy include the K–3 English standards from North Carolina and Texas as well as the New Standards primary literacy standards. Consider the following excerpts for kindergarteners and grade 1 students regarding one aspect of the early literacy standards, phonemic awareness:

	North Carolina	Texas
Phonemic Awareness	<p>Kindergarten: Phonemic Awareness and Alphabetic Principle</p> <ul style="list-style-type: none"> ▪ Demonstrates understanding that spoken language is a sequence of identifiable speech sounds. ▪ Demonstrates understanding that the sequence of letters in the written word represents the sequence of sounds in the spoken word. ▪ Demonstrates understanding of the sounds of letters and understanding that words begin and end alike (onsets and rimes) <p>First Grade: Phonemic Awareness</p> <ul style="list-style-type: none"> ▪ Can blend the phonemes of one-syllable words. ▪ Can segment the phonemes of one-syllable words. ▪ Can count the syllables in a word. ▪ Can change beginning, middle, and ending sounds to produce new words. 	<p>K.6 Reading/phonological awareness. The student orally demonstrates phonological awareness (an understanding that spoken language is composed of sequences of sounds). The student is expected to:</p> <ul style="list-style-type: none"> A. demonstrate the concept of word by dividing spoken sentences into individual words. (K-1) B. identify, segment, and combine syllables within spoken words such as by clapping syllables and moving manipulatives to represent syllables in words. (K-1) C. produces rhyming words and distinguish rhyming words from non-rhyming words. (K-1) D. identify and isolate the initial and final sound of a spoken word. (K-1) E. blend sounds to make spoken words such as moving manipulatives to blend phonemes in a spoken word. (K-1) F. segment one-syllable spoken words into individual phonemes, clearly producing beginning, medial, and final sounds. (K-1) <p>1.6 Reading/phonological awareness. The student orally demonstrates phonological awareness (an understanding that spoken language is composed of sequences of sound). The student is expected to:</p> <ul style="list-style-type: none"> A. blend sounds to make spoken words, including three and four phoneme words, through ways such as moving manipulatives to blend phonemes in a spoken word.(1) B. Segment one-syllable spoken words into individual phonemes, including three and four phoneme words, clearly producing beginning, medial, and final sounds. (K-1)

New Jersey’s earliest expectations are for grade 4, so the only explicit reference to early literacy skills in the New Jersey document is in a descriptive statement: “Proficient readers use a repertoire of strategies (including phonics, context clues, and foreshadowing) that enables them to adapt to increasing levels of complexity, and they develop lifelong habits of reading and thinking.”

North Carolina and Texas expectations are presented in a grade-by-grade format and consequently are better able to detail early literacy expectations. Because early literacy is so important, New Jersey should consider specifically addressing reading, writing, listening and speaking expectations for pre-kindergarten through grade 3. If students do not have foundational literacy skills mastered by the end of the early grades, they will struggle to meet the expectations for the following grades in all subjects, all of which depend on these skills. It is important that teachers are offered some guidance in this area, and that parents are provided a tool against which to mark their child’s progress towards literacy. The New Jersey standards document can and should serve this purpose. The state standards could be modified in the spirit of the Massachusetts document, which is organized in grade spans similar to New Jersey, but has made a point to detail expectations for early literacy in a special chapter.

Although the absence of early literacy expectations is of the most concern, there are other concepts and skills missing from the New Jersey standards. For example, our experts noted that the following topics in language arts literacy are weak or missing:

- Knowledge of various literary genres and forms
- Nonfiction or technical text
- Literature study and literary criticism
- Academic writing
- Aspects of essay writing
- Research
- Vocabulary development
- Grammar and usage

These are key topics that are included in exemplary English language arts standards, and sometimes the New Jersey standards do little more than reference a topic. For example, pieces of a research “strand” exist, but they are neither comprehensive nor particularly comprehensible. For example, spread in and around indicators for grades 8 and 12, students must do the following:

- “Gather and synthesize data for research”
- “Write to synthesize information from multiple sources”
- “Cite sources of information”
- “Write a research paper that synthesizes and cites data”
- “Demonstrate the ability to gain information from a variety of media”
- “Use simple charts, graphs and diagrams to report data”
- “Take notes on visual information from films, presentations, observations, and other visual media”

But the standards are silent on the development of open-ended research questions, the evaluation of various sources of information, or the use of dictionaries, thesauruses, or other reference materials. Nor do the knowledge and skills related to research develop logically over time.

5. In general, the test specifications for language arts literacy and mathematics are more explicit and helpful than the content standards.

While the precision and measurability of New Jersey’s standards in mathematics are not on par with either of the benchmark documents, it was not the state’s intent to have its standards stand alone. The burden of explication was left for the curriculum frameworks, and the burden of deciding what content will be assessed on the state tests was left for the test specifications. Therefore, Achieve’s standards and assessment experts also reviewed these documents.

The curriculum frameworks are comprehensive companion documents intended to provide guidance and assistance to curriculum developers and teachers as they implement the standards. In the mathematics framework, each progress indicator is enhanced with multiple activities meant to illustrate how it can be addressed in the classroom. This level of concreteness should be of help as teachers attempt to interpret what is meant by the standards. The state’s deep concern that the standards influence and improve instruction is also reflected in the way the framework is structured. Leading off with vignettes of ideal classrooms, the framework offers engaging descriptions of effective mathematics teaching and learning environments.

Also, whereas the standards are specified for grades 4, 8 and high school, the framework culls out progress indicators for shorter grade spans, K–2, 3–4, 5–6, 7–8 and high school, providing additional guidance for the interim grades. While these additions are a good start, clear expectations for the knowledge and skills students should learn at particular grade level groupings are still not set. Questions such as, “How large are the numbers that should be dealt with at grade 4?” and “To what extent are students expected to be fluent with algebraic manipulations by grade 12?” remain unanswered.

The language arts literacy framework provides nice descriptions of classroom practice, but fails to offer teachers clear choices about what is most important for students to learn or more detailed information about standards-based instruction than that provided in the standards themselves. While extraordinarily well documented and beautifully outlined, the vignettes in the framework suffer from the same key omissions that the standards do. Each lesson that is described is superb, but where, for example, are teachers to get ideas about how to teach reading? About teaching grammar? What the standards do not spell out should have been made explicit in the frameworks, yet this is not generally the case. For example, page 165 in the vignettes offers an assessment scale that makes reference to language skills as “incorrect grammar/usage, minor grammar/usage, and correct grammar/usage,” but how are teachers to have the language to talk to children about their “minor errors” unless students and teachers share the common vocabulary of subjects, verbs, pronouns, etc.?

Unfortunately, however helpful the frameworks may be to some teachers, their overall usefulness is hampered by their massive size. Weighing in at approximately 400 pages in language arts literacy, and 700 pages in mathematics, such documents are, in all likelihood, very intimidating

and overwhelming to teachers who do not have the time or the inclination to figure out how the document is organized and where the activities are that might assist them. Elementary teachers, responsible as they are for multiple subject areas, would be better served by more user-friendly documents.

The test specifications, on the other hand, provide both more detail and more guidance than either the content standards or curriculum frameworks. In language arts literacy, the appendices of the test specifications present a more coherent picture of the discipline than do the standards, though their organization could be improved. The descriptions of content and skills in the appendices address several of our experts' concerns about missing content in the standards, with the exception of concepts and skills related to foundational early literacy skills. Namely, the appendices of the specifications address elements of good writing, make clear that the conventions of language arts literacy are part of the writing and speaking requirements, and flesh out the knowledge and skills students are expected to acquire in reading, speaking, listening, writing and viewing.

The test specifications also offer a variety of sample reading passages to illustrate the complexity and sophistication of the reading level required of grades 4, 8 and 12 students. In fact, the extended reading passages included for grade 11 indicate that students will be held to a very high level of reading skill. And, the writing prompts found in the test specifications are comprehensive and demanding. Overall, the sample test items should assist teachers and students, as they provide a clear picture of what students are expected to know and be able to do.

The mathematics test specifications are organized a bit differently than the standards, probably in part to condense the 16 mathematics standards and related progress indicators. The specifications divide the knowledge and skills from the standards that will be assessed into "content clusters" and further divide these into macros. In contrast to the standards and framework, the specifications tend to delineate skills with more precision and clarity; at grade 4, for example, the specifications indicate that students will need to know the number facts with sums to 18. Also, sample test items are included and help illustrate the kinds of questions that might be asked of students.

In both subjects, although the test specifications and sample items are more explicit than the standards, it is not always clear that they sharpen the focus of the standards. One concern is that the test specifications are not always consistent with the standards and are sometimes weaker in terms of rigor. Contrast, for example, the way problemsolving is described in standard 4.1 and its accompanying progress indicators with the way it is interpreted within each content cluster in the mathematics standards. Nonetheless, it would be helpful to share the test specifications more broadly with teachers, parents and the public as they do provide a bridge between the standards and the tests.

Yet when all is said and done, curriculum frameworks and test specifications can certainly help to explicate overly broad standards, but they are not a substitute for a set of standards that clearly define the essential knowledge and skills of a discipline. In fact, since curriculum frameworks, test specifications and tests are derived from the standards, it is difficult for these documents to overcome the deficiencies in the standards themselves.

MAJOR FINDINGS — LANGUAGE ARTS LITERACY AND MATHEMATICS ASSESSMENTS

The New Jersey proficiency assessments are constructed on the basis of the test specifications. The specifications in language arts literacy include knowledge and skills corresponding to the five content standards in that subject area: reading, writing, viewing, listening and speaking. The mathematics specifications reorganize the standards into the following content clusters for grade 4: number sense, operations, and properties; measurement; spatial sense and geometry; data analysis, probability and discrete mathematics; and patterns and algebra. Similar clusters are used for grades 8 and 11, with the exception of measurement, which is subsumed by the remaining four clusters.

Achieve examined one form each of the language arts literacy and mathematics tests (Form A) of the Elementary School Performance Assessment (ESPA), the Grade Eight Performance Assessment (GEPA) and the High School Performance Assessment (HSPA). The language arts literacy tests, which are administered over a two-day period, consist on each day of a writing prompt, multiple-choice items (five on the ESPA and 10 on the GEPA and the HSPA), and two open-ended items. The GEPA and HSPA include a revising and editing item, and the ESPA and GEPA also include a field test component. On the ESPA this consists of an additional set of five multiple-choice items and two open ended items. On GEPA this consists of an additional revising and editing item.

In mathematics, the ESPA is administered over two days. The first day consists of 20 multiple-choice items, one open-ended, six more multiple-choice and another open-ended item. The second day consists of six multiple-choice and two open-ended items in addition to a field test component that consists of both multiple-choice and open-ended items. The GEPA, administered in one day, includes four parts, each consisting of 10 multiple-choice and two open-ended items; the fourth part is the field test component. The HSPA consists of three parts, each consisting of 10 multiple-choice and two open-ended items. Although the Achieve alignment analysis did not evaluate the field test items, we did consider these sections in forming overall impressions about the tests.

Achieve’s analysis of the alignment and quality of New Jersey’s assessments is designed to address three critical issues for states:

- Does each assessment measure only content and skills reflected in the standards? Or, put another way, can everything on the test be found in the state standards?
- Does each assessment fairly and effectively sample the important knowledge and skills in the standards? In other words, to what extent does each assessment measure the key content and skills for a grade level?
- Overall, is each assessment sufficiently challenging for students?

The following discussion summarizes and explains the most important findings related to these issues. More detailed findings are included in the accompanying technical report.

Strengths of the Assessments

Achieve’s review found that, overall, the assessments are extremely strong — some of the strongest the reviewers have examined in the three years Achieve has been using this process. The tests include a high proportion of well-crafted items that measure, in challenging ways, important knowledge and skills in which the state expects all students to demonstrate proficiency. A student who does well on the assessments can reasonably be said to have mastered important knowledge and skills. This is a commendable finding and one not characteristic of most states reviewed by Achieve.

1. New Jersey’s assessments are challenging and measure important knowledge and skills in language arts literacy and mathematics.

Reviewers found consistently that the level of challenge of the test items in most cases is appropriate for the grade level tested, and that the items tap both students’ basic knowledge and skills, as well as higher level skills, such as the ability to reason. Moreover, the reviewers found that for the most part the literacy assessments make good use of open-ended questions to get at skills that are not easily measured in other ways. Generally, the reviewers were further impressed with the quality of the multiple-choice items, which efficiently and effectively focus on salient aspects of the reading passages and provide legitimate distracters. These findings are laudable; the New Jersey tests stand out among those of the nine states Achieve has reviewed.

In addition, the reviewers found that the reading items consistently ask students to show their comprehension of the reading passages and to use their understanding to make inferences about what they had read. They do not simply ask students to recall facts from the passages. Reviewers also reported that the reading passages themselves are sufficiently rich to permit such analyses, and that they are appropriate for the grade levels tested.

In writing, the reviewers found that the assessments ably tap the qualities of effective writing, and that the six-point scoring scale — a particularly strong point of the assessment — allows distinctions between gradations of effective writing. The sample papers that earned the highest score are indeed exemplary. However, the writing samples are not widely available, and the scoring guide (the rubric) is not as clear as the samples about the qualities of high-quality papers. In fact, reviewers found that the scoring of the sample papers does not match the rubric in some cases. This may have been a factor in the problems the state experienced in scoring the grade 4 language arts literacy test this year. Once these problems have been addressed, the state should consider publishing the sample papers to provide teachers and students with examples of writing at each level of proficiency.

In mathematics, the tests also include many challenging items. On the ESPA, the items measuring spatial sense and geometry are particularly strong and show a good range from easy to difficult. Also noteworthy is the fact that the use of arithmetic operations is assessed without the use of calculators. Items assessing number sense and data analysis are less challenging, with most items placed at the easy level. Several items asked questions about money, which is important, but one of the least demanding contexts for mathematics questions.

On the math GEPA, the level of challenge is generally high for items assessing number sense, patterns, relations, functions and data analysis. The test makes good use of open-ended items to elicit knowledge and thinking not always tapped by a multiple-choice question. The items assessing spatial sense and geometry are less challenging. They assess important mathematics; they just do not ask enough of grade 8 students. Some items seem more appropriate to students in upper elementary school; others test content grade 8 students ought to know but only ask them to make rudimentary use of their knowledge.

2. The match between the tests and the standards is sometimes strong and sometimes unclear, in large part because the standards are vague.

Our reviewers in language arts literacy and mathematics took two different paths in analyzing the alignment between standards and tests. For purposes of designing tests and reporting the assessment results, New Jersey matches its test items to the content clusters listed in the test specifications. This decision has merit. A standards document typically outlines the universe of important knowledge and skills, only some of which can be measured on a test. For example, a language arts literacy expectation such as 3.4.17, “Read more than one work by a single author,” is a worthy curriculum goal, but not one appropriately assessed in an on-demand statewide test. Outlining the subset of knowledge and skills that will be tested in a set of test specifications makes clear to test designers what is “fair game” on a test, and it provides a potential means of communicating this same information to students and teachers. We underscore that this is only a *potential*; the state’s standards are disseminated much more widely than the test specifications, and the alignment between the content standards and the test specifications is not perfect.

In mathematics, the reviewers agreed that the knowledge and skills in the test specifications were precise enough, and more so than the actual content standards, to serve as “standards” for the purposes of analyzing alignment. And since New Jersey mapped the tests to the specifications, it made sense to the mathematics reviewers to make constructive comments about what New Jersey actually did.

The language arts literacy reviewers, on the other hand, found that the reporting categories in the test specifications to which the state mapped its test items were overly general — clustering important knowledge and skills under vague categories like “Working with Text,” “Interpreting Text” and “Analyzing/Critiquing Text.” To the language arts literacy assessment experts, it made little sense to match test items to such all-encompassing and potentially overlapping categories. They concluded that analyzing the match between the test and content standards would provide more useful information to the state.

Thus, the results of the alignment analysis in the two subject areas are widely different. In mathematics, there is a remarkable and exemplary consistency between the content the test measures and the expectations laid out in the test specifications, and the degree of agreement between the performances expected in the test and in the specifications is almost as strong. This level of consistency is much higher than we have found in other states. However, as we noted above, in some cases the match between the test specifications and the standards is not always perfect. Often, the specifications are more concrete and explicit than the standards; in other cases, the immeasurable performance required of students in the standards (such as “explore” or

“investigate”) is changed to a more easily-assessed expectation. While these changes are helpful to test developers and sometimes necessary to improve upon the vague standards, the knowledge and skills the public believes to be important to do well on the tests may be a very different thing than what the test actually measures.

In language arts literacy, the results are of more concern. The reviewers found the test items to be only somewhat consistent with the content and skills laid out in the standards, even though the items are of consistently high quality and value. For the most part, this ambiguity about alignment reflects the language of the standards, which the assessment reviewers found to be overly general. For example, it is difficult to see how an item could not be matched with standard 3.4.6, “Read literally, inferentially, and critically” (and, in fact, a majority of items were matched to this standard throughout the three tests we reviewed). Yet our reviewers felt it is equally difficult to determine with certainty if the test item measures what this indicator intends, because no specific description of particular reading skills exists in this or any of the other progress indicators. As a guide for instruction and to enable students to do well on the test, this kind of general, vague language leaves a lot to be desired.

While these findings confirm the concerns the experts documented about the vagueness of the standards, they should not signal a reason to lack confidence in the state assessments. Instead, they signal that using the current standards to guide curriculum and instruction may not be sufficient to prepare for the state assessments. The state should take steps to revise and realign the standards and test specifications so they clearly communicate the knowledge and skills embodied in the assessments that the state considers most important for all students to master. Once these steps have been taken, the state can feel confident that schools and students who use the standards will not be surprised by material on the tests.

Areas for Improvement

As discussed above, the New Jersey tests embody challenging, meaningful expectations that are likely to raise the bar for most students. At the same time, there are some aspects of the tests that could be changed to make them even better models for curriculum and instruction and to ensure that the results are meaningful to students, schools and the public.

1. Because of the standards’ vagueness, the tests sometimes focus on some standards to the exclusion of others. And the high school mathematics assessment being fieldtested is substantially less rigorous than what the standards imply is expected of all students.

This is an issue Achieve has uncovered in several states, often because of vague, sometimes haphazard standards, and sometimes because harder concepts found in the standards are omitted from tests. We found an interesting pattern in New Jersey: The language arts literacy tests are of very high quality overall, yet they are related to only a few of the state’s standards, probably because of the standards’ inherent weaknesses; while the mathematics tests appear to make some deliberate choices about what content to cover and what to exclude.

On the grade 8 math test, for example, the algebra items focused largely on the most important algebra and functions concepts for the middle grades. On the other hand, the geometry item set

skirts several of the most important geometric concepts central to the test specifications for grade 8 students. These include similarity, the Pythagorean Theorem, circles, surface area, volume and coordinate geometry, all of which are commonly introduced to American middle school students. The specifications also indicate that students will encounter test items requiring fluency with formulas for the area and circumference of a circle and for the surface area and volume of rectangular solids and cylinders. Yet, the assessment includes no items assessing any of these concepts.

The high school field test items assessing algebra and geometry also deserve special attention, given these subjects' importance in high school. Of the 12 items related to geometry on the test form we reviewed, 10 are about direct and indirect geometric measurement (concepts that are mastered in Japan before and during middle school), with only one item related to geometric properties and one for coordinate geometry. There are no items focusing on any of the following mathematical ideas found in the test specifications: spatial relationships, relationships between geometric figures, inductive/deductive reasoning, symmetry, congruence, transformations, vectors, understanding and application of measurement formulas for volumes, surface areas, areas of polygons other than rectangles and triangles, or the coordinate system. By ignoring such critical concepts, the set of items fails to focus on several concepts central to measurement and essential to the study of geometry at the high school level.

Nor does the grade 11 field test do justice to fundamental algebraic ideas that should be taught to all students by grade 11. The HSPA does not assess students' understanding of nonlinear functions in a significant manner, nor does it build on the assessment of basic understanding of exponential functions that began in the GEPA. Reviewers expected, based on the content outlined in the test specifications, to see some assessment of quadratic, exponential and/or periodic functions, and some assessment of students' understanding of slope, but these were missing.

In addition to the concerns about the content that is emphasized on the grade 11 mathematics test, Achieve is concerned about the exam's rigor, which is weak overall and in every strand. Of the four strands assessed, the level of challenge for number sense came closest to what reviewers considered appropriate. There was a moderate range of difficulty among the items within this set, though there were no items reviewers characterized as having a high level of challenge. In the other strands, particularly for geometry and data analysis, reviewers found that the level of challenge is too low for high school students, and alignment to the rigor denoted by the standards is weak. After examining the whole test, reviewers concluded that the HSPA contains too much low-level assessment of concepts and too much "revisiting of middle grades" across all four mathematics strands.

For example, a number of items on the test appear more appropriate for grade 8 students than high school students. For example, items assessing patterns, functions and algebra overemphasize linear relationships, a concept that should have been mastered earlier. In addition, a data analysis item that asks students to read a bar graph (featured in the grade 8 standards only) represents a poor use of the open-ended format. A more appropriate use of such a format would have asked for a deeper interpretation of a more challenging type of graph that high school students would study.

Just what content should be measured and how rigorous the grade 11 test should be are difficult issues that most states are grappling with, particularly when students will be asked to pass the tests as a graduation requirement. We understand how New Jersey may have made a deliberate decision to de-emphasize certain concepts and measure others with less depth on both the grade 8 and grade 11 math tests. Yet, this is a troubling trend: Recent research from the College Board and the U.S. Department of Education also confirms what parents, students and teachers have known anecdotally for some years — that completing rigorous and meaningful high school courses in algebra and geometry is a primary indicator of college success. Consequently, giving all students a solid foundation in algebra and geometry starting in middle school is essential. This is not to say that algebra and geometry should be the sole focus of middle and high school mathematics. Business people, college faculty and math educators alike appreciate, for example, that many concepts in data, probability and statistics have increasing applicability both in the workplace and in many other disciplines. The state assessments should reinforce this message and clearly signal that all students will be held to 21st century high standards.

2. The open-ended items on the language arts literacy and mathematics tests are worthwhile and engaging, but some need further attention to ensure that they are appropriate and well crafted. Additionally, the grade 4 language arts literacy test may be too long.

The language arts literacy team noted that some of the writing tasks should be improved to make sure that they align to the standards, that students understand what is expected of them during the test, that no students are unfairly privileged because of prior background knowledge, and that scorers consistently and accurately score the essays. For example, because the writing standards are too broad, the writing prompts are only somewhat consistent with them. The directions in the tasks do not always point out precisely what the student is expected to do, who the intended audience is, or how the writing will be judged. Moreover, the picture prompts — writing exercises that ask students to respond to a picture, rather than a written statement — are not as effective as they might be. While such tasks can ensure that topics are accessible to all students and enable students to interpret the meaning of the picture in their own way, the tasks in the New Jersey assessments are not always clear. The technical report discusses this finding in more detail.

Similarly, several of the open-ended reading items presented multiple issues as the reviewers determined whether they are fairly constructed and interpreted. Some of the student papers that were scored below acceptable are, in our view, of the same high quality as some of the high-scoring student work. It often seems as if the hardest part of responding to the open-ended and writing prompts for students is not producing a response in line with the demand of the task, but in figuring out what the demands are.

In evaluating the open-ended mathematics items, the experts determined that they could be crafted more effectively to ensure that they make good use of the time required to complete them. In some cases, particularly in elementary and high school, the time students spend on these items does not appear to yield information about students' understanding in proportion to the extra assessment time required. In other cases, the items are not as challenging as they should be, nor do they require enough depth of understanding to warrant the open-ended item format. Some

of the open-ended items are too similar to multiple-choice items to be on the same form of the test, and other items do not yield the level and kind of information they might, due to flaws in the related scoring guides.

Our experts are also concerned that the grade 4 language arts literacy test may be too long. Like all states, New Jersey is trying to accomplish a number of goals with its testing program, and multiple goals sometimes forces tradeoffs among them. First, the tests attempt to measure the challenging expectations the state has set for all students and provide information about the progress students are making toward those expectations. Second, the tests intend to use a variety of tools to measure student performance, based on the sound understanding that one type of test cannot measure all important skills. Third, the state is attempting to gather this information without placing an undue burden upon teachers and schools.

As a result, the tests pack a lot into a relatively short period of time. On one day in the language arts literacy section, a student must read a complex text, answer a series of challenging multiple-choice questions to demonstrate his or her comprehension, answer one or two open-ended questions to show extended understanding and produce a writing sample in response to a prompt. Then the student must do the same thing the next day, on a different set of tasks, and the day after that as well.

We are concerned that these demands may tax students' ability to perform well and consistently on the grade 4 tests. We urge the state department of education to collect data to determine if students are failing to complete the examination because of the extensive number of tasks they are asked to complete. If so, they should be cut back, and we propose starting by reducing the administration of the field test items. Unless the state needs a huge number of fresh items each year, there is no reason for every student to participate in the field test; a representative sample would do.

RECOMMENDATIONS FOR MOVING FORWARD

New Jersey has made substantial strides in developing a standards-based system of education. The state has adopted content standards in core subjects and a related set of curriculum frameworks and test specifications. The state has also developed assessments based on the standards. These tests are critically important to improving teaching and learning, since they indicate progress toward an adequate education throughout the state.

Under New Jersey law, the state is required to revise its standards every five years; the next revision is scheduled for 2001. The state is also refining its high school test, which is scheduled to be put in place in 2002. Our review suggests that New Jersey has a substantial and strong foundation on which to build as it goes through the process of reviewing its standards and tests. We believe the standards and particularly the tests now go a long way toward driving instruction and achievement in a positive direction, and we believe they can go much farther in the next round.

Achieve recommends that New Jersey:

- ❑ **Build on the strengths of the test specifications and revise the standards to provide explicit guidance to educators and the public about what all students are expected to learn.**

In creating such a document, New Jersey will have much to draw on, starting with the existing standards and test specifications. The revised standards should include a number of components to illustrate their meaning to teachers and parents, including sample math problems, sample reading passages to show the complexity of reading materials children can read as they grow older and sample student writing that meets the standards.

In addition, the state should consider the benchmark standards our review focuses on — Arizona and Japan in mathematics, California and Massachusetts in English language arts, and North Carolina and Texas in early literacy. Other state models that are worth looking at include: Indiana, Ohio, Pennsylvania and New York's online supporting materials in English language arts. In mathematics, noteworthy documents include Achieve's MAP expectations, the New Standards performance standards, North Carolina's course of study for grades K–5 in particular and the National Assessment of Educational Progress framework (NAEP).

We urge the state in revising the document to consider writing the standards in small enough grade clusters — two years at a time, rather than elementary, middle and high school — to make them useful to schools in organizing themselves to prepare students to meet the standards. The current four-year spans may be too broad; it is difficult for a grade 5 teacher to know what to do to prepare students for standards they must meet at the end of grade 8.

- ❑ **Strengthen the assessments through relatively minor revisions to ensure that they provide solid information about student progress toward the standards.**

Most critical for the state to address is the level of rigor in the high school mathematics test that is scheduled for administration in 2002. The test must reflect the challenging expectations the state has for high school students if teachers are to provide the instruction students need to succeed and students are to put forth the effort to learn challenging material.

The state also should address some issues surrounding the scoring of open-ended items throughout the language arts literacy tests. We were pleased to see that, in August, Commissioner Hesse took steps to modify scoring of the ESPA language arts literacy component. He moved very quickly to conduct an evaluation of test data and release recalculated scores for the 1999 and 2000 ESPA test administrations. By taking two additional steps — adjusting the scoring guides and tightening up the language in the questions students are asked to respond to make sure they are fair to all students — the state can be further assured that students will be accurately assessed. More detail with specific suggestions for addressing these issues is included in the secure technical report.

In addition to these two important issues, the state can buttress support for the statewide tests and assist schools' and districts' efforts to help students reach the new standards by publicizing the expectations embodied in the tests more widely. The sample tests currently available to the public on CD-ROM are quality examples of the kinds of knowledge and skills measured on the statewide tests and should be disseminated broadly to educators, parents, business and community leaders, and the media, among others. The state could also consider releasing most or all of the actual test items, scoring guides and student papers each year. While this may have cost implications, other states such as Massachusetts and Texas use released items to build support for the tests and provide important information to the public about what the state expects all students to know.

- **Reinforce the state's goal of high standards for all students, by launching an early reading initiative and joining the Mathematics Achievement Partnership (MAP), a multistate consortium focused on raising mathematics achievement to world-class levels.**

When revising the standards, the state should also place special emphasis on providing a balanced program in primary literacy. As reflected in the exemplary documents named above, the research is clear that an approach to reading instruction in the early grades that provides skills instruction in literature-rich classrooms is the most effective way to develop children's reading abilities. As these documents also make clear, there are steps teachers in each grade should take, so the revision should specify expectations for reading, writing, listening, speaking and viewing in each grade, K–3. Given the high expectations for student performance embodied in the grade 4 literacy test, clear guidance for primary teachers is essential.

In addition to being more explicit in the standards for the early grades, the state should consider launching an intensive initiative to provide research and information, diagnostic tools and classroom assessments, technology and resources, and training to K–3 teachers responsible for teaching children to read. Other states, notably Texas and Ohio, have undertaken such initiatives that will enable all students to master the foundation literacy skills.

In mathematics, the view that is emerging in standards documents produced after New Jersey's emphasizes mastery of basic and advanced skills combined with conceptual, reasoning and problemsolving abilities at all grades. A solid sequence of algebra and geometry for all students beginning in middle school is also expected. The recent revision of the National Council of Teachers of Mathematics standards for school mathematics also makes this emphasis clear, and high-performing nations in TIMSS reflect such high-minded goals.

For these reasons, New Jersey should consider joining with other states trying to raise mathematics standards and achievement. MAP, sponsored by Achieve, marks an important new chapter in cooperation among states to address common educational obstacles, offering proof that each state does not have to go it alone to address what amount to national problems. Partner states (currently 11 leading reform states) are working together to develop a framework outlining the core concepts students need to learn to meet internationally competitive math standards; high quality middle school math curricula and professional development opportunities to help all students achieve at world-class levels; and an internationally benchmarked grade 8 math assessment that will allow states to compare annually the performance of their schools with those in other states. Joining the partnership will help New Jersey ensure that its mathematics expectations are world-class and that all children are provided with opportunities to meet these standards. It will also allow schools to compare achievement against schools in other states.

APPENDIX

ACHIEVE’S BENCHMARKING CONSULTANTS, EXPERTS AND STAFF

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.

Language Arts Literacy

Standards Experts

- Carol Jago, Director, California Reading and Literature Project, and English teacher, Santa Monica High School
- Louisa Moats, PhD, Project Director, NICHD Project, and Clinical Association Professor of Pediatrics, University of Texas at Houston
- Susan Pimentel, JD, Co-founder, StandardsWork
- Karen Wixson, PhD, Dean, College of Education, University of Michigan

Assessment Reviewers

- Sheila Byrd, Education Consultant
- Sue Craig, Education Consultant
- JoAnne Eresh, Coordinator for Staff Development, Community School District Two, New York City Public Schools (lead reviewer)
- Vernon Gettone, PhD, California Teachers’ Association
- Laura McGiffert, Senior Project Associate, Achieve
- Art Schuhart, Chair, English Department, Woodrow Wilson High School, District of Columbia Public Schools

Mathematics

Standards Experts

- Susan Eddins, NBCT, Curriculum and Assessment Leader for Mathematics, Illinois Mathematics and Science Academy
- Kaye Forgione, EdD, Assistant Director, Systemic Research Collaborative in Education for Mathematics, Science, and Technology (SYRCE), University of Texas at Austin
- Ralph Raimi, PhD, Professor Emeritus, Department of Mathematics, University of Rochester

Assessment Reviewers

- Pam Beck, Director of Mathematics Assessments, National Center on Education and the Economy (lead reviewer)
- Barbara Bethel, Senior Advisor, City Heights K–16 Educational Pilot
- Richard Escobales, PhD, Professor of Mathematics, Canisius College
- Joan Heller, Education Consultant, Heller Research Associates

-
- Donald King, PhD, Professor of Mathematics, Northeastern University
 - Lynn Raith, Assistant Director of Mathematics, Pittsburgh Public Schools
 - Norman Webb, PhD, Director, Wisconsin Center for Ed Research, University of Wisconsin–Madison

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

- Ellen Clark, Benchmarking Consultant
- Matt Gandal, Vice President, Achieve
- Lauren Resnick, PhD, Director, Learning Research and Development Center, University of Pittsburgh
- Robert Rothman, Senior Project Associate, Achieve
- Jean Slattery, EdD, Senior Benchmarking Consultant
- Jennifer Vranek, Director, Benchmarking and State Services, Achieve
- T. Jason Weedon, Project Associate, Achieve

SELECTING BENCHMARK STANDARDS

Before conducting standards benchmarking, Achieve asked 10 respected experts to examine several sets of exemplary standards documents. The original sets of standards the experts examined in English and math were:

English Language Arts

California content standards
 Massachusetts curriculum framework
 NAEP reading framework
 NAEP writing framework
 New Standards performance standards

Mathematics

Arizona content standards
 Delaware content standards
 Japan curriculum framework
 NAEP framework

From these documents, Achieve selected California and Massachusetts in English and Arizona and Japan in mathematics. Achieve acted on additional expert advice to use the K–3 English standards from New Standards, North Carolina and Texas as well.

The content and standards experts who participated in the analysis of these documents were:

English Language Arts

- Sheila Byrd, Education Consultant
- Carol Jago, Director of California Reading and Literature Project, UCLA, and English Teacher, Santa Monica High School
- Louisa Moats, PhD, Project Director, NICHD Project, and Clinical Association Professor of Pediatrics, University of Texas at Houston

-
- Sandra Stotsky, PhD, Deputy Commissioner of Academic Affairs, Massachusetts Department of Education
 - Karen Wixson, PhD, Dean and Professor of Education, University of Michigan

Mathematics

- Lawrence Braden, Mathematics Teacher, St. Paul's School
- Susan K. Eddins, NBCT, Curriculum and Assessment Leader for Mathematics, Illinois Mathematics and Science Academy
- Ed Silver, PhD, Senior Scientist, Learning Research and Development Center, University of Pittsburgh
- Harold Stevenson, PhD, Professor of Psychology, University of Michigan
- Lucy West, Director of Mathematics K–12, Community School District Two, New York City Public Schools

Achieve also asked 15 experts in science, history, geography and civics to review several sets of exemplary science and history/social science standards. More information about these experts and standards is available upon request.

* Detailed biographical information about Achieve's experts, consultants and staff is available upon request.