

Pay Now or Pay Later

The **Hidden** Costs of High School Exit Exams



Center on Education Policy

MAY 2004

Based in Washington, D.C., and founded in January 1995 by Jack Jennings, the Center on Education Policy is a national independent advocate for public education and for more effective public schools. The Center works to help Americans better understand the role of public education in a democracy and the need to improve the academic quality of public schools. We do not represent any special interests. Instead, we help citizens make sense of the conflicting opinions and perceptions about public education and create the conditions that will lead to better public schools.

The Center on Education Policy receives nearly all of its funding from charitable foundations. The statements made and views expressed are solely the responsibility of the Center.

© Center on Education Policy May 2004



Center on Education Policy

1001 Connecticut Avenue NW Suite 522
Washington, D.C. 20036

Phone 202-822-8065

Fax 202-822-6008

E-Mail cep-dc@cep-dc.org

Web www.cep-dc.org

Pay Now or Pay Later

The Hidden Costs of High School Exit Exams



Center on Education Policy

MAY 2004

Table of Contents

Summary: The Hidden Costs of High School Exit Exams	1
<i>Keith Gayler and Nancy Kober</i>	
Part 1: Measuring the Cost of State High School Exams: An Initial Report	11
<i>Douglas Rose and John Myers</i>	
Executive Summary	13
Introduction	15
Chapter 1: The Task.	21
Chapter 2: The Method Chosen	25
Chapter 3: Procedures Used in Indiana.	29
Chapter 4: Indiana: The Cost of the GQE.	37
Chapter 5: Methodological Summary and Recommendations.	53
Chapter 6: Conclusion	57
Appendix I: District Panel Instructions, Current Costs	59
Appendix II: District Panel Instructions, Improvement Costs	61
Appendix III: Review Panel Instructions.	63
Appendix IV: Participants	65
Part 2: Comparing High School Exit Exam Costs in Minnesota and Massachusetts	67
<i>Douglas Rose with Justin Silverstein and John Myers</i>	
Executive Summary	69
Introduction	73
Chapter 1: Background and Expected Findings	75
Chapter 2: Method	79
Chapter 3: Minnesota	89
Chapter 4: Massachusetts.	103
Chapter 5: Comparing Costs, Programs, and Education Across States.	117
Chapter 6: Expectations and Results, Future Research, and Conclusion.	127
Appendix I: List of Panel Participants and Interviewees	131
Appendix II: New Bedford (MA) High School Services to Members of the Class of 2003 Who Have Not Met Their MCAS Competency Determination.	133

List of Tables

Measuring the Cost of State High School Exams: An Initial Report

Table 1-1: Indiana Panel Member Positions	31
Table 1-2: Indiana Public School Data	38
Table 1-3: 2003 Indiana Salaries	38
Table 1-4: GQE Pass Rates	39
Table 1-5: Per Pupil Local GQE Costs	41
Table 1-6: Local GQE Programs	41
Table 1-7: Per Pupil Local GQE Improvement Costs	49
Table 1-8: Local GQE Improvement Program Cost	49

Comparing High School Exit Exam Costs in Minnesota and Massachusetts

Table 2-1: Minnesota Panel Member Positions	83
Table 2-2: Massachusetts Panel Member Positions	83
Table 2-3: Process Evaluation	84
Table 2-4: Response Samples	84
Table 2-5: Minnesota Public School Data	89
Table 2-6: Estimated Average Public School Salary and Benefits in Minnesota, 2003	89
Table 2-7: Per Pupil Local BST Costs	92
Table 2-8: Local BST Programs	93
Table 2-9: Per Pupil Local GQE Costs	96
Table 2-10: Local GQE Programs	96
Table 2-11: Added Per Pupil BST Higher Standard Costs	98
Table 2-12: Local BST Higher Standard Program Cost	98
Table 2-13: Total Per Pupil BST Higher Standard Costs	99
Table 2-14: Local BST Higher Standard Program Cost	99
Table 2-15: Per Pupil Local Average State Exam Costs	100
Table 2-16: Local Average State Exam Cost	100
Table 2-17: Massachusetts Public School Data	103
Table 2-18: Estimated Average Public School Salary and Benefits in Massachusetts, 2003	103
Table 2-19: Per Pupil MCAS Exam Costs	105
Table 2-20: District MCAS Cost Programs	106
Table 2-21: Per Pupil Local Average State Exam Cost	109
Table 2-22: Local Average State Exam Cost	109
Table 2-23: Per LEP Pupil MCAS Exam Costs	111
Table 2-24: Per Massachusetts Student LEP Exam Costs	111
Table 2-25: LEP Programs for MCAS Exit Exam Success	112
Table 2-26: Per Massachusetts Student Costs for a Raised MCAS Standard	113
Table 2-27: Programs for a More Demanding MCAS Exit Exam	114
Table 2-28: Examination Difficulty and Failure, by State	118
Table 2-29: Teacher Salaries in Indiana, Massachusetts, and Minnesota	119
Table 2-30: Adjusted Costs of Exams, by State	119
Table 2-31: Costs by Exam Type and State, Adjusted for Price, Special Education, and Limited English Proficiency	122
Table 2-32: Expectations and Results	127

SUMMARY

The Hidden Costs of High School Exit Exams

By Keith Gayler and Nancy Kober

Key Findings

- **No cheap fix.** Although state policymakers may view exit exams as a low-cost way to raise student achievement, the extra costs of helping students pass these exams are considerable. To simply maintain the state's current level of exit exam performance costs an estimated \$171 per student per year in Minnesota, a state with an 8th grade exam; \$385 in Massachusetts, a state with a more rigorous 10th grade exam; and \$557 in Indiana, a state with a 10th grade exam of average difficulty.
- **Cost savings through coordinated reforms.** There are sound ways to lessen the costs of exit exams. Massachusetts has the most rigorous exit exam system of the three states studied, but its costs are in the middle range. This is probably because the exam is part of a well-integrated state system of standards, assessments, and accountability measures aimed at preventing general academic failure up front—thereby reducing the costs of exam-related remedial services later.
- **Hidden costs of remediation and prevention.** The direct costs of developing and administering the tests themselves make up a tiny fraction of the total costs of implementing an exit exam policy. The bulk of the costs go toward other “hidden” expenses necessary to give students a strong chance of passing the mandatory exams. These include remedial services for students who fail, programs to prevent failure, and professional development to upgrade the skills of teachers who must prepare students for the exams.
- **Invisible local costs.** The true costs of an exit exam policy are often invisible to state policymakers, because the expenses are being borne mostly by local school districts—and often by shifting existing funds away from other educational priorities.
- **Extra costs for special needs groups.** The costs of helping students with disabilities and English language learners (ELLs) pass exit exams are typically underestimated and often ignored. But these costs are significant, amounting to an extra \$26 for each special education student in Minnesota and adding \$101 to the overall per pupil costs of the Massachusetts exit exam for both ELL and non-ELL students.
- **Higher costs for higher performance.** The costs of exit exams in a given state go up substantially when the state seeks to increase its pass rate on the current exam, raise the cut score for proficient performance, or change the test to a harder one.

Lessons for Policymakers

- **Acknowledging local cost burdens.** States that have or are planning to adopt exit exams should acknowledge up front the cost burdens that these policies create for school districts and should consider ways to address these costs more equitably.
- **Starting out with rigor.** It is much cheaper to start with a rigorous exit exam than to change from an easier to a harder test at some future point.
- **Doing it right the first time.** Exit exams that are part of a comprehensive, adequately funded reform package, including early detection and prevention of student learning problems, are more cost effective than exit exams that are inserted into an unchanged system.
- **Remembering special needs.** States that have or are planning to adopt exit exams should make realistic estimates of the additional costs of helping students with disabilities and English language learners pass these tests.
- **Investing in higher performance.** Increasing initial pass rates without watering down the difficulty of the exam is a worthwhile goal from both an educational and a political standpoint, even if it does involve higher costs.

Background on the Center's Cost Studies

High school exit examinations—tests students must pass before they can get a high school diploma—are a popular state strategy for improving education. Currently, twenty states, enrolling more than half of all public school students, require their students to pass exit exams, and by 2008, four more states plan to introduce mandatory exit exams. Some governors and state legislators may see exit exams as a low-cost way to raise student achievement by spurring students to study harder and encouraging teachers to focus instruction on the important knowledge and skills every high school graduate should learn. After all, the thinking goes, the exams are just testing what students should be learning anyway, so they should involve few costs beyond test design and administration and some extra remediation for students who fail.

But recent studies commissioned by the Center on Education Policy suggest that this narrow view of exit exams costs is flawed and unrealistic. Few of the costs associated with exit exams involve actual testing expenses. Instead, the existence of a mandatory state exit exam leads to a shift in how local school districts use resources—for example, districts may shift their educational programming to devote more resources to remedial services for students who have failed the test or preventive services for students at risk of failing, and devote fewer resources to advanced education for students who have already passed. So when policymakers support exit exams, they are tacitly endorsing this shift and its accompanying costs, and when they oppose exit exams, they are opposing this pattern.

This is just one of several intriguing findings to emerge from two studies commissioned by the Center on Education Policy to look at the nature, type, and extent of costs associated with exit exams. The first study examined the costs of exit exams in Indiana, and the second looked at these costs in Minnesota and Massachusetts. Supported by a variety of foundations, this research on costs is part of our larger, ongoing national study of exit exams, underway since 2002.

Although costs are critical to every policy decision, we want to make clear at the outset that cost should not be the main driving factor for decisions about how we educate students. The cheapest fix is not necessarily the best, and a cost-benefit analysis cannot take the place of meeting our nation's moral obligations to our children. Nevertheless, our commissioned studies of exit exam costs suggest important lessons that can help guide state policymakers as they make crucial decisions.

Contents of This Volume

This volume contains the full reports of the two cost studies conducted for the Center by Augenblick, Palaich and Associates (APA),¹ a consulting firm with considerable experience in school finance issues. The first report, written by researchers Douglas Rose and John Myers, analyzes exit exam costs in Indiana. It was completed in February 2003 and formed the basis for a chapter of the Center's 2003 report, *State High School Exit Exams: Put to the Test*. The second report, produced in February 2004 by Douglas Rose of APA with Justin Silverstein of APA and John Myers, now of the JL Myers Group, looks at exit exam costs in Minnesota and Massachusetts and compares findings from these two states with those from Indiana. Keith Gayler, the Center's associate director, participated in and advised on the researchers' work, with assistance from Madlene Hamilton, research associate for the Center.

This overview chapter by Keith Gayler and Nancy Kober, a consultant to the Center, is based on the commissioned studies by APA. It looks across both studies to extract key findings about exit exam costs and lessons for policymakers.

Professional Judgment Method of Estimating Costs

Determining the full costs of an exit exam policy is no simple matter. Although states may have data on the state-level costs of developing and administering the tests and providing remedial help for students who fail, these costs are just the tip of the iceberg. The full costs of implementing a system of exit exams also include additional local expenses for remediation, plus other "hidden" costs necessary to give students a strong chance of passing the tests. These include the extra costs of programs at all grades to prevent student failure and raise test scores and to upgrade the test-related skills of teachers who must prepare students to pass the tests. Because school districts typically do not separately categorize or track these types of test-related costs, researchers must use other ways to determine what they are.

Our studies relied primarily on the professional judgment of expert panels of experienced service providers to generate cost data. In each of the three states, the APA research team assembled multiple panels of teachers, principals, curriculum and assessment specialists, superintendents, special education professionals, and finance directors. These panels were asked to specify the kinds and amounts of programs and resources that would be necessary for a hypothetical average school district in the state to achieve a particular outcome on the exit exam, such as maintaining the current average level of performance or raising the percentage of students who pass on the first try.

The panels were cautioned to identify only those extra resources that are necessary because the exam is a graduation requirement instead of an ordinary test with lower stakes. For example, the costs of designing the exams are not included, but the costs of admin-

1 Formerly Augenblick and Myers

istering retests are, because retest costs tend to be much higher with high-stakes exams. The costs of instruction to prevent general school failure are not included, but the costs of special, exam-related programs for students who fail a key exam in 8th grade are counted. Similarly, professional development activities are included if they are targeted on helping teachers prepare students for the high-stakes exam, and only that exam.

An expert review panel in each state scrutinized the work of the earlier panels and reconciled any differences in their conclusions. The APA research team then calculated the costs of the resources specified by the panels, using price estimates based on state educational cost data and approved by the review panel. The research team also collected additional information from interviews with various statewide education professionals and analyses of state data files. A more detailed explanation of the expert panels and other study methods can be found in the two papers.

Key Findings about Costs of Exit Exams

The two commissioned studies contained in this volume produced several key findings with implications for states that have or are planning to adopt exit exams.

1. Exit exams are no cheap fix.

The APA studies found that the costs of implementing exit exams are substantial, although they vary among states. To produce the current level of achievement on the Minnesota Basic Skills Test (BST) costs \$171 per student per year. This exam system is geared to an 8th grade level of learning in math and language arts and a 10th grade level in writing.

In Massachusetts, it costs an estimated \$385 per student per year to maintain the current level of student success on the exit exam component of the Massachusetts Comprehensive Assessment System. The MCAS exit exams are aligned to 10th grade expectations and are widely considered to be among the most rigorous in the nation.

In Indiana, the costs of maintaining the current performance level on the Graduation Qualifying Exam (GQE) were \$557 per student per year. The GQE is a 10th grade level exit exam of average difficulty.

These state variations reflect differences in the difficulty of the exams, their initial student pass rates, their degree of connection to the state's broader reforms, and other factors. Generally, the more difficult the exam, the higher its costs, because students must receive more instruction to pass it. Minnesota's BST, for example, is one of the less demanding state tests and involves relatively low per-pupil costs. By comparing costs across states, the research team concluded that in an average state, the cost difference between an easier exam (one geared to a 6th through 8th grade level) and a harder exam (one aligned to 10th grade standards) was about \$280 per student per year. But our commissioned studies also suggest that the most challenging exams do not always have to cost more.

2. The cost of a challenging exit exam can be lessened by implementing a coordinated system of standards-based reforms and assessments.

In light of the difficulty of the MCAS, one would expect exit exam costs in Massachusetts to be very high, but they fall in the middle range of the three states studied. A likely

explanation is that Massachusetts' exit exams are part of an integrated package of standards-based reforms and aligned assessments. Although the state just began withholding diplomas from students in 2003, the broader testing system and related reforms were put in place over several years. These reforms emphasize early efforts to monitor student performance, identify problems, and take steps to prevent academic failure. This approach enables school districts to target exam-related assistance on groups that are having difficulty passing the tests, rather than having to address basic education deficiencies for a considerable share of the secondary school population. So although the state has invested significant resources over the years in services to prevent student failure, most of these costs were not considered to be exam-related expenses by the Massachusetts panels. As a result of this emphasis on prevention and other factors, the state has a high initial pass rate for its exit exams, which translates into fewer costs for remediation.

3. Direct testing costs make up a small fraction of the total costs of implementing mandatory exit exams. The bulk of the costs goes toward remediation, prevention, and professional development activities to help students pass the exams.

In all three states studied, the direct testing costs of developing and administering exit exams constituted less than a fifth of the current exam costs—less than 7% in Minnesota, 9% in Massachusetts, and 18% in Indiana. In Minnesota, where the BST has been a requirement for many years, 63% of the estimated costs of maintaining the current level of exam performance goes toward prevention services, mostly for students who have had trouble passing earlier exams. These services include summer school, extended-day and half-size classes, and programs to improve the attendance of students with marginal performance, among other activities. About 29% of Minnesota's costs goes for remediation, while only 2% is for professional development.

In Massachusetts, the current exit exam costs are divided roughly equally among remediation (27%), teacher professional development (32%), and prevention (32%). In this state, prevention includes such expenses as tutoring for students who have failed earlier exams and prep classes for those in danger of failing. In Indiana, an estimated 29% of the current exam costs is for remediation, 25% for professional development, and 28% for prevention.

These different cost patterns seem to be related only in part to variations in the difficulty of the exams. Minnesota spends relatively little (\$3 per student) on professional development to help teachers prepare students for its basic-level exam, while Massachusetts spends \$95 per pupil and Indiana spends \$117 per pupil on professional development related to their more challenging exams. But Minnesota's exam is also not coordinated with its curriculum, making exam-related professional development unlikely to be productive. Similarly, Minnesota spends just \$38 per student on remediation, while Massachusetts and Indiana spend two to three times as much. This is probably because there are enough years between the time students in Minnesota first fail the 8th grade exam and the time they graduate to allow most of them to reach the passing threshold simply by maturing and completing their regular education programs.

4. The vast majority of the costs associated with exit exams are invisible and untalied, because they are being picked up by local school districts through reallocating existing funds.

In all three states, more than 96% of the current costs of exit exams is being borne at the local level. The state-level costs associated with exit exams are roughly \$2 per

student in Indiana and Minnesota and \$7 in Massachusetts. Although states may pass through appropriations for remediation to local districts, these state contributions tend to be small—about \$12 per student in Indiana, for example—not enough to cover the full local costs of remediation.

Local expenses related to exit exams tend to be invisible, because districts do not, and often cannot, track or tally them. Typically, districts cover these costs by shifting existing resources toward programs critical to exit exam performance and away from other priorities. For example, a district might increase the size of advanced math classes for students who have already passed the exit exam in order to free up teaching resources to provide extra instruction to students who have failed the math exam. Or a district might shift resources from a science class taught in two languages for English language learners to a program to help these same students acquire English language proficiency so they will pass the exit exam. In addition, many states allow students to take multiple retests throughout the year, which involves pulling out some students and teachers from their regular instruction and using substitutes to cover the teachers' classes. With a lower-stakes high school exam, these extra expenses would not arise. Exit exams also require guidance counselors to spend some of their time helping students sort through which tests they have and have not passed and what their testing options are—time that might otherwise be spent on college and career counseling.

This strategy of shifting resources only works up to a point, however, and in the current fiscal climate, districts are probably close to reaching that point. As explained below, the costs of *improving* student performance on exit exams are much higher than the costs of *maintaining* the status quo. With many states and districts facing large budget cuts, school districts will have difficulty finding additional resources to devote to increasing the percentage of students who pass.

5. The extra costs of helping students with disabilities and English language learners pass exit exams are often underestimated or ignored, even though these expenses are significant.

When determining the costs associated with exit exams, the state panels tended to underestimate the additional costs of preparing students with special needs for these tests—in fact, this topic was rarely mentioned by the regular panels. Nor were these costs calculated in the estimates for Indiana included in the first APA report. To fill these gaps, the APA study team assembled a special panel in Minnesota to consider the costs for students with disabilities and a similar panel in Massachusetts to determine the costs for English language learners.

The Minnesota panel concluded that the existence of the BST graduation requirement leads to extra costs of about \$26 per year for each special education student. Because special education in Minnesota already focuses on the types of 8th grade skills embodied in the BST, the marginal costs attributable to the exit exam are relatively low. In states with exit exams aligned to higher standards, however, the costs of preparing special education students would probably be greater.

The extra costs of preparing English language learners for exit exams are much higher, according to the special Massachusetts panel. The costs for these students add about \$101 per pupil per year, or about one-third, to the total costs associated with the MCAS exit exam for both ELL and non-ELL students. Well over half (60%) of these extra costs

for English language learners are designated for remedial programs, and a notable share also goes toward professional development to help regular classroom teachers meet the exam-related needs of students with limited English language skills, whether or not they are currently classified as English language learners.

Based on these findings about special needs students from Minnesota and Massachusetts, APA recalculated the costs of Indiana's exit exam system to include students with disabilities and English language learners. This recalculation raised the estimated costs of maintaining the current level of exam performance from \$442 to \$557 per student per year; the revised figures are contained in the second report in this volume.

6. Costs rise when states take steps to increase pass rates on their current exit exams, raise the cut score for proficient performance, or change to a more challenging exam.

The APA studies also considered what it would cost states to make various kinds of changes in their exit exam policies. In Minnesota, a set of panels estimated the resources required to move from 8th grade exams in math and language arts to 10th grade exams in these subjects, while maintaining the current pass rate. The panelists concluded that these harder tests would bring *extra* costs of \$377 per student, tripling the cost of the BST.

In Massachusetts, a set of panels considered the costs of raising the minimum MCAS score required for graduation from the “needs improvement” level, where it currently stands, to the “proficient” level. The panelists concluded that this change would add exam-related expenses of \$575 per student per year, more than double the costs of sustaining the current level of MCAS performance. Most of these additional costs would go toward prevention and professional development and would be targeted mostly at students with special needs.

The Indiana study determined that to raise the pass rate on the current exit exam would add \$685 per pupil per year to the costs of the exam, a 150% increase.

In sum, more difficult exams or more stringent cut scores bring higher costs, because larger numbers of students need extra help to learn more so they can surmount the passing hurdle.

Efforts within a state to reduce the percentage of students who fail an exit exam on the first try—in other words, to raise the initial pass rates—are the most costly of the changes analyzed by the study team, but among those with the greatest benefits in the long term. Reducing initial failure rates from a high to a moderate level, or from a moderate to a low level, would increase exit exam costs by about \$380 per student in any state, the team concluded, because this strategy involves a large dose of prevention.

Lessons for Policymakers

Exit exam costs are a pertinent and timely issue for state policymakers for several reasons. First, many states are dealing with budget crises, and a reality check of the cost implications of exit exams could be helpful as states debate critical budget decisions.

Second, many state policymakers, as well as state courts, are looking deeply into the issue of what it costs to provide students with an adequate education—a concept that has gained urgency due to the number of lawsuits challenging state school finance systems. As part of

these efforts, policymakers, judges, attorneys, and others are taking a close look at exit exam requirements. For example, a recent study by the American Institutes for Research (AIR) and Management Analysis and Planning, Inc. (MAP) defined an adequate education in New York State as the opportunity to acquire the knowledge specified in Regents Learning Standards—an outcome measured by whether students pass the Regents exit exams¹.

Third, most states are in the process of changing their testing programs to comply with the federal No Child Left Behind Act (NCLB). Although the Act does not require states to have a mandatory exit exam, it does require them to test students at least once in high school, as well annually in grades 3 through 8, and to ensure that students in all subgroups and every grade reach the state's benchmarks of proficient performance by 2014. The NCLB goals are likely to be more difficult—and probably more costly—to reach than the more limited objectives of state exit exams. But understanding the costs involved in preparing students to pass exit exams may help states make decisions about testing and resources under NCLB.

Several lessons for state policymakers can be drawn from the two studies collected in this volume.

1. State policymakers should acknowledge the cost burdens that these policies place on school districts.

Many state policymakers seem to view exit exams as a relatively low-cost reform, as indicated by the small amount of state funding earmarked for exit exam programs. Few of the states surveyed by the Center for our 2003 exit exams report said that they targeted funds specifically for exit exams, except for some limited appropriations for remediation.

Although exit exams are a state-mandated reform, targeted state appropriations cover only a minuscule portion of the total costs. Consequently, most of the costs associated with these exams are borne by local school districts, primarily by shifting existing resources. Because districts do not track these reallocations, it is easy for state policymakers to ignore the reality that districts must take from Peter to pay Paul. A fairer and more realistic approach is one that acknowledges the budget pressures state exit exam mandates put on school districts and considers ways to address these costs more equitably, or at least more honestly.

2. It is cheaper to start with a rigorous exit exam than to change from an easier to a harder test at some future point.

States sometimes start out with an easier exam, in order to give teachers and students more time to adjust to the policy and to raise student achievement gradually. The goal is to get the exam into place quickly, then increase its difficulty over time. Our commissioned studies suggest, however, that it is less expensive to start out with a challenging exam than it is to make the exam more difficult later.

In Minnesota, the cost estimates were quite high for the scenario of changing from the current 8th grade exit exam, which is not well aligned with state content standards, to a 10th grade exit exam that is aligned with state standards. Massachusetts took a different approach by putting in place a well-developed system of standards-based reforms and

1 American Institutes for Research and Management Planning and Analysis, Inc. (2004). *The New York Adequacy Study: Determining the Cost of Providing All Children in New York An Adequate Education*, Volume 1, Final Report.

challenging assessments before it began withholding diplomas. Now the state has high initial pass rates. Doing a reform the right way from the beginning—well-funded and linked to a comprehensive standards-based system—may lead to better outcomes and lower costs over time.

3. Exit exams are more cost effective when they are nested in a comprehensive, integrated, and adequately funded reform package that includes early activities to detect and address students' learning problems.

The experience of Massachusetts, which has moderate costs for its current exam, suggests that it is relatively inexpensive to implement an exit exam if it is part of a comprehensive and well-funded system of standards-based reform. Using an integrated approach that includes early efforts to monitor students' progress and prevent academic failure seems to be more cost effective than inserting an exit exam into a system before broad reforms are in place.

4. States should pay more attention to the costs of helping special populations pass exit exams.

Educators involved in the APA studies tended to overlook the resource needs of students with disabilities and English language learners. Yet the exam-related costs for these groups can be quite high, particularly if the test is geared to higher standards than the students' instructional program, as would be the case if Minnesota moved to a tenth grade exam. In Massachusetts, the cost estimates for English language learners are high, in part because the goal of getting a high school diploma is very important to these students and their parents. The costs associated with preparing these students also depends on whether the exam system includes testing accommodations or different standards for some students, and whether it includes waivers and appeals processes that allow students with special needs to receive a diploma without passing the exit exam. In these situations, educators may limit the amount of extra test-related assistance that these students receive.

5. Increasing initial pass rates while maintaining the difficulty of the exam is a worthwhile goal, even if it does bring higher costs.

High initial pass rates are attractive for both political and educational reasons. When the percentage of students passing on the first try is low, people are likely to raise questions about the fairness of the test and the adequacy of the educational system. Low initial pass rates can also lead to higher dropout rates. Although it is expensive to increase students' initial pass rates on the exams without “dumbing down” the tests or lowering the cut scores, this is an important enough goal to merit an investment of state resources.

Our commissioned research on the costs of exit exams also has implications for the No Child Left Behind Act, the major force shaping decisions about testing. Although the costs of reaching the NCLB goals are likely to be much higher than the costs of raising performance on exit exams, the types of resources involved are quite similar, encompassing prevention, remediation, and professional development. If states choose to spend resources to improve performance on exit exams, this could double as an intermediate step toward fulfilling the NCLB goals. As with exit exams, it may be wiser to get the NCLB testing system right the first time, rather than rushing to put something in place that will have to be changed later at greater expense.

In conclusion, states with exit exams should take into account the costs of providing students with meaningful opportunities to pass the exams. States that are phasing in or considering adopting exit exams are well advised to make the necessary investments from the beginning, instead of forcing school districts to pay for them later.

PART 1

Measuring the Cost of State High School Exit Exams

AN INITIAL REPORT

By Douglas Rose
*Augenblick, Palaich, and Associates*¹
and John Myers
JL Myers Group

For the Center on Education Policy
FEBRUARY 2003

1 Formerly, Augenblick & Myers

Executive Summary

In this report, Augenblick, Palaich, and Associates (APA) presents to the Center on Education Policy (CEP) a new method for estimating the cost of high school exit examinations and a report on the application of that technique in the test case of Indiana. The method succeeded in producing cost estimates for both current levels of performance and state-sanctioned higher performance targets. APA recommends modest revisions in the approach before it is applied again.

Because the exit exam culminates the education that precedes it, the methods used to study its costs should reflect the characteristics most likely to influence its cost. Exit exams so far are state-centered and serve as a unification point for the state education system. Hence, the methods to study their cost should be primarily state-centered rather than comparative. Similar arguments are presented here about why the study should include indirect costs for which the exit exam's role is clear, should focus on the costs of the "exit" aspect, and should include the costs of meeting higher standards as well as the cost of current performance. In turn, these adaptations of scope to substance support attention to current and future programs, not merely past ones, and the systematic use of interviews to supplement documentary evidence.

The method chosen by APA to study exit exams reflects the broad methodological choices dictated by the nature of the exams. The core technique is the use of diverse panels of experienced service providers to estimate the resources needed because of the Graduation Qualifying Examination (GQE) in a hypothetical school district. This "professional judgment" approach is adapted from APA's studies in various states of the cost of an adequate education. Along with the district panels are included supportive techniques such as a nominating committee and a statewide review panel. Interviews with key personnel and formal documentary evidence provide the basic information with which the panels can operate and with which APA interprets the panel results.

By agreement with CEP, we selected Indiana as the test case and, as agreed, applied the method there. The method succeeded in producing estimates of the costs of current and improved levels of performance, estimates that appear to be reliable within a 5-10% range. However, efforts to link cost estimates to programmatic features of exit exams or to specific resources did not turn out as well. While most procedures should be maintained intact, specific problems suggest a series of additions and revisions. Recommended revisions in the process include giving each panel a single task but offering more variety of tasks across panels. Among the possible tasks would be one or more of the hypothetical changes in exam structure, the use of a scenario that would be identical in every state, and topics studied only once in the set of states, such as the shifting of existing resources toward exit exams. We suggest that late September to mid-November and mid-February to mid-April would be better times to conduct panels.

The substantive results in Indiana indicate that the current costs of the GQE are \$444 per pupil per year. The costs are sizable, given the \$8,128 spent per pupil in 2001-2002, but generally invisible. The costs primarily occur for local programs, notably school personnel,

rather than the highly visible state programs. The \$442 million cost of the exit exam, while large, may not be highly visible in an \$8.1 billion total school budget. In addition to testing and remediation, professional development and failure prevention account for one-half of the costs. The largest share of the costs appears to be met by superintendents and principals shifting funds that are available from the state as per pupil block grants toward the students and programs critical to GQE outcomes and performance.

The costs for improving GQE performance to the state-mandated “commendable” level are estimated at 150% above current costs for the GQE, \$685 per pupil per year, for a total of \$682 million. As with the current costs, local programs account for almost all of the needed resources, and school-level teachers are the top cost item. Unlike the current costs, these could not be met by shifting resources. However, these programs are less specific to the exam than are the current costs, emphasizing education to increase the initial pass rates and emphasizing professional development, rather than remediation and testing. On the whole, the programs to improve scores follow the scenario of making interventions at critical points in the current education process in order to make permanent changes that allow students and teachers to improve the ordinary education process. They yield improved test performance because they improve education performance more generally.

Introduction

This study reports on work done by Augenblick & Myers for the Center on Education Policy during the fall and winter of 2002–2003. The aim was to design and produce original policy research. The design task has been to create a method for estimating the cost of conducting state high school exit examinations, while the research task has been to test out that method in Indiana. The test shows that our basic method worked well, and we learned enough to suggest some changes.

The Purpose of the Study

The growth of high school exit examinations arises from the standards-based reform movement in the states in the 1990s. Standards-based instruction clarifies what students should know, exit exams are a measure of whether they know it, and accountability provides schools with an incentive to help students prepare themselves. States have had to design and implement exit exams without research-based information about the costs of their choices. This study aims to take the initial steps toward providing information about the cost of exit exams.

High school exit examinations are high stakes tests for students, public school systems, and state education systems. Examinations have grown in popularity, and the capstone test in twenty-four states is or will be the high school exit examination. For students, the direct consequences of the test include failure to receive a high school diploma or substantive investment of time and effort in remediation. The failure rate of students on an exit exam can determine the reputation of the public school system in a district and, in some cases, can determine the allocation of resources within the system. For the states, an exit exam represents an investment in an educational strategy, one that can be used to determine the state's flow of resources.

This study focuses on the costs of these important exams that have been increasingly adopted in states. Thus far, the costs have been unknown and uncertain. In the absence of cost estimates, decisions have depended on the general desirability of the idea and goal of exit exams, on the one hand, and the undesirability of reliance on a single examination, on the opposing hand. Once exams have been adopted, standards for success have been selected without clear information on the resources needed to achieve those standards. One aim of this study is to develop a way to allow more fully informed decision-making, both by education policymakers and by the public. Additionally, the goal has been to specify the components of cost, in order to allow planning by budget-makers, departments of education, and school districts for how to meet the costs of high school exit exams.

The first step is to develop a method for estimating cost. While states vary in the form of their responses, they face similar problems involving similar resources and trade-offs. The purpose of this study is to develop a methodological framework within which the cost of each state's unique response can be estimated and common patterns clarified. Policy choices informed by research based on carefully considered methods are, in our experience, more likely to produce the expected consequences and have fewer unintended outcomes. This

report is thus largely about methods. Although the methods may continue to evolve, this first step is when they most need investigation and discussion. The second step is to test the method, evaluate and refine it. Method and substance support each other. The development of the method has as its purpose the illumination of policy choices, so part of this study is a review of what the costs of exit examinations are in an initial state. While the costs in other states are unknown and a single state is not a sufficient base for conclusions, the results in the initial state serve to suggest possible patterns to look for elsewhere.

This study does not address the costs of implementing the No Child Left Behind (NCLB) Act, though exit exams are part of the same broad movement toward standards-based assessment. NCLB does not require an exit exam, its consequences for schools start well before high school, and its standards for performance (100% pass rates) imply a quite different exam and preparation than does an exit exam. Hence, an adequate preparation for an exit exam and an appropriate implementation of NCLB might be quite different, and their costs would be the same only by coincidence.

The History of the Project

APA is a Denver-based consulting firm that specializes in providing technical assistance primarily to state-level policy makers, particularly legislatures and state education agencies, around education finance, governance, and school improvement issues. APA has developed expertise in analyzing the equity and adequacy of school finance systems, linking finance to accountability.

In the 1990s, APA conducted several studies for the Department of Education and the Department of Children, Families and Learning in Minnesota on the impact of the state's testing and graduation standards.² The cost of the high school exit examination that was about to be implemented was part of that work. These studies of the resources needed for such exams prepared APA to complete further studies in this area.

Based in Washington, D.C., and founded in 1995 by Jack Jennings, the Center on Education Policy is a national independent advocate for public education and for public schools that are more effective. In 2002, CEP published the baseline report in a three-year study of high school exit exams that was funded by the Ford Foundation.³

One outcome of CEP's initial investigation was the desire to conduct research into the cost of high school exit exams. CEP contacted APA to plan and test a method for estimating the resources needed for high school exit exams.

From the start, CEP and APA have worked to develop a method that could estimate both the current costs related to the exams and the costs related to providing students with a substantial chance of passing a test that measured what they should have learned.

APA proposed and CEP agreed to the following tasks for APA.

2 Augenblick & Myers, "Description of the work done in Minnesota to estimate costs associated with new high school graduation standards," November, 1997.

3 Naomi Chudowsky, Nancy Kober, Keith S. Gayler, and Madlene Hamilton, Center on Education Policy, *State High School Exit Exams: A Baseline Report*, Washington D.C., 2002.

1. Organize the framework in terms of the costs of programmatic choices faced by decision-makers in a state. Whenever possible, cost estimates would be specific to the program characteristics and will be separated into phase-in and recurring yearly costs. The cost estimates represent the experiences of other states that need to be adjusted by means of the use of standard units such as teaching days, average daily membership (ADM), average salary, number of schools, proportion of at-risk students, and 2001-2002 school year dollars. Attention will be given to the changing cost of providing resources for changing policies.
2. Develop a methodology appropriate to the states that have exit exams now and those that might have one in the near future. To identify the significant features and variations of state arrangements, APA would draw on CEP's 2002 report, plus interviews conducted in the winter of 2002-2003. The focus would be on the state variation in: policies on accountability, remediation, and retakes; test examination formats, alternatives, pass rates; and extent of implementation, grades tested, and coordination with other tests and the curriculum. APA would also draw upon the description of cost elements included in our report on Minnesota's exit exam. These include developing professional skills, testing, recording, remedying, informing the public, shifting resources, and saving on related expenses. Attention would be added to developments since the Minnesota report was written, notably the emphasis on a state's providing adequate resources for educational proficiency.
3. Apply the methodology to one state in order to test and illustrate the approach. Drawing upon the existing financial data and conducting interviews with those charged with administering aspects of the exit exams, APA would describe the cost of providing students an adequate opportunity to pass a high stakes exam. The case study state selected should represent typical state exit examination arrangements, provide adequate data to make estimates, have fully experienced the exit exam process, and offer feasible arrangements for adequate interviewing within the budget.
4. Report on the methodology and its application to one state within the programmatic choice framework, explaining the procedures and the reasons for them, and using the case study as an illustration.

This report covers those tasks. Because the high school exit examination is historically and programmatically included in a broad set of standards-based testing and teaching, the estimate of costs can only be approximate, rather than representing a precise, unarguable fact.

The team conducting the study has been the following.

John Myers, head of JL Myers Group, oversaw the development and application of the methodology, visited Indiana, conducted panels, interviewed key personnel, provided coordination with Jack Jennings of CEP and top officials in Indiana, and edited the report.

Douglas Rose, a partner in APA, developed the methodology, researched background statistics, visited Indiana, conducted panels, interviewed key personnel, provided coordination with Keith Gayler, Associate Director of CEP, and drafted and presented the report.

Katharine Christensen, an associate in APA, provided office assistance, arranged travel, visited Indiana to provide support, and coordinated interviewing schedules.

Keith Gayler, in addition to his role for CEP, visited Indiana where he helped to conduct the panels and interviews, as well as provide timely information and insight.

State Exit Exam Approaches

Broadly, there are three approaches to exit exams that we term the testing, remediation, and prevention approaches. Although there are many differences among state exit exams, the most important cost differences might trace to the broad approaches.

Testing emphasizes the actual testing of students. Supportive activities would include test development, administration, record keeping, creating alternative exams, providing accommodations for disabilities, training for administration, and arranging for retests. The state would provide a test based on the standards to be taught, and students are responsible for passing the exam. A state assessment office would bear the budgeted costs of the exam, while students would bear the largely invisible costs of their own efforts. The exit exam would serve as an incentive for marginal students to learn the material.

Remediation emphasizes efforts of service providers to move students from an initial failure on the exit examination to ultimate passage and graduation. Supportive activities would include summer school, remedial classes, tutoring, pre- and after-school programs, emphasis on test-taking skills, development of a remedial curriculum focused on core skills, adopting remedial software, learning to translate test evaluation into appropriate instruction, and alternative routing to demonstration of skills. The state and local district and school would be accountable for providing what students need to be able to pass the exam. The exit exam would serve as an incentive for schools to improve the performance of marginal students.

Prevention emphasizes changes in how instruction is conducted in order to make sure that students learn the skills they need in order to pass the examination. Supportive activities would include training teachers to better address the needs of special education students, targeting programs at students who have failed exams in earlier grades, cumulative record keeping on individual student skills and test diagnoses, emphasizing early reading, adopting teaching technology aimed at problem students, and developing techniques to present standards-based material to special education, limited English, and at-risk students.

From a cost point of view, the three approaches cumulate in part. Testing costs are likely to be maintained at the same level under all three approaches. The costs added by the remediation approach would be maintained under the prevention focus for those students who initially fail—but the aim of the prevention approach is to reduce the number initially failing, hence the total remediation cost. In general, the testing approach should be least expensive and the prevention approach the most expensive, with the testing approach producing the least change in individual student learning and education service while the prevention approach would yield the greatest individual and systemic change.

States may vary over time in their approach. The early steps might focus on testing, especially for exam proponents who hope for an inexpensive improvement in education by increasing student motivation. As a state confronts the reality of withholding diplomas from students who have failed the exam, the remediation approach might become more pronounced, perhaps speeded by a lawsuit. As the exam becomes institutionalized and school districts become accountable for results, service providers may move toward the prevention approach. The costs might be expected to change over time as well.

In setting standards to which schools and districts will be held accountable in the long term, policymakers might wish to be able to anticipate not just the costs of testing and not just the costs of remediation but also the costs of prevention. In estimating cost, current costs

are part of the costs of meeting the standards set. From a policymaker point of view, the two sets of costs are the costs of “what is” and the costs of “what if.” The costs of the current arrangement, or “what is,” are already being paid for, while new funding would have to be found for the costs of meeting the performance standards, or “what if.” We study both.

An Outline of the Report

The following three sections discuss methods, proceeding from the general to the specific. The next section considers the larger framework of the options available for studying costs, including both what to study and how to study it. Within these options, the nature of state exit exams influences APA’s choice of methods to be used to estimate costs. APA’s methodological choices in turn shape the specific methods, discussed in the succeeding section. The central method adapts a “professional judgment” approach to estimating costs and adds a series of supplementary techniques. One is to use two different scenarios for two sets of panels, one focused on current costs and one focused on the costs of improvement to meet state-set levels of performance. The third methodological section reviews the techniques actually used in the test case of Indiana and includes an evaluation of the successes and topics for improvement. The technique produced the sort of specific estimates intended, but the experience suggests refinements that could be used in future studies.

The section on the cost of the GQE in Indiana provides the substantive findings. The unique characteristics of Indiana and the particular instructions to the panels shape the cost estimates. After a background review of the GQE, the section discusses the hypothetical district and test results for which panels of experts provided estimates of the required resources. The initial set of cost estimates is for the current costs of the GQE, or “what is,” and includes costs of ongoing local programs, one-time costs, resources shifted from other programs, and state programs. For state programs, a distinction is made between state-provided programs and state funding for local programs. The second set of cost estimates is for the costs of improved performance on the GQE, or “what if.” These costs are entirely for local programs, include more start-up expenses, and do not involve many shifted resources. Even though influenced by the particulars of Indiana and the premises provided to panelists, the cost estimates may also reflect patterns that could be found in other states.

The final section summarizes the methods that will be continued, summarizes the cost estimates for Indiana, and offers recommendations for improved techniques. Recommendations include the use of an identical scenario in each state studied in order to ease cost comparisons across states. The study ends with a conclusion that suggests using the method in a study of multiple states. Appendices presenting the instructions given panels and the names of the participants in the study follow the conclusion.

CHAPTER 1

The Task

Developing a method to estimate the resources needed for high school exit exams involves facing choices about what to study and how to study it, the subject of this section. These tests are unique in being both capstone exams, so that resources expended to pass them can occur well before high school, and high stakes exams, so that resources may be expended on a series of retakes and remedies. The combination of early prevention and subsequent remediation costs makes the approach to costing out high school exit exams different from the approach to costing out other standardized exams.

Issues and Choices About What to Study

Single-state or Comparative?

In focusing on a single state, the state's education system and history form the main context of interpretation. Causes can be identified with historical changes in the system. Because changes come in bundles, it is difficult to disentangle a package of causes and effects. For instance, the exit exam system is often changed in many aspects at once, leaving obscure which aspect produced the observed costs. In a cross-state study, the context is formed by the variations in the aspects of the exit exams. We look for similar exam provisions to produce similar effects. Though states differ in many ways, differences in exit exams provide the privileged account for differences in costs.

APA has chosen to emphasize the single-state focus. In part, this is inevitable as the actual format, content, and difficulty of the exams differ, along with the standards being taught and the cut points for passing the exam, as do also key elements of costs, such as teacher salaries. In part, this reflects the early stages of the project, as we are not ready conceptually for a cross-state study—not ready to make definitive expectations of which elements of the exams will be most important, for instance. State-focused studies, as case studies, can prepare for cross-state comparisons later on. Having taken this primary focus, however, we intend to conduct the studies in a way that permits cross-state inferences later on. For instance, we collect data on resource use separately from data on the cost of that resource. At a later stage, states can be compared in their use of resources even if their prices differ. A second technique to permit cross-state inference is to use hypothetical questions about the impact of changes in major exam variables—what would be the cost of using open-ended questions, for example? This permits comparing costs by the exit exam characteristics across states.

Direct Costs or Also Indirect Costs?

The costs of an exit examination may be arranged by closeness to the test. At the near end are the costs of administering the exam that are close to the test because they are solely attributable to the test, occur at the same time as the test, and have a budget. At the far end are the costs of revising the education curriculum in state colleges, because these revisions would be

influenced by many factors, among them the national push to standards-based instruction, state tests, and—in states that have them—exit exams for graduates who expect to teach in the state; and the changes do not immediately follow from the adoption of exit exams but come about slowly and in piecemeal fashion, without benefit of a concise budget. If the more faraway effects are counted as costs, the total estimated cost of exit exams increases, possibly too much. If only proximate costs are counted, however, some inevitable aspects of the costs of exit exams may be missed.

APA has leaned toward including two types of less direct costs: those lacking a clear budgetary record and those occurring at a later time than the establishment of the exam system. In part, these choices reflect our understanding of how state education systems operate. The linkage between budgets and programs appears at first glance to be more direct than it is on further acquaintance, and recent changes have been in this direction of providing resources and demanding accountability, rather than in bureaucratically specifying the use of resources in each system. Given the diminished bureaucracy and the emphasis on professionalism, the indirect consequences in a change such as the adoption of an exit exam take time to develop. They do develop, however, because public education is among the most systematic sectors in America, with changes in one major element—such as high school standards for graduation—inevitably affecting every other major element.

APA decided against including effects where the exit exam is not identified as the primary cause of a cost or where its share of causation is unclear. Precisely because causes tend to be bundled together, costs might be equally attributable to any one of them—exit exams, increased testing in lower grades, teaching to standards, district accountability, etc. In this situation, we are reluctant to attribute costs to the exit exam. For joint costs of a clearly delineated program—as with the assignment of the exit exam’s share of an assessment budget—we will attribute costs in part on the basis of costs where shares can be identified, as with the costs of publications on exit exams and other topics.

What Is or What Ought To Be?

The costs of the current exam arrangements and results are distinct from what the costs would be were the state to achieve its goals for the exit examination. The case for measuring current costs is that measuring hypothetical costs is difficult. The case for measuring the costs of future success depends on the purpose of the exit exams. Legislatures adopt exit exams to improve education, not simply to measure what already exists. In this sense, the cost of exit exams ought to be the cost of their success: if scores did not improve, why have the exams? At a minimum, the cost of current efforts and results should be distinguished from those needed to be more successful. Many states have some official standard of adequate performance—for individuals, for schools, for districts, and for the state as a whole—below which improvement is demanded and above which future improvement has no impact. These states thus provide a standard for performance under which costs can be assessed. An alternative would be to use a single national standard, such as NCLB’s goal of 100% passage. Because state exams vary and were not designed for 100% initial pass rates, the national standard may offer more the appearance than the reality of a single standard.

From the start, APA and CEP have worked to develop a method that could estimate both the current costs of exams and the costs of an arrangement in the future, where students would have a substantial chance of passing. This seems to us to be necessary for decision-makers and educators to make informed choices about standards, the required level of resources, and the best allocation of resources.

Exit Exam or High School Exam?

Should the entire costs of the high school exit examination be included, or should the cost of a high school low-stakes test be taken out? Given that all states will be required to have a high school examination of some sort under NCLB, the cost of a high school exit exam may be considered as something in excess of a minimum exam—either an additional test or additional costs associated with a high stakes test. While that is a clear standard for costs, its disadvantages are twofold. First, because the adoption of exit exams preceded NCLB, state legislatures in exit exam states may opt for low-cost NCLB supplements to the exit exam. As a result, costs may be attributed to the NCLB tests that are properly accounted for by the exit exam. Second, there is no easy way to separate the costs of an imagined test with low stakes for the student from the actual high stakes test.

APA chose to focus on the costs of the exit aspect of the high school exit examination. Whatever the past events, states now have options of a non-exit exam, an exit exam also meeting the NCLB standard, or separate exams for separate purposes. Focusing on a separable exit component allows us to assess a state's cost under current choices and to make comparisons across states.

Choices About Information to Trust

Historical or Current Data?

By examining historical records, particularly around the time of the adoption of a major change in the use of the exit examination, the cost of the test can be seen in the increase in resources used. This is especially true if the program is easily separable from other programs, has a separate budget, and is implemented through a single state agency. The advantage of historical data is that they are less disputable and can be matched in time to known policymaking events. The disadvantages are that the data cannot be controlled or easily added to when they are insufficient and that other, complicating causes in changes in expenses may be happening simultaneously—time itself a cause. Because many state exit exams have been phased in over time, historical records, especially of school and district-level expenditures, can be more difficult to match to purported exit exam causes. Because high school exit exams have been part of a broader movement toward increased testing, it can be difficult to separate out the exit exam from the other tests, especially when efforts have been made to link them together and tie them to curricular changes. Additionally, historical data provide at best a hint of what would be required to meet official standards of performance in the future.

By examining information about current programs, the needed current resources can be specified, most easily in the case of single agency programs. The advantages of current information can include the control over the data collection and design, information about the cost of currently preferred alternatives, and a diversity of data types and sources, especially interviews. The disadvantages include the absence of a clear baseline against which to measure current costs, producing difficulty in estimating what share of current costs are due to the exit exam, a problem we address later in this report. APA has chosen to rely primarily on current information about programs, supplemented by historical information, especially about state level costs. The focus on current information allows the study of costs separated in time from the exams and that are not directly accounted for in official budgetary records.

Official Documents or Also Interviews?

Should sole reliance be placed on existing documents of costs and resource use, or should additional reliance be placed on collecting new information by way of interviews with informed sources? Existing documents have the advantage of being unobtrusive measures—we do not change or influence estimates in the process of studying them. Moreover, they usually follow conventions and standards that are understood and important, as in the case of budgets, and have been subject to scrutiny by informed and sometimes suspicious observers. However, existing documents rarely cover all the desired information, and the contexts in which they are developed and used take study before the documents can be understood and used as cost data. Interviews can provide an additional array of desired information, including information about hypothetical situations, intentions and causality. Disadvantages of interviews can include finding the best people and getting their agreement, having interviewees choose what to say to promote their aims, and resolving disagreements among interviewees.⁴

APA placed reliance on interviews as well as documents for this study. Documents are ill-suited to the study of local costs of exit exams. In our experience, educators are among the best subjects for interviews. Because of our emphasis on hypothetical standards and possible future costs, interviews are the only source of data on key topics. By bringing interviewees face-to-face around a table and focusing on a common topic, we plan to avoid having to resolve disagreements ourselves.

One type of cost likely to be underestimated in interviews is one-time, nonrecurring costs associated with the development of a program. People tend to “telescope” in time, which has the effect of understating costs that are not current or ongoing. For one-time costs, documents are primary sources. Conversely, resources that are shifted from another purpose in order to cover the needs of exit exams can only be easily studied with interviews. Given resources that increase more slowly than demands, educators have shifted resources toward high priorities. We suspect that high school exit exam performances are among the very highest priorities for state and district officials and expect a corresponding shift in how resources are deployed. There is no use of official documents that could distinguish GQE shifts from resource shifts for other reasons. For instance, if all new software purchases are allocated away from 11th and 12th grade toward middle school, this might be due to a strategy for exit exam success, hence a cost of exit exams, or might be due to a perceived relative surfeit of high school software, and the way to resolve the issue is to ask.

In sum, the characteristics of the exit exam influence the choice of method. The difficulty in separating the historical costs of exit exams from the costs of related reforms enacted in the same period, as well as the state-normed contents and standards of the tests, plus the influence that capstone high stakes testing has on the rest of education, and finally the change-oriented aim of the exams, the undocumented decisions about choices of how to allocate district resources—all these factors lead to a cluster of related choices about what to study and how to study it. A single method should embody these preferences. The next section discusses that method.

4 One possibility would be to survey service providers, as do Lauress L. Wise et al., Human Resources Research Organization, *California High School Exit Examination(CAHSEE): Year 3 Evaluation Report*, June 28, 2002. However, we use interviews to find the best information, rather than to represent a population, so a survey is ill suited to this research task.

CHAPTER 2

The Method Chosen

The choices among what and how to study shape the methodology applied by APA to study the costs of high school exit examinations. The focus on a single state, the reliance on interviews, and the concentration on current arrangements together lead to a method of going into a state and conducting interviews and group meetings to learn all aspects of exit exam costs, their causes, and their effects. CEP has already done a historical and comparative cross-state study based primarily on existing documents. APA can rely on this broad background, and CEP's case studies, in designing single state investigations. The single state interviews also aid in investigating the cost of achieving standards in the future, separating the "exit" aspects of exam costs, and accounting for indirect effects and those that take time to develop. This section presents the techniques that match the method.

Our method of using interviews to achieve an overall portrait of state costs and investigate hypothetical questions follows techniques developed by APA and other investigators in school adequacy studies.⁵ While four approaches exist, the professional judgment approach is the most commonly used.⁶

The professional judgment approach relies on the views of experienced service providers to specify the kinds of resources, and the quantities of those resources, that would be expected to be available in order to achieve a set of specified objectives (where the objectives are not determined by the service providers). The approach uses multiple panels of "experts" to specify the way education services should be delivered in prototypical schools and school districts. Once the services have been specified, with a focus on needed numbers of different types of personnel, costs are attached and a prototype per pupil cost is determined.

This approach best reflects the experiences of people who are actually responsible for delivering education services and may be combined with research results as the basis of a rational way to specify the magnitude of resources that are expected to produce some level of results. The advantages of the approach are that it reflects the views of actual service providers and it is easy to understand; the disadvantages are that it tends to be based on current practice. Additionally, the approach shares with all approaches the disadvantage that there is little evidence that the provision of money in designated amounts, or even a speci-

5 For papers discussing adequacy approaches and results, see studies listed at <http://www.aandm.org/papers.htm>. A useful initial study is A&M's "Calculation of the Cost of a Suitable Education in Kansas in 2000-2001 Using Two Different Analytic Approaches, Prepared for Legislative Coordinating Council, May 2002. See also the most recent of a series of studies in Wyoming by MAP, "Proposed Revisions to the Cost Based Block Grant," James R. Smith, Management Analysis & Planning, Inc., Submitted to the Wyoming State Legislature, January, 2002

6 A discussion of the advantages and disadvantages of alternative approaches, along with a review of their application in one case, may be found in "Enabling 'Adequacy' to Achieve Reality: Translating Adequacy into State School Finance Distribution Arrangements." Guthrie, J.W., Rothstein, R., Ch. 7, *Equity and Adequacy in Education Finance: Issues and Perspectives*. Ladd, H. F., Chalk, R., Hansen, J.S., eds. Washington, D.C.: National Academy Press, 1999.

fied deployment of resources, will produce the indicated outcomes. Based on the adequacy experience, the cost estimates appear to be reliable within approximately 5–10%. That is, panels in the same state with the same task but different members and a different moderator on a different day will often differ from the average by between 5 and 10%.

The professional judgment approach forms a base for our method. By bringing together panels of informed educators and posing a single core problem in a hypothetical district the method uses their diverse experience and expertise but brings it to a sharp and unified focus on the problem of interest, high school exit exam costs. Further, panels provide detailed information on what programs and resources are needed to accomplish the goal.

The rubric of professional judgment covers differences in the methodology of the adequacy panels and the exit exam cost panels. Where the adequacy panels tackle a single, future-oriented, broad task of designing an adequate education, the exit cost panels have a narrower focus on exit exams, sometimes look at current costs, and have an additional task of considering the cost of hypothetical changes in the exam system.

Operational Methods

Advisors

Before visiting a state, we establish a relationship with one or more advisors, including someone in the department of education. Additional advisors may include members of legislative committees, university experts, business or foundation proponents of high stakes testing, and officials of the statewide school board, and teaching and administrative organizations. Advisors help provide information about potential participants; help solicit cooperation; provide news about developments; suggest places to meet; offer suggestions about proposed procedures; and provide background information about events and personnel. They may be asked to serve on an advisory committee.

Panels

PURPOSE

Panels of informed personnel provide the primary information about the use of resources in implementing high stakes tests. District panels provide information about school-level and district-level resources. Later, a statewide panel reviews their work, converts resources into costs, and provides information about state-level costs. Additional information from the panels about resources required for an altered, hypothetical examination is used to compare costs in states with different examination configurations.

TASK

The task of estimating the resources required for an adequate chance for students in a hypothetical school district to pass the high school exit examination is split into two stages: estimating the resources required for the current exam and standards and estimating the resources required to increase performance to the state-certified adequate level. These two tasks are assigned to two different sets of district panels. Additionally, each district panel is asked to specify the resources required under a different, hypothetical examination arrangement.

COMPOSITION

District superintendents have more of the desired information than any other group, and they form the core of the panels. In larger districts, assistant superintendents, business managers, and curriculum supervisors may be more knowledgeable. In addition, panels include representatives of groups such as teachers, principals of schools at each level, assessment specialists, special education providers, and alternative education experts.

SELECTION

Advisors of the project are invited to suggest names and contact information for potential panelists. Calling upon a diverse group of advisors helps provide a diversity of suggestions. Potential participants are sent a letter of invitation, including information about their recommendation by an advisor. A follow-up phone call solicits their commitment. An acceptance rate of 50–60% is expected. For teachers, some released time—perhaps for professional development—may be requested from principals and district superintendents.

LOCATION

Panels of 5–7 members are convened in the capital and at convenient locales around the state. The panels represent different regions of the state and are located to ease access, notably for small districts in rural areas. Panels with a concentration of personnel knowledgeable about at-risk students and alternative schools normally are held in urban areas, and those with a bilingual education focus meet in areas close to the Hispanic population. Facilities may be provided by the state department of education, a school district, or supporting associations, or may be a local restaurant or meeting place.

DURATION

Panels last for two to three hours. The reporter takes notes and tape records the proceedings for verification. While it is important for all interviewers to attend and jointly debrief an initial panel of each type of task, a single moderator and a recorder normally conduct a panel.

PROCEDURE

After refreshments and introductions, the moderator presents the outline of the project and the task of the panel. Panelists are presented information about a hypothetical, typical school district and asked to estimate the resources required to provide students with 1) the current chance of passing the high school graduation examination or 2) an increased chance of passing at the minimum level officially described as adequate—where no further improvement is required of districts by the state. The panel first decides upon programs, and then estimates the required resources. The moderator may offer an initial list of possible programs and a draft spreadsheet for specifying the resources used for programs. Interviewers record consensus estimates, take notes and use a tape recorder to keep track of the conversation. Panelists respond to a summary of the findings of the session, fill out a questionnaire about the procedures used, and then are thanked and the panel ends.

FOCUS

We ask the panel to identify required supplies and staff time, professional and support, for each category of program. Later, they distinguish between time shifted from other tasks and new positions and between nonrecurring and recurring costs. When possible, panelists identify resource use with programs. Toward the end of the panel, attendees are asked about how resource use would have differed had the programs been changed in one or

more fundamental ways, such as being end-of-course exams or lacking accountability or having a much lower pass rate. A final questionnaire asks about the process in order to make improvements.

REVIEW PANEL

In addition to the informed sources panels, a statewide panel meets to review the findings. While the district panels focus on resources, the review panel helps to translate them into costs. This panel reviews the use of resources estimates and decides how to resolve differences in estimates among panels. Additionally, the review panel estimates the state-level resources required. The statewide review panel includes specialists in finance at the district level, as well as specialists in exit exam budgeting, legislative staff for education and finance, those knowledgeable about state programs and policy, and experts in remediation and assessment from the department of education. Representatives of additional groups, such as parents, may be invited to participate as well.

Interviews

To provide supplementary information from a variety of points of experience, interviews are arranged with those knowledgeable about legal costs, curricular changes, legislative concerns, remediation programs, policymaking, assessment, teachers' concerns, parental concerns, in-state academic advice, and public information. These interviews are commonly one-on-one but may be held in small groups when convenient. The agenda and format vary, but the meetings usually take about an hour and focus narrowly on the contribution of special knowledge. Some of these meetings may be over the phone or through e-mail if necessary.

Documents

The main print sources of information are official documents of the state government and department of education describing budgets, state programs and grants, organization of the state education system and department, characteristics of students and teachers, prices and costs, the exit exam and its administration, standards for and rates of passage overall and for disaggregated results, state standards, and exam remediation. Additionally, state and district websites provide information about personnel, intended changes, district exam results, contact information, job responsibilities, publications, recommendations to districts, federal funding, experimental programs, and legislative priorities. Sources of background information and analyses for state exit examination costs include the *Journal of Education Finance* and the National Center for Education Statistics, as well as the books and articles cited in this report.

In choosing the core method of professional judgment panels, APA also chooses the related methods it has evolved, such as a review panel and a nominating group. A method devised for another purpose can be rearranged to study the cost of high school exit examinations with adjustments in the tasks of the panels. The central feature of this method is that for an average school district, panels are asked to specify the resources needed to produce current levels of performance or state-sanctioned levels of performance on the high school exit tests.

CHAPTER 3

Procedures Used in Indiana

APA's aim was to find a state that would tell us the most about how the technique would perform when applied to a fuller set of states. Indiana best met the following criteria for selection of a state in which to test the methodology. Indiana has completed the early and middle stages of the exit exam process and continues to innovate, moving beyond the testing approach to include remediation and to begin to consider prevention programs. Indiana's Graduation Qualifying Examination typifies exit exams in its major characteristics: student pass rates and extent of ethnic differences; standards-based test format that includes short answer; options in meeting the graduation requirement that include something for general education students, and offer accommodations only for disabilities and limited English; and local requirements that include school and district accountability, with student remediation efforts required of schools and funded by the state. Indiana offers adequate data; the award-winning state website; a comparatively stable exam since 1997; results disaggregated by race/ethnicity; and experience with seniors failing. Further, Indiana offers feasible arrangements for interviewing within the budget: a moderate-sized population with one large, central, capital city; reliable transportation within the state in January; and interviewing contacts from recent APA work and a CEP case study.⁷ Overall, Indiana offered the best chance of completing a study under the available budget that could both estimate costs in the test state and help prepare for studies in additional states.

The procedures used in Indiana closely follow the methods proposed above. This section will provide two types of additional information: specifics of the procedures used in Indiana that illustrate the method, and an evaluation of successes and shortcomings. The narrative offers particulars about the practical operation of the method, while the evaluation forms the basis for recommendations in a concluding section. Our intent is to be open and frank about our methods. The subsequent section discusses the substance of what APA found out about the costs of the GQE in Indiana.

Narrative of Procedures

We spent four days in Indiana. On Monday, APA conducted interviews at the Department of Education (DOE) offices and at the headquarters of the Indiana State Teachers Association (ISTA). Three interviews we had arranged fell through, each for a different reason—a traffic accident, a change in staff—which led to additional interviews later on.

⁷ Indiana provides unusually full documentary information on district (corporation) finance and assessment performance. At the department's award-winning website, see <http://ideanet.doe.state.in.us/htmls/performance.html>. A companion study to this one might use cross-district, over time statistical models to analyze the documentary evidence, such as the yearly corporation expenditures for remediation and the percentage passing the exit exam. Even this relatively full documentary evidence is, however, a patchwork of available measures, rather than a systematic collection of data germane to the cost of the exit exam.

On Tuesday, two district panels convened in Indianapolis, one focusing on current costs in the morning and one focusing on costs of meeting standards in the afternoon. Debriefings from these panels led to clearer task descriptions for panelists, the addition of information about the number of teachers in the hypothetical district, and the dropping of the hypothetical questions in future panels.

On Wednesday, the team of Doug Rose of APA and Keith Gayler of CEP conducted a panel addressing meeting standards in Bloomington, and then held a panel focused on current costs in Lafayette. The team of John Myers of JL Myers Group and Katherine Christensen of APA conducted a panel on current costs in South Bend. In the evening, results from all the district panels were readied for presentation to the review panel.

On Thursday, the head of assessment from DOE was interviewed, and then the state-wide panel met. It adjusted estimated prices; suggested pooling the results of the three panels on current costs; selected the output of one panel on costs of meeting standards as a much better estimate than the other; made non-numeric suggestions about state-level programs; and did not have time to consider the hypothetical questions. In the afternoon, the delayed interview with the communications specialist was held, as was an interview with the chair of the House Education committee.

The names of our advisors, panelists, hosts, and interviewees are included in the appendices. We are grateful for all the help we received.

CEP's initial contacts were most useful, notably Lowell Rose of the University of Indiana, George Kersey of Phi Delta Kappa, and Marc Steczyk of the DOE. Though APA drew early on upon its contacts, such as Terry Spradlin in DOE and Dan Clark of ISTA, we did not start from our list of participants from an earlier adequacy study in Indiana. Many of these prior contacts turned up in the course of the study. In general, making an on-site visit several months before the panels to make arrangements and establish relationships would have avoided difficulties.

Lowell Rose kindly put together an advisory panel of six well-respected individuals who were collectively well informed about the categories of people we were seeking, and the panel came up with a list of 80 names of potential panel participants. We contacted 81, of whom 52 agreed to participate, an acceptance rate of 71%. The acceptance rate was highest for the panels in Indianapolis. In the end, 37 individuals turned out for the five district panels, a turnout rate of 64% among the 52 who agreed to participate, and an overall yield of 46% on the original names. The turnout was held down by conflicting meetings for assessment held during the same week, by unusually cold and snowy weather, and by statewide budgetary emergencies that kept superintendents home, focused on their budgets and boards.

The number of panels, their central tasks, and their locations underwent changes, primarily in response to the availability of panelists, either near the locations or with the appropriate expertise. We planned six panels and ended up with five. In addition to Indianapolis and Bloomington, we originally planned panels in Terre Haute and Fort Wayne but actually conducted panels in Lafayette and South Bend because of the distribution of acceptances.

The original plans had called for some concentration on special topics such as alternative education and special education, at-risk, and limited English students. However, as the plans evolved the main distinction became one of focusing on current costs ("what is") and focusing on the cost of achieving stated goals ("what if"). The remaining differences in substantive topics centered on the hypothetical questions unique to each group, which were matched to the corporations and individual experience of the participants. When the hypothetical ques-

tions were ineffective in the first two panels and were dropped from the remaining meetings, the panels differed simply in location and in their what is or what if topic.

In combination, the composition of the district panels was as follows:

Table 1-1: Indiana Panel Member Positions

Superintendent	10	27%
Assistant Superintendent	4	11%
Business Manager	6	16%
Principal	8	22%
Assessment, Curriculum	5	13%
Special Ed, Counselor, Teacher	4	11%
Total	37	100%

The participants were told to expect to spend two to three hours at the panel. The two initial panels stayed for the maximum time, including a break, without completing the hypothetical question. Panelists were not capable of tackling the new task posed by the question. The remaining panels met for around two and one-half hours. Time constraints limited two of the panels, one of which completed its work and one of which did not have time to discuss resources for the more minor programs it suggested.

The instruction sheets for district and review panels are included in Appendices I, II, and III. The instructions indicated the purpose of the group, described the hypothetical district, set forth the standards for performance on the GQE, and emphasized the task of specifying resources for programs to meet the standards. In a reminder e-mail the Friday preceding the meeting, participants were reminded of the time and place, provided transportation instructions, and given a suggestion to think about programs for producing exit exam results. At the meeting, panelists were provided with a starter list of programs, some of which were never discussed by any panel. Once the programs had been agreed upon, APA handed out a spreadsheet whose empty cells represented a resource (row) to be used for a program (column) for the exit exam. Again, some resources and most combinations were never used. While panels differed, they tended to focus on a few resources and a few programs. All panels discussed the resources needed for the major programs they had specified. If the panel was not short on time, the moderator reviewed with panelists the choices made and prompted for additional suggestions.

Part of the process of refining our methods was to get feedback from panelists. At the conclusion of the panel, participants were asked to complete and return a questionnaire covering the following questions:

- Do you feel that the instructions and materials presented to you were clear?
- Did you have an understanding of the tasks that you were being asked to complete?
- During this meeting, have you been able to openly express your ideas and concerns?
- Do you feel these have been used in the final conclusions of the group?
- Do you feel that we have addressed all of the resource needs involved in implementing programs that would yield the specified GQE outcomes?

All but the four participants who left early returned the questionnaire.

All indicated that they felt able to express their ideas, and no one said that his or her ideas were not used in the conclusions, so a positive consensus existed on the participation aspects of the process. There were some concerns expressed about the instructions and the understanding of the tasks, especially at the initial panel. In general, participants gained confidence in their understanding as time progressed in each panel. Finally, a substantial minority expressed skepticism that “all” the resource needs had been addressed. While they were proud of their work, there was often doubt about whether the task was possible at all, much less in a morning or an afternoon.

Two of the review panel participants did not make the meeting due to a snowstorm, and one participant had to leave early. The need for expertise about state-level programs was filled with interviews and follow-up exchanges with DOE personnel. The group met from 8:30 a.m. until noon.

The interview about legal costs was conducted via e-mail before the visit, and an interview about special education was conducted by telephone after the visit.

Evaluation of the Successes and Shortcomings of the Procedures in the Test State

This is a new method and that newness implies some mistakes. Indeed, part of the purpose of the test case in Indiana was to discover the weaknesses so that they could be corrected. Because the basic professional judgment approach has already been used by APA, however, the approach itself was not expected to produce problems and errors if it fit the task.

Overall, the methodology was successful in yielding informed estimates of the cost of the GQE. Most of the specific methods also succeeded in accomplishing their basic purpose. Therefore, the list of positives will be kept short, limited to aspects that worked notably well or without a glitch. The list of negatives will also be kept short by omitting mistakes of the sort expected even in a mature methodology.

Success

The advisory group was well informed and produced early a quality list of nominees who well represented the populations desired. Unlike the adequacy studies, this project had no in-state sponsor or policy payoff, so the motive for participation was weak. The reputation and established relationships of CEP and APA helped secure participation of well-respected figures, which in turn persuaded panelists to accept our invitation.

APA had forecast a 60% acceptance rate and a 75% turnout, for an overall yield of 45%. The acceptance rate was even higher than expected. Because we made no early visits, a high acceptance rate based solely on letters, phone calls, and e-mails is a success. The earlier work by APA and CEP in Indiana helped. Before the weather turned worse, turnout was high, reflecting in part the success of our reminder, map, parking information, and introduction to the task. Throughout, interest in the topic of exit exams was pronounced.

The separation into two tasks for two types of panels, what is and what if, allowed the panels to have a clearer focus and to complete work on time. The use of a common, average

hypothetical district eases the comparison of the two types of costs. The two and one-half to three hour duration was enough time for panels to complete their basic task.

The conduct of the panels—instruction handout, separate program and resource discussion stages, program list handout, keeping to task, chance to contribute—was successful in the senses of yielding expected results and generating few complaints and some compliments.

The review panel completed its most important tasks. Its substantive decisions, from diverse perspectives, about the district panel output agreed on the key points with the assessments the interviewing team had formed while conducting the panels.

The use of multiple panels improved the output. The later panels were more productive in less time than the early panels. Comparing panel outputs in terms of programs, resources, and total costs provided a context that the review panel felt it needed to arrive at an overall judgment. The three estimates of the current recurring local costs were all within 13% of the average, close enough for the review panel to make confident judgments. In statistical sampling theory, if three estimates have this amount of variability, then their average would be expected to err by about 5%.

State-level costs were adequately estimated from interviews, documents, and follow-up questions. DOE officials were accommodating in providing information.

Improvements Needed

No African-Americans attended the panel. As 19% of public school students and 5% of teachers in Indiana are minority, the omission is important. Because the African-American share of superintendents and specialists is probably well below the proportion of teachers, some race-conscious measures needed to be taken to ensure participation. An interview with the chair of the legislative Indiana Black Caucus covered aspects of the achievement gap.⁸

The early use of the list of prior participants in APA's Indiana adequacy study would have eased the creation of the panels, including the review panel. Specifically, it would have provided an alternative source of names when acceptances were low for a particular geographic site. Such contacts are not available in most states.

The timing of the visit caused problems. From mid-November through mid-January, it is difficult to make arrangements or conduct panels. Our visit conflicted with state-sponsored assessment meetings in Indianapolis, which reduced attendance at the Wednesday panels around the state. An early state visit might have prevented this problem. Our January visit also overlapped with a change in the legislature, which caused some problems with the first day's interviewing.

The hypothetical questions failed. There was not enough time for them, they were too unrelated to the core task, and they required more creativity and energy than panels had available toward the end of the session. The review panel then had nothing to review and no time in any case. Without the hypothetical questions, it becomes more difficult to compare Indiana to other states and to cost out provisions of the exams.

8 Legal issues surrounding this gap are discussed in *The Use of Tests as Part of High-Stakes Decision-Making for Students: A Resource Guide for Educators and Policy-Makers*, U.S. Department of Education, Office for Civil Rights, December 2000. www.ed.gov/offices/OCR.

There were too many tasks, at least with the hypothetical questions included. One consequence was time pressure; another was an over-reliance on the moderator to run the group, with a corresponding lesser commitment by panels to owning the process. The lowest estimate of current costs came from a panel that skipped low priority programs because of time pressure. A guaranteed availability for three hours would have helped, providing for a break and enough attention to the question of shifted resources.

While the topic was interesting, the task generated little enthusiasm. Exit exams may be important, but educators and administrators do not love them. Without the hint of a policy change in the offing, panelists took at best a work-like orientation to estimating the use of resources for the GQE as it is. More surprisingly, panels were not much more interested in figuring out the programs and resources needed to achieve state-set goals. In the case of the what if panel among large and urban schools representatives, this reflected some dislike and skepticism about exit exams themselves, as well as the absence of the sort of pragmatic leadership usually provided by superintendents. The reluctance to attempt a grant proposal-like formulation may also reflect the current fiscal situation, where resource cuts rather than increases dominate thinking and concern.

The procedures used for the panels produced some additional problems:

- None of the Wednesday panels had three hours available if needed, and 2.5 hours should have been the minimum length.
- The early panels lacked clarity on the basic task; better analogies were created the next day; in general, moderators improved with practice.
- The suggested list of programs and resources was too long, partially reflecting our need to discover which topics panelists would be most likely to find appropriate.
- The spreadsheet was too imposingly large, making the task seem difficult.
- As expected, the use of a single hypothetical district discouraged exploration of programs appropriate to atypical districts.

Programs targeted at populations—special education, at-risk, and limited English students—were scarce, given that exit exams programs have to succeed with these groups in order to produce high levels of success. In part, this represents an emphasis on mainstreaming these students. While there was some discussion, panels usually avoided programs targeted to these populations. Instead, targeting tended to be toward those who had failed state assessments, with programs being appropriate to any student who fails.

Because panels only addressed the topic of a typical school corporation, they made no estimates of the differences among districts in the costs of exit exams. However, initial differences among districts are apparent in the varying initial failure rates, presumably followed by inequalities in the funds required for remediation. While maturation takes away a share of the problems in language arts, failures in mathematics routinely require intensive remediation. Districts with high rates of math failure will need to spend more on remediation. Districts also vary in the professional development of their staff, and the professional development requirements fall most heavily on districts and schools with the least preparation. Finally, districts having the highest proportions of students who begin schooling with the least preparation should have the highest need for preventive funds. If these problems are severe for a district, then the usual method of shifting resources will not be adequate,

as the need outruns the per pupil allocation. When the problems are present together, need could easily outrun available funds. Conversely, school corporations with low rates of failure, few math problems, well prepared staff, and well prepared entering students will, because of the low need for exit examination funding, have a relative surplus of funds to spend on programs of all sorts. Differences among districts may increase as attempts are made to raise scores—by far the highest estimates of the resources needed to reach the commendable level of performance came from the what if panel composed of representatives of large and urban school districts with more at-risk students than average.

The expected costs of providing test accommodations to special education students could not be measured from available data or be estimated by the panels. The resources are time and staff, but these are invisibly shifted from other uses and leave no documentary trace.

Federal programs came out only in the course of discussing model or effective or adequately funded programs. They were not presumed germane to the typical district.

The panels did not feature one-time, non-recurring costs. Questions explicitly addressing these resources came late in the session and were cut if time was short. Even when asked, however, panelists assumed recurring resources except for the obvious cases such as technology purchase or the use of outside consultants. For large changes, panelists tended to assume a slow process that required a constant expenditure and would eventually yield persistent benefits. Several noted that preventive costs should eventually diminish the costs of remediation.

In summary, Indiana offered a typical, feasible state. The actual methods used closely tracked the planned methodology. The narrative reveals the usual surprises found in any specific application. However, the methods basically worked as expected and produced the desired types of information. In addition, the test in Indiana suggested improvements to the methods, including an early visit, a better time of year for conducting panels, an eye to adequate representation of African-Americans on the panels, and the replacement of the hypothetical questions. Panelists judged the process a success, though some were skeptical about any group's ability to adequately specify all the resources needed for performance on the GQE.

The methodological discussion is complete, having moved from the broad choices, to the method in principle, to the evaluation of the applied techniques. Next comes substance, the estimates of the costs of the high school exit examination in Indiana.

CHAPTER 4

Indiana: The Costs of the Graduation Qualifying Examination

This section presents the substantive results of our investigation into the costs of the Graduation Qualifying Examination in Indiana. After an introduction to Indiana’s schooling and GQE, we present the premises of our study, followed by the actual cost estimates. The estimates for the costs, local then state, of the current level of performance come first, and APA’s estimates of the costs for improved performance on the GQE come at the end. The estimates should be viewed as being within a range of approximately +/- 10%.

In presenting current costs, first come the estimates for the recurring costs of local programs, followed by estimates of one-time costs and of funding by shifting resources from other programs. For state involvement, state funding for local programs—already covered in the local programs—is distinguished from state-level programs administered by DOE. The costs are then summed into total yearly costs, first for current performance, then for improved performance.

Background Information on Indiana’s GQE

As indicated in **Table 1–2**, Indiana public schools reflect the characteristics of public education nationwide. Indiana has fewer non-native English speakers than average, spends more than average, and divides responsibility among more districts than average, but is not so different as to be extreme.

Most of the yearly cost of education occurs for personnel. **Table 1–3** reports the estimated average salary and benefits for the personnel most germane to the costs of the GQE. The base calculations are discussed in the APA report on its adequacy study for Indiana, available at <http://www.aandm.org/INDIANAFINALREPORT.pdf>. The review panel adjusted, confirmed and added to the estimates. Finally, the latest data on average teacher salary available on the DOE website were inserted, with all categories adjusted proportionally.

Indiana conducts statewide assessments in mathematics and language arts for grades 3, 6, 8, and 10 as part of the Indiana Statewide Testing for Educational Progress (ISTEP+). These assessments are criterion-referenced, relative to standards and cut-points established by the state. The exams are standards-based to reflect the curriculum guides for Indiana education.

Indiana students must meet the Graduation Qualifying Examination requirement, in addition to earning the course credits needed, to receive a diploma. Students must demonstrate mastery of 9th grade skills in English and math, and may do so in one of three ways:

- Pass the GQE in mathematics and English/language arts.
- Complete all components of the Core 40 curriculum, a more rigorous curriculum than the general education track, with a “C” or better in each course.
- Appeal their test results with a 95% high school attendance rate, attain a “C” average in the 22 credit hours required of all Indiana high school graduates, take the test at least once a year, participate in remediation opportunities provided by the school, gain the recommendation of a teacher in mathematics or English, backed up by documentation of mastery of the subject area, and have the principal concur with the teacher’s recommendation. A similar, though not identical, appeals process is available for students with disabilities.

Table 1–2: Indiana Public School Data*

	2001-2
STUDENTS	995,407
Minority	19%
Free lunch	24%
LEP	2%
Graduation rate	91%
EXPENSE/PUPIL	\$ 8,128
PUPIL/TEACHER RATIO	16.7
CORPORATIONS (DISTRICTS)	293

*Data are adapted from the tables listed at <http://mustang.doe.state.in.us/TRENDS/trends0.cfm>

Table 1–3: 2003 Indiana Salaries

	AVERAGE	BENEFIT RATIO	TOTAL
SCHOOL PERSONNEL			
Teachers	\$44,884	30%	\$58,349
Technology/Assessment Specialists	\$53,071	30%	\$68,992
Guidance Counselors	\$48,603	30%	\$63,184
Clerical/Data Entry	\$23,474	30%	\$30,516
Instructional Aides	\$13,169	0%	\$13,169
Substitutes	\$44,884	9.5%	\$49,148
DISTRICT PERSONNEL			
Assistant Superintendent	\$78,528	30%	\$102,086
Supervisor/Coordinator/Director	\$71,826	30%	\$93,374
Technology/Assessment Supervisor	\$71,826	30%	\$93,374
Curriculum Coordinator	\$71,826	30%	\$93,374

In the fall of 2002, 68% of the Class of 2005 passed the mathematics portion of the GQE, and an identical percentage passed the English part, with 60% passing both exams. Students who fail have an opportunity to retake the exam every semester of their junior and senior year, for a total of four times. For those not passing, remedial instruction time is offered, performance on specific skills is reported, and the student’s work is returned for diagnosis and remediation. The first class required to meet the GQE requirement, the Class of 2000, went from an initial two-subject passing rate of 54% to a final passing rate of 86%.

As reported in **Table 1–4** as well as on pages 80 and 107 in CEP’s *State High School Exit Exams: A Baseline Report*, the most recent disaggregated results, for the 2001 exam, show wide achievement gaps. Most of the low-income students and students with disabilities in Indiana are white, while most of the English language learners are Hispanic.

Based on their students’ average passing rates on both the English and mathematics exams, schools are classified by the state by their level of performance and by the amount of their recent yearly improvement. The category of “commendable,” where no further improvement is demanded, is achieved by an average initial pass rate of 80% on the exams.

Changes in the mathematics examination may make it difficult to improve these results without substantial change in supporting programs. By 2003–4, 30% of the GQE math test will be based on Algebra I content, though only 70% of Indiana’s students currently take that course before they take the test.⁹

In general, Indiana has a well-developed system of assessment, capped by the GQE. For instance, DOE makes available for downloading off its website PDF guides, at each grade

Table 1–4: GQE Pass Rates

	ENGLISH/LANGUAGE ARTS	MATHEMATICS
All students	68%	65%
White	73%	70%
Hispanic	49%	46%
Black	38%	31%
Free or reduced lunch	45%	42%
Disabilities	19%	21%
Limited English	28%	33%

⁹ The Education Roundtable by law reviews and recommends the content and format of Indiana’s state-wide assessments. The governor and the superintendent of public instruction appoint it jointly. Its “Resolution Determining the Mathematics Content for the New Graduation Qualifying Exam (GQE),” October 8, 2002, reads in part: “Whereas to recognize outstanding student performance on the Graduation Qualifying Exam, the assessment must include questions designed to differentiate students excelling at the highest level; Now, therefore, the Education Roundtable recommends the content of the mathematics section of the new Graduation Qualifying Exam to be as follows: 70% of the test will cover the cumulative academic standards for mathematics through grade 8 and 30% of the test will cover all of the Algebra I academic standards.”

level tested, for parents, for educators, for language arts remediation, and for mathematics remediation.¹⁰ The tests are tightly linked to standard-based instruction as it is practiced by experienced teachers.

The Current Costs of the GQE

Premises of Our Study of the Current Costs of the GQE

To discover the school and corporation costs of the GQE, APA created a hypothetical district that resembled one the average student attends. The hypothetical district has 5,350 students in 11 schools, including two high schools (9-12) with 800 students each and an alternative and vocational school with 150 students. Of the hypothetical total, 16% of students are in special education, 22% of students are eligible for a free lunch, and 3% of students are in limited English proficiency programs.

In this hypothetical district, the GQE tests, test-takers, and results closely followed current statewide figures:

■ Initial (10th grade) total pass rate on both exams	60%
■ Pass rate for English/Language Arts & Mathematics, singly	68%
■ Students passing both tests within 2.5 years	86%
■ Special education students taking the test (84% of them with some type of accommodation)	11%
■ ESL/LEP students taking the test (66% of them with some type of accommodation)	3%
■ Initial pass rates for special education students	19%
■ Initial pass rates for free lunch students	45%
■ Initial pass rates for black and Hispanic students	40%
■ Initial pass rates for ESL/LEP students	28%

Panels of superintendents, principals, specialists, and teachers were asked to specify the programs and resources needed for this average district to produce average results. One analogy offered for their task was to advise the superintendent of this hypothetical district in a state similar to Indiana but that had just adopted an exit exam what programs and resources he or she should use to achieve the results. In sum, panelists were provided a state microcosm at the district level, then asked what programs, with which resources, would be required to produce the GQE results given the district characteristics. On a per student basis, these estimates apply statewide as the local portion of costs.

10 For preparation for the GQE, the most useful guide is the *ISTEP + GQE Item Sampler*, DOE, 2001, that covers the multiple choice portions of the test.

Local Yearly Costs of the Graduation Qualifying Examination

All three panels examining the resources currently needed for local GQE programs arrived at similar overall estimates of the costs, so the review panel recommended pooling them to arrive at an average estimate. The type of resources and the location of expenditure for the costs per pupil of the GQE are summarized in **Table 1–5**. The yearly costs are substantial, and primarily arise at the school level to pay for personnel. Over one-half (54%) of the personnel costs is for teachers, while another one-third (34%) goes for specialists, notably in assessment and curriculum. The remaining personnel costs are for supervisors such as assistant superintendents, nonprofessional personnel such as aides and clerical assistants, and costs applying to all personnel in the district. Among the nonpersonnel costs, most (62%) goes for instructional materials and supplies, one-ninth (11%) purchases technology such as software, and one-quarter (26%) is spent on transportation.

Table 1–6 indicates the allocation among types of programs of the same per pupil local cost. The four major types of programs share relatively equally in the cost. Remediation includes classes for students who have failed an ISTEP+ exam, either in the 10th grade or earlier. Preventive programs often make use of specialists, technology, or special instructional materials. Professional development includes professional development days for teachers, with the use of substitutes for school-year days, and the employment of assessment specialists or consultants. Testing includes the actual administration of the exams, as well as data entry, analysis, and administering exam accommodations for students in special education.

Some of these local costs are paid for by state funds. The state provides an estimated \$6 per student for summer schooling that is due to the GQE, and another \$9 per student in remediation funds for the GQE. The federal government and other non-local sources provide small but unmeasured contributions to the GQE costs.

Table 1–5: Per Pupil Local GQE Costs

	SCHOOL	CORPORATION	TOTAL
Personnel	\$ 318	\$ 60	\$ 378
Other	\$ 34	\$ 30	\$ 64
TOTAL	\$ 352	\$ 90	\$ 442

Table 1–6: Local GQE Programs

	SHARE OF COSTS
Remediation	29%
Prevention	28%
Professional Development	25%
Testing	18%
Other	1%
TOTAL	100%

One-Time and Shifted Resources

Where, then, do the funds come from to pay for the cost of the GQE? The funds for schools and school corporations in Indiana largely come in bloc grants from the state on a per pupil basis, with adjustments for students with special needs. Because there is local flexibility in the use of these funds, superintendents and principals can respond to demands such as the GQE by reallocating resources. Our panelists described how funds available for multiple purposes are shifted from other educational programs when GQE results became a priority. For example, advanced math class sizes are increased in order to free up teaching funds for more basic instruction. In general, as having marginal students pass the math and language test becomes more important, funds are shifted from instruction for other students and from instruction in other subjects. For programs such as special education and Title I that are already targeted at academically marginal students, resources are shifted toward preparation for and remediation after the GQE. Though an earlier study of remediation found math remediation classes somewhat effective, panelists emphasized earlier instruction in mathematics and, even more, early emphasis on reading to prevent problems with mathematics.¹¹

These shifted resources are relatively invisible to outsiders, as they are not met by budget increases or at least not by budget increases described as due to the GQE. Indeed, even our expert panelists had difficulty attempting any precise estimate of the magnitude of these reallocations, and so we offer none. No single program or category of expense dominated the discussion of where the funds come from or where they go.

The use of shifted resources appears to be more notable for the GQE costs that are indirectly connected to the test. While direct state funding is available for testing itself and for remediation of failures, it has not necessarily been available for mandatory professional development or for early reading programs that head off failures. When superintendents believe these programs are critical to success on the GQE and that the success on GQE will be how students, parents, and state officials evaluate the district, then resources may be shifted from programs benefiting students unlikely to fail or subjects that are not tested.¹²

In addition to shifted resources, panelists indicated relatively modest one-time, non-recurring costs for the GQE. These costs averaged \$12 per pupil for software for tutoring and record keeping. Historically, the GQE did not come in with a large infusion of initial costs and funding. Nor has the GQE been publicly linked with visible one-time costs, such as new buildings, hardware, or buses. Nonetheless, the panel's estimate is more likely to be too small than to be too large. For instance, the use of funds to bring in specialists on a one-time basis to help schools prepare in some way for the GQE was not mentioned by panels. However, it may be characteristic in moderate-sized and larger districts leading to higher one-time costs than indicated.

Though hard data are lacking, the shifting of resources appears important for the special education costs of the GQE. Due to broad reform efforts, special education students increasingly have been included in the general education curriculum. With the movement toward

11 "An Investigation of Indiana High School Remediation Programs," CTB/McGraw-Hill, Monterey, CA, March 2000.

12 The shifting of resources is a subject of controversy. For a favorable view, see Allan Odden and Carolyn Busch, *Financing Schools for High Performance: Strategies for Improving the Use of Educational Resources* (Jossey-Bass Inc., 1998).

standards, special education students have been exposed to standards-based instruction. However, without the GQE, students might be listed as receiving more grade-appropriate material than they actually were exposed to. With the elimination of the “diagnostic option” in Indiana, the pass rates of special education students counted in a district scorecard, and special education students were eligible for remediation of a test failure. Special education teachers now had to learn how to teach to standards, and general education teachers had to learn how to teach special education students. The funds for this training may have come from increases in state and federal streams of funding for special education. The federal funds now account for one-quarter of Indiana special education funding and increased by 25% two years ago and 17% last year, and out of such a sizable fund can come the ongoing costs of professional development associated with the GQE and special education students.

What the Local Costs Imply About the Sources of GQE Problems

The pattern of costs ill fits what might be called the “motivational understanding” of the high school exit examination. Exit tests in the motivational understanding are an inexpensive route to improved education. When motivation is perceived to be key, then the student's anticipation of a high stakes test leads to increased application and learning in earlier grades, which yields improved education at no cost to the public except the cost of testing. In addition to this sort of prevention by motivation, students who fail the exam may be motivated by failure to apply themselves in class or in remediation classes in order to pass. While this remediation is more expensive to the public than prevention, it is still relatively inexpensive to provide essentially the same material on an additional occasion.

The amount and pattern of costs outlined by panelists go beyond the motivational understanding. While participants accept the motivational premises for some students, the bulk of the cost of the GQE goes to improve performance of students who have problems that start in the early grades and that are not caused by motivational problems. These students are often at-risk or have learning disabilities. Motivation problems are as likely to be a consequence of their educational problems as they are to be the cause of them. The preventive and pre high school programs specified by the panelists are targeted toward students with early learning difficulties. While in a typical district it is possible to shift resources from students without difficulties to those with them in response to the GQE, in a district with a high concentration of students with difficulties, there is no place from which to shift the resources. These costs of meeting the GQE standards are substantial.

A third type of difficulty illustrated by the indicated costs is a training problem. Teachers trained in earlier eras have to learn how to teach to standards and how to translate test reports into a teaching strategy for a student who fails and who is likely to fail. The professional development emphasis reflects this problem. As the exams and standards keep changing, this cost appears recurring for the immediate future. The costs of professional development are substantial.

State-Level Costs

Because of reforms, including the No Child Left Behind Act, a high school test is inevitable. In looking at the costs of the Graduation Qualifying Examination, we attempt to subtract out the costs of an ordinary, low stakes exam. That is, only the cost associated with the “graduation qualifying” aspect of the GQE is estimated. This is perhaps the most difficult aspect of the estimate, but it aims at answering the question of what the costs would be with a 10th grade exam but without an exit exam.

“State-level” refers to programs that operate statewide, rather than being carried out at the school or corporation level. The state government is the author of the state-level programs we know about, and the Department of Education conducts them. In addition to these state-level programs, the state government makes funds available to school corporations and schools under varied programs, including basic tuition grants and programs targeted to remediation of GQE failures.

In general, we count as state-level costs of the GQE the resources used by the state to conduct those programs that:

- Prepare students for the GQE, such as the pamphlet that informs 9th and 10th grade students and their parents about the test;
- Prepare and conduct testing and scoring, such as the cut-point setting conference of January 2003; and
- Remedy failures, such as providing the remediation guide. Funding from the federal government dedicated to these state programs is included. State payments to school corporations do not count, even if targeted for the GQE, if they are counted under local programs and resource use. We do not count testing programs for earlier grades, at least in the absence of evidence that these costs are incurred in order to produce success on the GQE. Indirect, partial resource effects, such as changed course offerings in education departments of Indiana’s universities, are not considered direct state-level costs.

We do not attempt to estimate separately the state share of grants from the federal government for GQE-related programs. Though federal programs under Title I, for instance, target populations at a higher risk for GQE failure and are often the source of model programs for improving GQE performance, they are not explicitly linked to the GQE, and we have no information about their role in state-level programs. The \$84 million federal grant received in January 2003 for “Reading First” in 65 corporations is not explicitly tied to any exit exam criteria, for instance.

Neither have we estimated the cost of GQE public relations and legal costs. Although the GQE figures prominently in deliberations and pronouncements of top state officials, there is no “public relations” budget line item for agencies and no GQE subheading. For legal costs, while there has been a court case seeking a remedy as a consequence of the GQE and more might be expected, no record of attorneys’ hours was kept, no outside counsel was hired, and no payment made, since the state successfully defended the suit. The Attorney General’s office would normally handle cases as part of its legal role for the state. Consequently, the Department of Education’s General Counsel, Kevin McDowell, believes that there is no method for calculating the actual legal costs incurred to date in defense of the GQE, nor is there any way to project potential legal costs.

STATE-FUNDED LOCAL PROGRAMS

Before considering the state-level costs, we review two prominent state-funded programs that we count as corporation-level programs: summer school and testing remediation.¹³ For summer school, there are 51,643 9th–12th grade students taking “category 1” courses that

13 For a summary of state-funded programs, see the Department of Education’s *Digest Of Public School Finance In Indiana, 2001-2003 Biennium*.

are required for graduation, at a cost to the state of \$9.4 million. If one-half of these courses were taken because of the GQE, in preparation for the exam or its retake, then \$4.7 million would be the cost. If 10% of the 17,647 students taking other sorts of summer school courses, such as noncredit remediation, were due to the exit exam, then another \$0.3 million would be state cost, for a total of \$5 million. As the localities pay 15% of the total, local costs would be \$0.9 million, for a total summer school cost of \$5.9 million for the GQE.

For state-sponsored remediation, \$20 million is available to corporations to pay for remedying the results of tests for grades 3, 6, 8, and 10. The 10th grade share is more than the proportional one-quarter. The amount given depends on the distance between the student's actual score and a passing score, with extreme differences concentrated at the later grades. Additionally, corporations may be more likely to make available remediation programs for later grades and students may be more likely to take advantage of them. We estimate the distribution of the total cost among exams at the grade levels as follows: 2/9 for the cost of any exam; 1/10 for extra costs for the 10th grade exam due to the GQE; 1/90 for extra costs for a 10th grade exam due to age rather than the GQE. We thus estimate the 10th grade share of the remediation funds at 1/3 of the total, and estimate the portion of this attributable to the GQE, as opposed to being an exam for older students, at 30%, 10% of the total funds, or \$2 million. The remediation funds require a corporation match of 50% of the total, or another \$2 million. Additionally, there is \$5 million in state funds available for remediation for the GQE alone. We estimate that these funds would not be available if the GQE were not a high stakes test, so the total remediation resources for the GQE linked to state programs is estimated at \$9 million. The resources for summer school and for remediation under state grants are allocated to corporation and school efforts.

On a per pupil basis, this state-induced funding of local programs accounts for \$15 of the \$442 presented in **Table 1-5**, as part of the 29% of funds spent on remediation. While the estimates of the graduation-qualifying share of expenses might be raised or lowered under different assumptions, no assumptions would lead to an estimate that the state directly induces or pays for a sizable share of the local program expenses.

ESTIMATE OF THE STATE-LEVEL YEARLY COST OF THE GQE

Thirteen million dollars in direct appropriations for ISTEP+ and \$8 million from the secondary market (interest) fund the state-level programs in the area of assessment. The costs currently are \$44.50 per student tested, though the cost should drop soon under a recently renegotiated contract. Costs for the GQE are somewhat higher than those for other tests: the construction of open-ended questions is expensive, especially since they must be changed every year; 10th graders write longer answers; 10th grade open-ended items are double-scored; and rescoring requests are higher for the GQE because it is a high stakes exam. While the costs are part of a broad contract, the contract would have been different had there been no GQE. We estimate costs of the 10th grade exam at \$59.70 the per student tested (one-third greater than the average) under the current contract, and estimate that 30% of this, or \$17.90 per student tested, is due to the graduation qualifying aspect of the exam, leading to double-scoring and high rescoring, for instance. As 70,973 students took the GQE in the fall of 2002, \$1.3 million is the estimated cost of developing and administering the GQE as part of a state-level program.

The Communication division of the Department of Education produces public service announcements, pamphlets, and a teacher-targeted publication all focused on the GQE at a cost of over \$0.3 million, and with the addition of an Assessment-funded staff person (salary and benefits), the cost of these programs reaches \$0.4 million. Additionally, the Commission on Higher

Education picks up the printing costs for a parent-student guide to remediation on the GQE. We assume that a low-stakes 10th grade examination would have few, if any, of these costs.

Within the Assessment division, the GQE should bear some share of the \$5 million in expenses for developing, administering, and following through on tests. Based on our estimates so far, where the GQE costs about 50% more than other exams, with 90% of the difference due to the graduation qualifying aspect, the GQE's share of expenses would be 10%, or \$0.5 million, of the departmental budget. This includes salaries, costs of conferences and workshops to test items and set cut scores, and costs of developing an item-sampler as a test guide, as well as a share of the department's 10-person website team.

Overall, then, an estimated \$2.2 million is the cost of state-level programs for the GQ aspect of the GQE. This is about \$2 per pupil.

Total Yearly Costs

The total yearly costs of the GQE are estimated to be, per pupil in the public school system:

Local programs	\$442
State-level programs	\$2
TOTAL	\$444

Included in the local amount is \$15 for state-sponsored local remediation and summer school programs, \$12 paid by designated funds from the state and \$3 paid by localities.

Because the local share of costs is virtually the entire amount, most of the discussion of total costs is simply the discussion of local programs. Additionally, one-time initial costs of \$12 per pupil occur at the local level.

Reporting results on a yearly per pupil basis accurately represents the costs in terms of grade level. The panelists report spreading resources far more evenly spread across grades than the notion of a high school exit exam would suggest. Remediation resources are devoted more to remedying of failures on the 3rd, 6th, and 8th grade examinations collectively, than to remedying solely the failure on the GQE itself. Professional development programs are targeted at teachers at all grade levels. Specialists are brought in to help as a preventive measure in elementary and middle school. Special instruction materials and supplies are targeted to P-8, while software is aimed equally at all grade levels. Only in-school tutoring during the school year is targeted primarily to high school students.

The emphasis on yearly per pupil costs misrepresents the reported use of resources in implying a type of regular, routine expense. While this may be true of professional development in part, panelists in general emphasized critical intervention. Most of the expenses occur for programs that seek to provide some missing ingredient that would allow the normal routines of education to operate effectively. The ingredient might be teaching teachers how to translate standards or test results into teaching, might be software, or might be early reading diagnosis. Given this emphasis on intervention to get education back on track for an individual student, a better way of reporting costs might be in terms of the cost due to the GQE across thirteen years of schooling. That is, when the expense occurred is less relevant than the fact that the bill must be paid at some point. Because educators tend to believe the cost is less if the problem can be solved early, resources are often spent in the early grades.

However, for students who presently lack that early help or who have motivational problems, the cost may come now. If students attend school for an average of 12.5 years, then the cost of the GQE across the student's public school education is \$5,550, out of a total cost for the student across that schooling of over \$100,000.

One context for understanding the cost of the GQE is the total cost of public education per student in average daily attendance, \$8,128 in Indiana in 2001–2002. The exit exam resources use up 5.6% of the total spent, about one of every eighteen dollars, or \$442 million in a budget of \$8.09 billion. The cost of any single grade level averages 8% of the total, so the cost of the GQE is most of the cost of a year of school. Whether this is too much or too little will depend on the evaluation of the GQE, which is outside the scope of this report. If the GQE is seen as a distracting waste of time, then any amount will be too much. If the GQE is seen as the crucial pinnacle of a program of standards-based testing that vastly improves an otherwise ineffectual education, then no amount will be too much. In the middle, the amount may be judged by what it is spent upon, and the funds are spent less on testing itself or even high school remediation than they are spent on remedying earlier failures, preventing failure, supporting intervention specialists, and training professionals, all linked into the system of local schooling.

This concludes the discussion of the costs of the current level of performance on Indiana's GQE. Next comes the consideration of the question of what it would cost to improve the student performance to the minimum level specified by the state as being "commendable" for a district.

The Cost of Improving Performance on the Graduation Qualifying Examination

Panels estimated the cost of raising the initial pass rate on both the mathematics and language arts examinations to 75%, from the present 60%, and raising the share of students who pass the test within 2.5 years from 86% to 94%. No one suggested that this change could be produced quickly, especially for students in the upper grades who have most of their education behind them. Therefore, the programs proposed emphasized long-term strategies. While panelists might have wished for a blank check, they were told to expect that their proposals would be reviewed by skeptics and should be a defensible means of achieving the goals.

One-Time Costs and Shifted Resources

In a departure from the estimates for current costs, the panels focusing on needs for changed outcomes (the what if panels) did not consider shifted resources to be a realistic way of paying for improved performance on the exit exam. Panelists felt that current resources were already successfully employed for high priority tasks and that shifting resources would not necessarily help performance for the long term.

The what if panels recommended three types of initial, one-time costs: teacher training, technology, and materials. While all have ongoing expenses, the bulk of the cost occurs at the time of initiation. In sum, the cost of these programs amounts to \$26 per pupil or about a \$2.7 million one-time statewide expenditure.

State-Level Programs

Panelists and interviewees have not been able to arrive at the estimates of the state-level costs of change. One reason for the lack of estimates is that the causal connection between state-level programs and student performance on the GQE is not clear.

The current state-level programs do not seem to be a bottleneck for future progress, so there has been little support for substantially increasing the funds spent on testing, communication, or high school remediation. Rather, the panels emphasized the ability of the state to implement more effective reading programs in the early grades and stressed the capacity of teachers to learn to teach to standards and to translate exam results into targeted teaching. These developments would not primarily be state-level programs, though five state-funded professional development days have been a DOE legislative priority since 2000.

State priorities in part seem to be to avoid change for the worse rather than to achieve change for the better.¹⁴ In the short run, scores may drop rather than increase when Algebra course content is included in the exam in 2004–2005. Even with several years of preparation, the current changes in the teaching of 4th to 6th grade math will not entirely show up in the initial results. While the state can implement programs that have fairly immediate effects on 3rd grade tests, which reflect 1–2 years of teaching, affecting 10th grade results involves affecting what happens nine years before the test. Some of the most effective state-level programs for future GQE results may be the current efforts aimed at improving early grade test results.

In the long term, according to the review panel, the state-level change most influential in raising future GQE scores might be changing the curriculum of state universities so that future teachers are trained on how to teach to standards, how to interpret test results, how to translate them into teaching, and how to motivate students for high-stakes testing. We offer no cost estimates for the GQE share of this change because the causation is partial and indirect.

Estimate of the Yearly Cost of Local Programs for GQE Improvement

Table 1–7 presents the estimated yearly cost of local programs to bring the GQE results up to the minimum “commendable” level. The cost per pupil of improvement is higher than the cost of the existing program. Given the realities of motivational problems and the difficulties of raising achievement for students with disabilities and low-income students, the commendable level may be a practical maximum of statewide average performance. Reaching that level would, in the judgment of panels, require an additional cost of 150% in resources devoted to the GQE. By itself, this would be an 8.4% increase in per pupil expenditures or an additional one dollar for every twelve currently spent. It is an open question whether the legislature or the public would be willing to make this expenditure in the future in order to accomplish the improved performance.

As with current costs, school-level personnel constitute the most sizable expense. Personnel at corporation headquarters are a relatively minor expense. The corporation’s larger expenses would be for expenses such as materials, training, consultants and transportation. Aside from personnel, the major costs at the school level would be for instructional supplies.

14 The *DOE 2003-2005 Budget Issues*, <http://ideanet.doe.state.in.us/legwatch/2003/page02.html>, stated top priorities do not include the programs proposed by panelists. However, teaching to standards and professional development to learn how received top support at the August 2002 P-16 Pathways to College Policy Forum, full-day kindergarten is already being created with the General Assembly funding, and Early Intervention Grant research-based programs have been positively evaluated by the Indiana University Education Policy Center.

The programs for improved scores feature preventing student failure and professional development, which together account for one-quarter of the cost (see **Table 1–8**). By contrast, remediation and testing move from nearly one-half of the current costs to slightly over one-quarter of the resources for improvement. While the suggested programs were diverse, the overall emphasis was on intervention as early as possible. Although testing and remediation programs such as improved remediation materials or transportation for summer school or better tracking of student test performance over time do attract improvement funds, the focus is less on test-oriented programs than it is on programs that could be justified on broad educational grounds such as improving reading or teaching to standards. The reasons that they count as costs of improvement on the GQE are: 1) they are perceived as needed for improvement on the GQE; 2) they are perceived as unlikely to occur without the spur of the GQE; 3) achieving “commendable” success on the GQE as a result of the programs would lead to the continuation of the funding.

An emphasis on programs that directly improve existing teaching is evidenced in the distribution of resources among the personnel. Costs associated with teachers account for 72% of the proposed funding, while costs associated with specialists are 19%, notably less than the estimated one-third they receive of current expenditures on the GQE. The other personnel costs are for special assessment roles bringing assessment into the existing structure, such as an assistant superintendent of assessment and test-into-instruction coaches for teachers.

Two-thirds of the non-personnel costs for the proposed programs occur for instructional materials and supplies to support diverse programs in remediation, prevention, and professional development. Almost all of the remainder occurs for expenses incidental to the programs, notably for transportation for after school and summer school.

Table 1–7: Per Pupil Local GQE Improvement Costs*

	SCHOOL	CORPORATION	TOTAL
Personnel	\$ 521	\$ 39	\$ 561
Other	\$ 50	\$ 74	\$ 125
TOTAL	\$ 572	\$ 114	\$ 685

* Totals are correct because other data in the table have been rounded to the nearest dollar.

Table 1–8: Local GQE Improvement Program Cost

	Share of Costs
Remediation	16%
Prevention	43%
Professional Development	31%
Testing	10%
Other	1%
TOTAL	101% *

* Percentages do not add up to 100% due to rounding.

Overall, the panels' recommendations for improving GQE performance focus on ways to improve broad educational performance. Reading, for instance, is stressed because it is seen as a crucial prerequisite for accomplishment in both mathematics and language arts. Panelists discussed current proposals for broader educational reform, including full-day kindergarten, pre-school, and alternative yearly calendars redistributing summer vacation time. We have not included costs for these reforms except when, as with the Jump Start program, the panelists directly tied them to improved performance on the GQE. The panels' implicit assumption has been that performance on the GQE traces to skills taught in the standards-based curriculum that almost all children should learn.

The Total Cost of Improvement on the GQE

As with current costs, the perceived costs of change are largely for yearly local programs. The cost of classroom teachers accounts for the largest share (\$521) of the estimated per pupil cost of improving performance to the commendable level for the average district. An additional \$26 per student would be a one-time charge for supporting materials and equipment. In contrast to funds spent simply to teach or funds spent to remediate and test, the proposed expenditures aim at critical interventions in the current education process that would, in the eyes of our well-regarded participants, yield improvements in the GQE performance.

Total yearly per pupil cost of improvement: \$685

Across a half-day of kindergarten and grades 1–12 for a single student, the \$685 yearly cost of GQE improvement translates into \$8,562 per student for a full public school education, on top of the existing \$101,600. In a year, the total cost would be \$682 million. The cost of improvement would be slightly more than the average cost of a year's instruction.

Averages can be misleading when the costs vary widely. Education costs in general are higher for at-risk, special education, and small-district students. Given the current rates of failure, the costs of improving GQE schools would disproportionately apply to the task of improving the performance of students with difficulties.

At this point, a methodological digression is called for. In calculating improvement standards for exceptional learners, we took the reduction in the overall rate of initial failure on the GQE for all students as a baseline. The minimum commendable level reduces this rate by 37%, from 40% to 25%. We assumed an equivalent reduction in the rate of failure on the math and English exams considered separately. In calculating improved initial pass rates for special education, at-risk, and African-American and Hispanic students, we took the overall rate as a base and then adjusted for how difficult it would be to lift these students' scores; how many resources were dedicated to these students; and whether the students' scores could be improved by less expensive collective solutions or would require more expensive individual remedies, remedies that might wait until after the student's initial failure on the GQE. Therefore, we hypothesized a decline of 32% in the initial failure rate of special education and minority students and a decline of 24% in the initial failure rate on the GQE of students eligible for free or reduced-price lunches, as compared with a decline by 37% for the average student. That is, we expect reductions in the initial failure rate for most of the hardest to teach students to be near the levels for the average students.

Even within special education, the cost of improvement varies with disability. The severest disabilities have no cost associated with the GQE presently or in the future. Communications disorders, prevalent among special education students in the early grades,

largely disappear by the 10th grade. Only 3% of students with mild mental handicaps who take the GQE pass the test, a rate unlikely to dramatically change. The important group is those with learning disabilities, who comprise the bulk of the special education students who take the GQE. Because the passage rate is low, most of these students are retaking the exam. If the initial pass rate could be raised to 45%, the total number of such students taking the exam could be cut by one-quarter or one-third, due to fewer retakes, with associated savings in preparing for the test. However, because learning disability diagnosis often comes relatively late, compared to other disabilities, these students are hard to target for early intervention. For learning disabilities, then, the cost of improvement might involve new funds and shifting funds from remediation to prevention in upper elementary school.

We have not attempted to estimate either the current cost of the GQE or the cost of improvement for various categories of students or for schools with different compositions. In an adequacy study of Indiana last year, our estimate was that special education students cost about twice as much to educate as did average students and that hard-to-serve at-risk students cost substantially more than half-again as much as the average. In lieu of targeted estimates, the broad estimates might stand as starting points for a consideration of the concentration of GQE current and improving costs among categories of students. Removing the achievement gap will require investing more in the groups that are behind but that can improve. The cost per school district would vary dramatically across the state. While some corporations have more than 80% of their students passing both English and mathematics on their first attempt, the highest rate among the 29 schools in Gary, South Bend, and Indianapolis is 51%, with only two schools above 38%. Some districts have four times as large a share of students failing as do other districts.

In summary, Indiana's well-developed assessment system is capped by a 10th grade Graduation Qualifying Examination in English and mathematics. Despite a hefty initial failure rate, marginal students usually pass in time to graduate. APA created a microcosm of the state for a hypothetical district. For current levels of performance on the GQE, panels of qualified service providers indicated that the costs primarily occurred at the local level, notably for school personnel. The magnitude of the costs, \$444 per pupil, dwarfed budgeted state expenditures for the GQE and went for a variety of testing, remediation, professional development, and prevention programs. By shifting resources from other programs and students, administrators covered the cost of the GQE without additional budgeted state funding targeted to the GQE. The estimated cost of meeting the minimum state prescribed standard of performance, \$685 per pupil per year, would be 150% above the cost of the current level of performance. In the view of the panels, these funds could not be shifted from existing programs but would be devoted to prevention and development interventions at critical junctures that promise educational improvements well beyond simple performance on tests.

The cost estimates indicate that Indiana has moved well beyond the approach of simply testing students to motivate them to improve themselves. The emphasis on remediation for state-sponsored programs represents a further step. Current costs, however, go beyond remediation as local districts and students attempt to improve their initial pass rates by targeting problems before students arrive in high school. The emphasis on prevention increases as panelists considered future improvements, up to the state-defined "commendable" level. However, testing and remediation continue as important components of improving scores.

CHAPTER 5

Methodological Summary and Recommendations

Legislative decision-makers choose among programs and features of reform in light of their principles and of the desired consequences. Service providers, by contrast, concentrate on implementing programs within the context of the students and school they deal with daily. In the middle, educational supervisors and specialists think in terms of resources and priorities among competing demands. The aim of our methods is to draw upon the experience of providers and administrators in order to formulate cost considerations in terms useful to policymakers. We ask those with respected experience to address “what is” and “what if” questions of the cost of standards and programs. APA’s chosen method allows for the estimation of the costs of current and improved performance no matter what the state’s approach to exit exams, so it could be useful to decision-makers no matter the state of development of their exit exam arrangements.

Summary of Approach

Interviews

In the professional judgment approach, experts are asked what resources would be needed to supply a given level of educational output or outcome. Experts in school expenses and finance then translate these resources into costs. APA starts out with a hypothetical, average district that is close to the experience of the participants. We focus on resource use under current levels of performance. Later, estimates of required resources are translated into costs.

Ought and Is

The task of a second group of panels is similar except that it focuses on outcomes at a higher, state-certified level. To the degree that the states are not wholly satisfied with the preparation of students for high stakes testing, we ask informed experts for what else would be needed to meet state standards of adequacy. We combine cost and resource approaches, as well as estimates of current and adequate resources. At the state level, we focus on actual costs for current programs. At the state level, it is easier than at the district level to separate out expenses by their purpose and program, and there is not the problem of highly varying district programs and expenses, so we rely more on existing documents.

Direct or Indirect

The focus is overall cost of producing a specified outcome or output, including costs that develop over time and those that come from adjusting priorities. This is also the approach taken by Michael Rebell and colleagues to figuring the cost of an adequate preparation for the New York Regents exams, in which they broadly consider all costs of preparing for the exams. Our approach differs because we wish to consider the cost of alternative programs

and policies for exit exams. When policies are compared to each other for a single state or across states, marginal pricing becomes appropriate. We take a marginal approach to the cost of an adequate preparation for and conduct of high school exit exams.

Exit or Exam

The marginal cost of exit exams can be considered as the entire package of cost differences between the state's exit exam policies and an alternative set of policies lacking an examination required for graduation. We get at this cost by asking about the resources required for the current testing situation that would not be needed without the graduation requirement. Because the No Child Left Behind Act mandates high school testing, the cost of testing per se is, in our approach, distinct from the cost of high stakes testing. The programs associated with high stakes testing—remediation, retakes, counseling—have the costs we are interested in estimating. Then we are interested in estimating the additional costs of accomplishing the state policy regarding the level of performance on these high stakes examinations.

Case or Compare

To some degree, every state is unique. This uniqueness is evident in the actual content and standards of the graduation exams themselves. No matter what we try, there is an irreducible basic difference among states in their exit exams and hence in our cost estimates. By focusing on an average district for each state rather than the same hypothetical district in every state, we choose a state-centered analysis. Additionally, we focus on the examination system as it exists in each state, rather than a hypothetical package that would be the same in every state. The state-mandated standards for adequacy are again unique to the state, and, in principle, standards could be identical in each state. Along with pricing and other aspects of the school arrangements, these methods mean that it will be difficult to directly compare costs across states. We do use some methods that aid comparison. By focusing on resources that are later translated into costs, we allow the comparison of resources across states with different costing arrangements. Further, the hypothetical changes permit inferences about similar exams.

Present and Future or Past

Implementing officials face change caused by exit exams as a package, perhaps spread over time. Funding may be provided under the name of a program with either tight or loose control on the actual use of the funds. In any given year, district superintendents must request and organize the resources needed for all programs that year and will allocate available funds according to priorities. The programmatic labels or funding sources for resources are not a reliable guide to the actual delivery of services. Instead, we use a hypothetical, average district and hypothetical questions to elicit information about the allocation of resources toward exit exams.

Summary of Results

Panels produced estimates for the cost of the Graduation Qualifying Examination in Indiana. The three panels provided similar estimates of the cost of current resources needed for school and district programs for the GQE. The overall estimate is \$444 yearly per pupil, which is concentrated on providing personnel at the school level. The programs combine testing, remediation, professional development, and prevention and operate at all grade levels. The contribution of state-level programs and start-up costs are relatively minor. Shifting

resources from other programs and students produces most of the resources used for the GQE, which total about one dollar of every eighteen spent on Indiana public education.

The estimated cost of improving education up to Indiana’s minimum “commendable” level is \$685 per pupil, all at the local level. The state programs are not the bottleneck for improvement. Panelists recommended programs emphasizing prevention and professional development to make interventions that allow normal education procedures to be more effective with students at risk of failure on the GQE. Panelists rejected the suggestion that a substantial share of the resources needed for improved performance on the GQE could be shifted from other programs. Initial costs of \$26 per pupil provided for one-time investments in training, technology, and software. According to these estimates, the state’s targeted improvement on exit test results will not come automatically from current programs but will instead require additional funds targeted at education improvements that also improve test scores.

Panels produced no estimates of the cost of different features of exit examinations. While specific programs are tied to the number of teachers or the number of students, no linkage between specific components and cost has been made. Rather, the cost estimate is for a typical district and the entire state of the GQE as it is, at current levels of performance and at improved levels. The estimate should be viewed as being reliable within about +/- 10%.

Recommended Revisions in Methods

Studying a set of states in a single year would clear up questions and test assumptions. For instance, results from states at different stages of development would be needed to answer the question of whether the costs of exit exams increase over time as remediation and prevention approaches become pronounced. Comparison of the costs in states with exams at similar stages of development would shed light on the role of student and exam characteristics.

All panels should be provided with a single clear task, a short list of programs and resources, and three hours of time. Across panels, membership should be diverse in age and race as well as in occupation and type of district.

Given practical limitations on the number of panels, consideration should be given to the tradeoff between having multiple panels address the same task, which provides an error check of diminishing value per panel, and having tasks that are more diverse. The number of panels estimating current costs could be reduced to two. Additional possible types of tasks include having a district with a greater than average share of hard-to-serve students, permitting inferences about how the student composition of districts influences costs; having panels address exit exam programs for special education or at-risk students alone; and addressing questions of the cost of changes in the exit exam features.

Questions about programmatic changes to the exit exam cannot be simply added to a panel’s task. An entire panel could be devoted to any of the major differences that divide state exit exams, such as the overall pass rate or the availability of accommodations for students with special needs.

The techniques to ease cross-state inferences—hypothetical changes, tying costs to programs, specifying costs by resources used—have not been notably successful. An alternative would be to create a scenario—an exam, a district, and standards—that was used in identical form in every state for at least one panel.

The best dates on which to visit states may be the end of September, the second week in November, mid-February, and mid-April, depending on the timing of Easter. From the week before Thanksgiving to late January, arranging and conducting panels is difficult, as it is toward the end of the year and during the summer. Panels require an early in-state visit and two months to arrange ahead for participation, so a September panel requires arrangements before the summer vacation.

One possibility within a multi-state study would be to investigate one special topic in each state, providing at least a benchmark for costing in all states. For instance, the topic of shifting resources toward exit exams under changing priorities might be the subject of an entire panel in a single state. Similar emphasis might illuminate the role of federal funds and grants in developing programs for exit exam performance. An in-depth examination of special education resources for an exit exam in one state might help to understand patterns in every state.

For special education, where cost data cannot easily be garnered, APA might develop cost models that could be verified and corrected by state and local officials. The costs of accommodations and alternative routes to a diploma cannot be calculated with existing data.

After the refinements suggested have been tested in a study, the method would be ready to be applied to populous states that require a large, diverse set of panels exploring a variety of topics or to rural states with small districts and a dispersed population.

CHAPTER 6

Conclusion

In this study, APA provides to the Center on Education Policy a method to estimate the costs of high school exit examinations. After a series of methodological choices, the core innovative technique has been the adaptation of professional judgment panels to the estimation of exit exam costs. The method has been created to be appropriate to all states with exit exams, and has been tested in Indiana. The test case validates the basic method, which produced cost estimates, and has led to recommendations for specific revisions that refine the professional judgment approach.

Indiana has a well-developed statewide assessment program with leadership at the top from the Department of Education, the Governor, the General Assembly, and the Education Roundtable. Nonetheless, the main estimated costs of the GQE are borne at the school level and the corporation level, because Indiana has moved beyond the testing and even remediation approaches. Even state programs specifically supporting testing and remediation at the local level fund only a small fraction of the local resources employed because of the GQE. Service providers in our panels responded to the demands of the exit exam with a variety of programs for test administration, remediation, professional development, and failure prevention. Local leaders show every sign of being flexible and creative in adapting to the changed priorities represented by the GQE by shifting resources from other programs and purposes.

Improving average GQE results up to the minimum commendable level for a school district requires an estimated additional investment in excess of current spending on the GQE. Further, the funds would be targeted primarily into prevention and professional development programs that seek to intervene in a way that increases the productivity of the existing educational system. That is, raising the GQE scores means improving such educational basics as reading because they are requisite to acquiring advanced skills, including mathematics.

For future work, it would help to study a set of states at the same time to aid comparison. Also, several of the recommended revisions—applying identical scenarios in each state, holding panels on special topics with results that might be useful in all states, creating estimates of the costs of exit exam features one by one across states—presume the ability to plan out at least a multi-state design.

APPENDIX I

District Panel Instructions, Current Costs

Instructions to High-Stakes Testing School District Panel

Augenblick & Myers, Inc.
Denver, CO

January 14 & 15, 2003
Indianapolis and South Bend, IN

1. You are a member of one of two panels of people that are being asked to consider programs for the Graduation Qualifying Examination in a prototype school district. The prototype school district is hypothetical and does not actually exist. It is a convenient way to specify the resources that schools need to have in order to accomplish a specific set of objectives regarding the high school exit examination.
2. Each panel is being asked to indicate the resources needed to provide students an average opportunity to pass the Graduation Qualifying Examination (or an equivalent path to a high school diploma). Please identify the resources that are needed for the examination as it is and that would not be needed if the examination were not a requirement for high school graduation. That is, what resources would be saved if the GQE had always been just another ISTEP test? These resources may or may not be the ones being used in your school.
3. The characteristics of the prototype school district are:
 - 5350 students in 11 schools
 - 6 elementary schools (K-5) with 400 students each
 - 2 middle schools (6-8) with 600 students each
 - 2 high schools (9-12) with 800 students each
 - 1 alternative and vocational school with 150 students
 - 16% of students (about 875) are in special education
 - 22% of students (about 1175) are eligible for a free lunch
 - 3% of students (160) are in programs for limited English proficiency

4. The GQE objectives accomplished by the prototype schools and district under the 2002 standards and procedures are:

■ Initial (10th grade) total pass rate on both exams	60%
■ Pass rate for English/Language Arts & Mathematics, singly	68%
■ Students passing both tests within 2.5 years	86%
■ Special education students taking the test (84% of them with some type of accommodation)	11%
■ ESL/LEP students taking the test (66% of them with some type of accommodation)	3%
■ Initial pass rates for special education students	19%
■ Initial pass rates for free lunch students	45%
■ Initial pass rates for black and Hispanic students	40%
■ Initial pass rates for ESL/LEP students	28%

5. We need you to provide some specific information so that we can calculate the cost of the resources needed to meet the objectives identified above. The fact that we need that information should not constrain you in any way in considering the programs of a prototype school district. Your job is to outline a set of programs designed to serve students with diverse needs in such a way that the objectives specified above are fulfilled. Use your expertise to organize personnel, supplies and materials, and technology in any way that you feel will produce the desired outcomes.

6. For programs you believe are needed but that are best provided at the state level, assume that the state does provide them and make a note of this assumption.

7. There is no single “right” approach. Your task is simply to create a reasonable and effective analysis.

APPENDIX II

District Panel Instructions, Improvement Costs

Instructions to High-Stakes Testing School District Panel

Augenblick & Myers, Inc.
Denver, CO

January 14 & 15, 2003
Indianapolis and Bloomington, IN

1. You are a member of one of three panels of people that are being asked to design programs for the Graduation Qualifying Examination in a prototype school district. The prototype school district is hypothetical and does not actually exist. It is a convenient way to specify the resources that schools should have in order to accomplish a specific set of objectives regarding the high school exit examination.
2. Each panel is being asked to indicate the resources needed to provide students an increased opportunity to pass the Graduation Qualifying Examination (or an equivalent path to a high school diploma). Please identify the resources that are needed to accomplish the objectives. These resources may or may not be the ones already being used in your school district—some current resources might not be needed while others might be needed that are not used presently.
3. The characteristics of the prototype school district are:
 - 5350 students in 11 schools
 - 6 elementary schools (K-5) with 400 students each
 - 2 middle schools (6-8) with 600 students each
 - 2 high schools (9-12) with 800 students each
 - 1 alternative and vocational school with 150 students
 - 16% of students (about 875) are in special education
 - 22% of students (about 1175) are eligible for a free lunch
 - 3% of students (160) are in programs for limited English proficiency

4. The objectives that need to be accomplished by the prototype schools and district under the 2002 GQE standards and procedures are:

AN INCREASE IN:	FROM	TO
■ The initial total pass rate on both exams	60%	75%
■ English/Language Arts & Mathematics, singly (the minimum average pass rate for a commendable school)	68%	80%
■ Students who pass both tests within 2.5 years	86%	94%
■ Special education students taking the GQE (85% of them with some type of accommodation)	11%	14%
■ Initial pass rates for special education	19%	45%
■ Initial pass rates for free lunch students	45%	58%
■ Initial pass rates for black and Hispanic	40%	59%
■ All students failing a single exam retake the next year, as do 90% of those failing both exams.		
■ One-half of students who remain in school and who do not pass both the GQE tests nonetheless graduate with their class under alternative procedures and standards.		
■ Three percent of those taking the test are English as a second language or with limited English proficiency, 65% of whom have an accommodation.		

5. We need you to provide some specific information so that we can calculate the cost of the resources needed to meet the objectives identified above. The fact that we need that information should not constrain you in any way in designing the programs of the prototype school district. Your job is to create a set of programs designed to serve students with diverse needs in such a way that the objectives specified above are fulfilled. Use your expertise to organize personnel, supplies and materials, and technology in any way that you feel will produce the desired outcomes.
6. For programs you believe are needed but that are best provided at the state level, assume that the state does provide them and make a note of this assumption.
7. There is no single “right” approach. Your task is simply to create a reasonable and effective package. This might focus on summer school remediation, standards-based curriculum, targeted preventive education in elementary school, tutoring or anything else that you can agree on.

APPENDIX III

Review Panel Instructions

Instructions to High-Stakes Testing Expert Panel

Augenblick & Myers, Inc.
Denver, CO

January 16, 2003
Indianapolis, IN

1. You are a member of a panel of experts, people who have been identified as having extensive knowledge of how schools and the Graduation Qualifying Examination operate and of the resources schools need to fulfill their GQE objectives. Your job is to extend the work of other panels that were asked to specify programs for the Graduation Qualifying Examination in a prototype school district. The prototype school district is hypothetical and does not actually exist. It is a convenient way to indicate the resources that schools should have in order to accomplish a specific set of objectives regarding the high school exit examination.
2. While there is only one expert panel, you will be reviewing the work of two types of district panels: 1) one set of three panels was asked to designate the programs and resources for the GQE as it currently exists as a high stakes examination; 2) a second set of three panels was asked to point out the resources needed to provide students an increased opportunity to pass the GQE. All panels were asked another, later hypothetical question about costs for a changed GQE
3. The characteristics of the prototype school district are:
 - 5350 students in 11 schools
 - 6 elementary schools (K-5) with 400 students each
 - 2 middle schools (6-8) with 600 students each
 - 2 high schools (9-12) with 800 students each
 - 1 alternative and vocational school with 150 students
 - 16% of students (about 875) are in special education
 - 22% of students (about 1175) are eligible for a free lunch
 - 3% of students (160) are in programs for limited English proficiency

4. The GQE objectives accomplished by the prototype schools and district under the 2002 standards and procedures are:

■ Initial (10th grade) total pass rate on both exams	60%
■ Pass rate for English/Language Arts & Mathematics, singly	68%
■ Students passing both tests within 2.5 years	86%
■ Special education students taking the test (84% of them with some type of accommodation)	11%
■ ESL/LEP students taking the test (66% of them with some type of accommodation)	3%
■ Initial pass rates for special education students	19%
■ Initial pass rates for free lunch students	45%
■ Initial pass rates for black and Hispanic students	40%
■ Initial pass rates for ESL/LEP students	28%

The objectives that need to be accomplished by the prototype schools and district under the 2002 GQE standards and procedures are:

INCREASE	FROM	TO
The initial total pass rate on both exams	60%	75%
English/Language Arts & Mathematics, singly (the minimum average pass rate for a commendable school)	68%	80%
Students who pass both tests within 2.5 years	86%	94%
Special education students taking the GQE (85% of them with some type of accommodation)	11%	14%
Initial pass rates for special education	19%	45%
Initial pass rates for free lunch students	45%	58%
Initial pass rates for black and Hispanic	40%	59%

- All students failing a single exam retake the next year, as do 90% of those failing both exams;
- One-half of students who remain in school and who do not pass both the GQE tests nonetheless graduate with their class under alternative procedures and standards;
- 3% of those taking the test are English as a second language or with limited English proficiency, 65% of whom have an accommodation.

5. We need you to review the work of the panels and find an overall set of resources for the present system and for a more commendable system.
6. Additionally, we would like you to review the estimates the district panels made in response to six hypothetical changes.
7. The next task of the expert panel is to check on the adequacy of our price estimates. We have some preliminary estimates of the cost of many of the resources on a separate list.
8. The final task is to estimate the resources needed for state-level programs under the present system, under the more commendable system, and under the hypothetical changes. We have an initial list of some current state programs.

APPENDIX IV

Participants

Corporation Panelists

Sandy Beaman	Ron Furniss	Sharon Pitts
Mike Benway	Karen Gould	John Prince
Jane Boultinghouse	Don Harrison	John Ritzler
Judith Bush	Russell Hodgkin	Lorene Sandifer
Mike Bushong	William Isaacs	Arlene Schultz
Kevin Caress	Tim Jackson	Ryan Snoddy
J.T. Coopman	Walter Jamrose	Tom Thornton
Edward Daihl	Virginia Land	Charlotte Totten
Tom Doyle	Jennifer McCreadie	Marvin Ward
Allan Essig	Sandra Nixon	Ruth Warren
Ron Etienne	Linda Nowling	Joanne Weddle
Tom Fletcher	Donna Osborn	
Joyce Fulford	Ray Pavy	

Review Panelists

Larry Grau
J. Brett Lewis
Susan McDowell
William Riley
Mark Shoup

Advisory Committee Members

Lowell Rose
Steve Heck
John Kline
Wes Bruce
Roger Thornton
Dennis Costerison

Hosts

Dan Clark, Indiana State Teachers Association (ISTA)

Ed Eiler, Lafayette School Corporation

George Kersey, Phi Delta Kappa (PDK)

John Ritzler, South Bend School Corporation

Interviewees

Wes Bruce, Assessment,
Department of Education (DOE)

Dan Clark, ISTA

George Kersey and staff, PDK

Robert Marra, Exceptional Learners, DOE

Kevin McDowell, General Counsel, DOE

Rep. Gregory Porter, Chair,
House Education Committee

Terry Spradlin, Legislative Affairs, DOE

Marc Steczyk, Communications, DOE

PART 2

Comparing High School
Exit Exam Costs in
Minnesota and Massachusetts

By Douglas Rose
Augenblick, Palaich and Associates

with contributors¹

Justin Silverstein
Augenblick, Palaich and Associates

and John Myers
JL Myers Group

A Report to the Center on Education Policy

FEBRUARY 2004

¹ Keith Gayler of CEP is not listed among the contributors only to avoid the appearance of CEP reporting to itself. His contributions are major and most welcome, as is the aid of CEP's Madlene Hamilton.

Executive Summary

This analysis is part of a multi-state study that examines costs of state high school exit exams. In the initial 2003 portion of the study (see Part I of this volume), we developed a method for studying exit exam costs and applied the method in Indiana, a state that in many ways typifies the exit exam movement. In this study, we report on the extension to Minnesota's less demanding test and to Massachusetts' integrated assessment, providing preliminary conclusions from a comparison of resource use in all three of the states studied. Later plans call for costs to be estimated for a number of other states including a smaller western state, a southern state with a sizable minority population, and a state whose large population offers the potential for economies of scale, after which we will offer firm conclusions about the patterns of costs for exit exams in the American states.

We define the costs of exit exams as the cost of those resources that would no longer be used for the current programs if passage of the examination were not required for student graduation. The resources currently required for success in state high school exams are \$171 per student in Minnesota and \$385 per student in Massachusetts for school and district level general education, less than Indiana's \$557 per student.

Using interviews and task groups convened in ten cities in two states, this research identified the programs and costs for four types of testing situations: the current exit exams, an exit exam reflecting cross-state average characteristics, an exit exam with raised standards, and for special education and Limited English Proficiency students taking the tests.

Minnesota. The Basic Skills Tests (BST) require an 8th grade mastery of mathematics and language arts. Compared to tests requiring command of 10th grade material, they are believed to be less difficult. Minnesota targets its BST programs to prevent initial failure, focusing school personnel on students who have had difficulty on earlier exams. Students needing remediation after failing the BST are offered summer school, special math classes, talented teachers, and other support programs.

The low difficulty of the BST causes the focus for improvement to be on students who have difficulty with exams. Special education students require resources that cost an additional \$26 per student because the BST is a graduation requirement. Exit exam or no exit exam, the school district should—and often does—provide special programs for these students that help bring their skills in language and mathematics up to an eighth grade level, so the marginal cost of the exit exam is relatively low.

State-level programs in assessment and oversight add another \$2 per student in Minnesota and \$7 per student in Massachusetts to the cost of exit exams. State programs that pass funds to the district level, such as grants for remediation programs, are included in the district-level spending.

Minnesota's BST involves fewer programs—for testing, remediation, professional development, and failure prevention—and much less cost than does Indiana's exit exam, the Graduation Qualifying Exam (GQE), reported on in 2003. Minnesota has a less demanding exam, a longer period in which to remedy student performance since the test is given

earlier in a student's education, and an exit exam not tied to a standards-based curriculum. However, many of the patterns of resource use are the same, beginning with the concentration of expenditures at the school and district level rather than the state level.

If the current rate of success were maintained, Minnesota educators estimated the added district cost of a 10th grade exit exam at \$377 per student, triple the cost of the current BST. Added to the cost of the current BST, the resources needed would exceed those required for Indiana's 10th grade GQE, and the resources would be more concentrated on school personnel and professional development.

For a hypothetical 10th grade, average state exit exam, with more students in difficulty and ultimately failing the exam than is the case in Minnesota, Minnesotans estimated a district-level cost of \$639 per student, the most expensive of the state's three estimates, with an emphasis on testing and professional development costs rivaling those of Indiana.

Massachusetts. The Massachusetts Comprehensive Assessment System (MCAS) includes a 10th grade test whose passage is a graduation requirement. In 2003, Massachusetts graduated its first class under the MCAS system, and educators graduated a high proportion of seniors. The MCAS has a relatively high initial pass rate, in part because it is nested within a statewide standards-based assessment system. Further, individual and aggregate results are closely monitored, and corrective action is mandated. For individual students, this includes a plan for improvement for any student failing the eighth grade exam. For districts and schools, this includes a state remediation plan.

The resources used at the district level because of the exit exam requirement constitute nearly 4% of the Massachusetts budget for K-12 education. The costs are spread fairly evenly among preventive programs, remedial programs, and professional development. The remedial programs have been chiefly funded by a \$50 million appropriation from the state.

Because MCAS is perceived to be a demanding 10th grade exam and because the per student costs for education in general are higher in Massachusetts than in Minnesota or Indiana, the cost estimates for the current MCAS seem low compared to cost estimates in other states. Further, the estimated cost of implementing the average state exam was dramatically lower than the estimate in Minnesota.

One key explanation stems from the high initial pass rate in Massachusetts, leading to low remediation and testing costs. Prevention as a strategy is cheaper than remediation, and Massachusetts' strategy of an integrated standards-based curriculum and closely monitored assessment program is a failure prevention strategy. Little of this prevention is identified as a cost of the exit exam, as it is associated with the broad package of reforms of which the exit exam was part.

Not all of the MCAS is low cost. The cost of preparing limited English proficiency (LEP) students to pass the exit exams was almost entirely omitted from the original panels' estimates of current resource use, so we convened a panel to focus solely on LEP costs. These Massachusetts panelists identified passing the MCAS exit exam as the top educational priority of LEP students, with considerable school-level resources developed to educate students so that they will be successful. The LEP emphasis is strongly on remediation. That adds \$101 per student, about one-third, to the total cost of the MCAS exam for a total cost of \$424 per student, LEP and non-LEP combined.

The cost of the MCAS would more than double if the standard for passing were raised from the “needs improvement” category to the “proficient” category. This is similar to the estimated impact of raising standards for Indiana’s GQE, and not far from the tripling of cost associated with changing Minnesota’s BST to a 10th grade exam. Significantly higher levels of performance on high school exit exams will involve the commitment of proportionally increased resources for education.

Comparison. There are enormous differences in the costs of the various exams. The major differences are more among exam situations than across states. To compare costs across states, the extra costs for special education students and for LEP students were added to all tests. We also adjusted costs by equalizing teacher salaries to an average salary.

After comparing results across states and types of exams, we find an important distinction between differences among exams and change in exams. More difficult exams are more costly—this is a difference among exams. In addition, however, the change to a more difficult exam involves extra costs, more than would be expected simply from the difference in exams. And the cost of change is asymmetric: improved education can cost a great deal, while change to a less stringent standard does not save an equivalent amount. Change in the initial failure rate is the most cost-consequential aspect of the exam. Using preventive programs to reduce initial failure rates from “high” to “moderate” or from “moderate” to “low” would reduce exit exam costs by an estimated \$380 per student in any state.

The costs of programs for success on the exit exams tend to follow the difficulty of the exam, from Minnesota’s BST to the raised standards for Massachusetts’ MCAS and Indiana’s GQE. The harder the test is, the higher the cost, at about \$280 per student, because more students will need help to get over the passing hurdle. In this sense, stringent exit exams stimulate more education, which has an additional cost. Higher initial failure rates are also linked to higher programmatic costs, especially for remediation, of \$266 per student. Finally, increases in the difficulty of the exam add \$134 per student to the costs, while changing to a less demanding exam is relatively costless. While tougher exams are more costly than easier ones, changing to a tougher test from an easier one adds to the cost and ends up with a higher cost than that incurred by adopting the tougher test to start with.

These costs can be added together, so maintaining an exam of low difficulty and a low initial failure rate would be relatively inexpensive, about \$66 per student. By contrast, an exam as difficult as the “proficient” MCAS, which represents a change from an exam similar to the current MCAS, combined with a high initial failure rate, which represents a drop from an even higher failure rate, could cost about \$2,300 per student, about 30% of the Massachusetts K-12 education budget. Both differences in the type of exam and changes among types of exams have costs, so the most expensive option is to change to the most difficult type of exam. To reiterate, these are the estimated costs of the educational programs to produce the success measured by the exit exam.

The recent introduction and reform of exit exams has the effect of narrowly focusing programs and costs on testing and remediation, while there is a tendency over time for an established system to evolve programs and costs incorporating prevention and professional development. However, the most notable impact of recency (the freshness of reform) is to change the perception, rather than the reality, of exit exam costs. In situations of recent reform, it is often easy to identify the cost of the reform package, though not of individual reforms. By contrast, once reforms have become an integrated part of the education system, it is difficult to identify the cost of individual parts of that system. Recent exams have costs

that are difficult to separate from the cost of associated reforms, while established exams have costs that are difficult to separate from the rest of the ongoing system.

A major finding is the important distinction between exam structure and exam change. Differences between states in structures have lesser cost consequences than do changes within states in structures. Change produces a different pattern of resource use, notably a surge in prevention programs and a shift away from remediation as the exams become more difficult or the failure rate becomes higher.

The Next Stage. The results in Indiana, Minnesota, and Massachusetts may not extend to differing states: large and small states, southern and western states, states with less educated populations, states with large minority populations, and low-performing states might show quite different patterns. More states and more variation are required before these cost estimates can be relied upon. Additionally, increased variation in state size might indicate whether there are economies of scale in providing exit exams and supportive structures.

Additional ideas for development in the methods of estimation include the following: add a single all-day panel that considers current costs in the morning and costs for a changed exam in the afternoon; conduct a panel aimed solely at an atypical, urban district with a high concentration of students having difficulty passing statewide exams; conduct an intense study of a single district, including close examination of its records; and let the special focus panel build from the district resource estimates of the general panel.

Conclusion. If the standards-based curriculum and integrated assessment are working, then it is comparatively inexpensive to administer a high stakes test, at least one with “needs improvement” as its threshold of passage. That appears, so far, to be the lesson of Massachusetts’ MCAS.

A basic examination is also relatively inexpensive once it is underway, as is the case in Minnesota. However, an exam such as the BST that is not integrated with a standards-based curriculum and with other assessments can be expensive to change.

The pattern is for costs, especially remediation costs, to increase with increases in exam difficulty and the initial failure rate. However, the largest costs are the costs of change. It is cheaper to get exit exams properly aligned from the beginning.

Because the exit exam costs occur primarily at the district level, they may be relatively invisible to state policymakers. Across states, the spending occurs in accord with the exam structure and difficulty. Using task groups to help estimate resources and costs, state policymakers can thus anticipate the education and the accompanying expense.

Introduction

This report covers an investigation of the costs of state high school exit exams in two states, Minnesota and Massachusetts. It compares results between these states and between them and Indiana, the subject of a similar study in 2003. Cost reflects the programs consuming resources in order to generate educational success. Exit exams involve distinctions between the education that they measure, the assessments themselves, and the changes in education and behavior that a high stakes exam produces.

To avoid studying all of K-12 education in high stakes test states, we limit our study in two respects. We only identify resources that have identifiable costs to the public school system. Hence, we do not input the value of volunteered time, private tutors, or psychic costs and benefits of failure and success, and we do not focus on the role of such resources as parental education. Additionally, we study those costs that depend on the existence of high stakes testing, which would be incurred differently or not at all if the tests were just an ordinary examination without much consequence for the individual student. For example, the programs that move student exam performance from a passing level to a more advanced level are not studied, as they do not affect the stake of high school graduation. This report covers programs with identifiable public school costs that depend on the existence of high stakes testing.

Though we use the shorthand phrase “costs of exit exams,” few of these resources are actually used for testing. Instead, the existence of the exit exams leads to a shift in the use of resources—for example, the increase in remedial education. The cost signals a shift in education and educational programming. To endorse exit exams is to endorse tacitly this shift and its accompanying cost; to oppose exit exams is to oppose the shift and its accompanying cost.

Purpose

The purpose of this study is to inform policy and programmatic decisions about high stakes testing by providing a realistic estimate of the resources involved. While the broad decision of whether to have an exit exam may not depend on cost considerations, the planning for implementation and the choice of types of exam systems can be improved by accurate cost data.

Cost data about high stakes testing is hard to arrive at because most costs do not come in the form of a line item in a state budget. Instead, resources are committed primarily as a result of a myriad of decisions at the district and school level, decisions that are nowhere categorized or summarized as related to exit exams. Correspondingly, state decision-makers lack basic cost information, and even district leaders may not have information about what to expect or how to estimate their costs. We hope this report fills these information gaps.

The reason for covering more than one state is to provide comparison and insight into the reasons for costs. Within a single state, it is difficult to disentangle the disparate impacts of historical events, such as when Massachusetts adopted simultaneously a comprehensive

assessment system and an exit exam with specific characteristics. States that resemble each other in one respect but not another allow some inference about the role of the characteristic that varies, such as the impact on costs of emphasizing prevention.

Sponsorship

Funding for this study has been provided to the Center on Education Policy by a number of foundations. Neither of the state governments concerned, Massachusetts nor Minnesota, provided any funding or any substantial assistance in non-monetary form. This is an independent, nonpartisan report, unconnected to any policy proposal, ideology, or court case. The findings and expert opinions are those of the authors.

Organization of Report

Chapter 1 provides background on the genesis of this project, reviews a prior study's method and results, and then introduces Minnesota and Massachusetts as the subject states. We cover why they were chosen and summarize what their exit exam experience has been. The chapter then presents our expected findings.

In Chapter 2, the methodology is discussed, from informed task groups to interviews. Several changes from the procedures used in Indiana are highlighted. Then we discuss the groups and topics used in the states in the fall of 2003. The tasks facing each group are also discussed.

The procedures and results from Minnesota are the subject of Chapter 3. We cover the exams themselves, then the various components of costs—types of programs, level of government, and type of resource. Costs for exams, both as they are and as they might be, are discussed, as are costs for special education students and for a hypothetical example of a prototypical average state exit exam. We compare Minnesota's costs to those in Indiana and explain the differences.

A similar analysis for Massachusetts is the point of Chapter 4, with comparisons extended to cover Minnesota. Comparisons are extended to cover all three states and to programs for limited English proficiency students.

Chapter 5 explicitly focuses on patterns across the states. Possible sources of similarities and differences, such as pricing and supportive structures, are reviewed. The special education results are considered from a cross-state perspective, to see whether the Minnesota results can be generalized to other states. A similar analysis is conducted for limited English program exam-related costs. In light of the cross-state comparisons and patterns, costs are recalculated to take account of additional factors. The impact of exam characteristics on costs in all states is estimated.

Chapter 6 provides a summary of our expectations and their fit to results, as well as directions for future research, including a possible focus on regional education culture, suggested changes in methods, and possible new substantive foci. A few concluding remarks end the report.

CHAPTER 1

Background and Expected Findings

This study evolved from two broader streams of work, the Center on Education Policy's (CEP) coverage of high school exit exams and Augenblick, Palaich, and Associates' (APA) analyses of state school finance employing professional judgment models. The two streams converged in 2002–2003 when CEP commissioned APA to develop a method to estimate the costs of high school exit exams and to test the method in one state. This report covers a second stage, where the method is extended to additional states and results are compared among states.

CEP, APA

Based in Washington, D.C., and founded in 1995 by Jack Jennings, the Center on Education Policy is a national, independent advocate for public education and more effective public schools. CEP is nonprofit and nonpartisan and has been conducting studies of high school exit exams and the No Child Left Behind Act with funding from a number of major foundations. As part of its second, 2003 volume on exit exams, CEP commissioned a study by APA on exit exam costs in Indiana that was included in the report.² From the start, CEP and APA have worked to develop a method that could estimate both the current costs related to the exams and the costs related to providing students with a substantial chance of passing a test that measured what they should have learned.

Augenblick, Palaich and Associates provide technical assistance to policymakers on education finance and governance issues. APA has designed the school funding formulae for many states, has estimated the cost of an adequate education for even more states, and is involved in several locales in designing and costing out compensation systems. In the 1990s, APA (then A&M) conducted several studies for the Department of Education and the Department of Children, Families and Learning in Minnesota on the impact of the state's testing and graduation standards.³ The cost of the high school exit examination that was about to be implemented was part of that work.

Method Used in Indiana

For the earlier study, APA proposed and CEP agreed to the following tasks for APA.

-
- 2 Ch. 4, "How much does it cost to implement exit exams?," in Keith Gayler, Naomi Chudowsky, Nancy Kober, and Madlene Hamilton, *State High School Exit Exams: Put to the Test*, Center on Education Policy, 2003, 49-58, <http://www.ctredpol.org/highschool/1/exitexam4.pdf>
 - 3 Augenblick & Myers, "Description of the work done in Minnesota to estimate costs associated with new high school graduation standards," November, 1997.

1. Organize the framework in terms of the costs of programmatic choices faced by decision-makers in a state. Whenever possible, cost estimates would be specific to the program characteristics and would be separated into phase-in and recurring yearly costs.
2. Develop a methodology appropriate to the states that have exit exams now and those that might have one in the near future. To identify the significant features and variations of state arrangements, APA would draw on CEP's 2002 report, plus interviews conducted in the winter of 2002-2003. The focus would be on the state variation in policies on accountability, remediation, and retakes; test examination formats, alternatives, pass rates; and extent of implementation, grades tested, and coordination with other tests and the curriculum.
3. Apply the methodology to one state in order to test and illustrate the approach. Drawing upon the existing financial data and conducting interviews with those charged with administering aspects of the exit exams, APA would describe the cost of providing students an adequate opportunity to pass a high stakes exam. The case study state selected should represent typical state exit examination arrangements, provide adequate data to make estimates, have fully experienced the exit exam process, and offer feasible arrangements for adequate interviewing within the budget.
4. Report on the methodology and its application to one state within the programmatic choice framework, explaining the procedures and the reasons for them, and using the case study as an illustration.

The primary method developed is an adaptation of the professional judgment model used in school finance adequacy studies. The task groups used here for estimating exit exam costs are usually smaller, meet for less time, are focused on a narrower task, and are organized around different topics than are professional judgment panels in adequacy studies, but the two share many strengths and weaknesses as techniques. The questions that they address, however, are quite different. Adequacy studies typically focus on the entire cost of an education, often to meet current state law, while the present study focuses on the marginal cost of the programs for exit exams, including costs of changing from present standards.

The primary conclusions in the study of Indiana were that the costs of implementing exit exams were substantial, that the costs of meeting performance goals for those exams would be even more substantial, that the costs were overwhelmingly borne by districts from funds not designated for use to implement high stakes testing, and that the costs were for a mix of programs ranging from testing itself, to remediation, to preventing initial failure, and to professional development of school personnel. The costs were met by shifting resources from other priorities.

Extension

This study extends the examination of costs to two additional states, making available results for different types of examinations. Each state is to some degree unique and is presumed to require an independent investigation of its costs. In addition to extending the count of state studies, the aim is to change the quality of available information, first by covering new types of exams and states and second by permitting comparisons across states of costs and their correlates.

Minnesota and Massachusetts

The decision to extend the study to the specific states of Minnesota and Massachusetts reflects the importance of diversity. Minnesota administers the Basic Skills Tests—8th grade math and language tests as well as a 10th grade writing test. The 8th grade tests are believed to be less demanding than the 10th grade examinations used by most states. We were interested in seeing whether a more basic examination had different costs than Indiana’s 10th grade Graduation Qualifying Examination.

Unlike Indiana and Minnesota, Massachusetts graduated its first class in 2003 under the high stakes testing requirements as part of the Massachusetts Comprehensive Assessment System. Does the cost structure of a newly implemented exit exam system such as MCAS differ from those of longer-established systems such as the GQE and BST? Unlike Indiana and Massachusetts, Minnesota does not link its exit exam to the standards-based curriculum—does this decoupling of reforms change the amount of costs or types of programs that follow from the exit exam?

Massachusetts has a high initial pass rate and a high graduation qualifying rate, by passing the tests, for high school seniors, while Indiana is much lower on both; Minnesota has a low initial pass rate but a high graduation qualifying rate. By implication, Massachusetts invests in the education before the exit exam is first administered; Minnesota concentrates on providing remediation to students in difficulty, as does Indiana, which has a shorter time period for remediation. Are these implied characteristics matched by actual programmatic expenditures?

Expected Findings

Are the costs similar to those in Indiana? Do different types of exams indeed generate different costs, and does the method uncover these differences? These questions both look backwards to the Indiana study, wondering whether to confirm its results, and look forward to the ability to compare states and analyze general patterns.

We undertook this study with expectations both procedural and substantive. We expected to be able to conduct the studies in a manner similar to that used in the Indiana study, with similar cooperation from state and local personnel.⁴

Substantively, we expected results similar to those in Indiana, with the sizable expenses at the district level, though we expected that Minnesota’s Basic Skills Test would require fewer resources than did Indiana’s more rigorous GQE. Massachusetts’ rigorous MCAS exam was expected to be the most costly of the three exams. We expected high resources to be involved with exam-related programs for both special education students and limited English proficiency students.

For the analysis of costs, we expected that a hypothetical, more difficult standard for passage would involve higher costs. We additionally expected that similar exams would have similar costs in all states. We expected that cost differences among the states for similar types of exams would trace primarily to the general cost of education—salaries, pupil-teacher ratios and the like; additionally, the exam-related costs would be related to how well the

4 For a discussion of methods used in Indiana, see Part I of this volume.

educational programs offered by the state fit together with the exit exam. Finally, our expectations about changed standards for an exam focused on differences in the exam structure.

One programmatic expectation was that the costs of a more basic exam would be not only less but also different from the costs of a more demanding exam. While testing itself should be as expensive and remediation might be as extensive, there would be less need for intensive remediation and for dramatic intervention to prevent failures, and less demand for professional development to meet changed requirements.

A second programmatic expectation was that states with recently implemented exams would focus their efforts more on testing and remediation, while states with more established systems would focus more on professional development and prevention, especially if a move toward a more demanding examination was planned or implemented.

The higher the initial failure rate, the more the education that must occur after the exam (remediation) rather than beforehand (prevention). The higher the final graduation qualification rate, the more the education that actually has occurred (and the more the expense that has been incurred). A large gap—high initial failure followed by high qualification—suggests a large amount of learning focused on remedying the deficiencies demonstrated by the initial exit exam test.

The concluding chapter summarizes these expectations and compares them to the results.

CHAPTER 2

Method

Our methods begin with an understanding of earlier findings and concepts from studies of related topics. Regarding exit exams, the most useful sources have been the CEP publications.⁵ For Minnesota and Massachusetts, the single most helpful sources have been the state department of education websites. The National Center on Educational Statistics provides comparative statistics.⁶ APA's earlier projection of the likely costs of Minnesota's BST provided helpful background.⁷ This secondary research provided the context for the original research reported here.

APA used a combination of interviews and group meetings to study the cost impact of exit exams in both Minnesota and Massachusetts. The interviews, usually held with statewide education professionals, were conducted early in the week of fieldwork visits and provided a framework of different perspectives from departments of education to teachers' unions. In the days following the interviews, APA convened groups of education professionals in various locations around each state. The groups included teachers, principals, superintendents, special education professionals, and finance directors. The process in both states involved panels early in the week discussing the existing exit exam program, mid-week groups estimating the cost of alternative exit exam arrangements, and a panel at the end of the week to review the work of previous groups and approve the price estimates upon which APA would conduct its analysis.

The main focus of this study is upon perceptions of the effects of exit exams on costs and resource use. However, there are four common polarities that we have not found useful in this study:

- Resource v. program
- Cost v. benefit
- Cause v. effect
- Perception v. reality.

The use of resources measured here is for programs. The general pattern that we have found is that the resources to be used for one program are already being used for another program and would have to be shifted. Resources might be shifted from science programs conducted in Spanish to programs to acquire English language proficiency in order to pass

5 Naomi Chudowsky, Nancy Kober, Keith Gayler, and Madlene Hamilton, *State High School Exit Exams: A Baseline Report*, Center on Education Policy, 2002, and Keith Gayler, Naomi Chudowsky, Nancy Kober, and Madlene Hamilton, *State High School Exit Exams: Put to the Test*, Center on Education Policy, 2003, <http://www.ctredpol.org/highschoolsexit/1/exitexam4.pdf>.

6 <http://nces.ed.gov/>

7 Augenblick & Myers, "Description of the work done in Minnesota to estimate costs associated with new high school graduation standards," November, 1997.

the exit exam. Even if resources are newly shifted into education, they are shifted from some other public or private use. For the most part, the resources involved in carrying out exit exams would, if it were not for exit exams, be used for other education programs.

The costs measured here have accompanying benefits. When the resources are shifted, the benefits associated with the former programs decline, and the benefits associated with the new programs increase. It is a mistake to simply ignore the benefits of one or the other set of programs. Teachers may bemoan the loss of education benefits to top students when too much time is spent on exit exam preparation, while parents may be puzzled by why their schooled children cannot read and calculate well enough to pass an exit exam.

Exit exams do not produce their most important effects by passing or failing students. While failure on exit exams does indeed generate remedial education to pass the exam before the student's class graduates, the main consequences of exit exams occur because of anticipation. In anticipation of the possible effect of exit exam failure, students, parents, teachers, and administrators are willing to change their behavior and programs. In the end, Massachusetts' exit exam failed few students in its first year as a graduation requirement, but the anticipation of possible widespread failure produced a major educational effort by all concerned. The anticipation of an effect has an effect.

While there are differences between perception and reality, this is another way of saying that there are differences among perceptions. The perceptions of informed people with good judgment are what people take as reality. These perceptions, even though they may later be reversed, are the basis for anticipations. The absence of any clear perception of informed people with good judgment—which has been the case up to 2003 for the district resources used for state exit exams—corresponds to a virtual absence of reality. Costs that are not measured seem not to exist.

While we focus on resources, costs, effects, and perceptions regarding exit exams, we attempt to keep in mind the linked programs, benefits, anticipations, and informed judgments.

General Information Sources

State and National Data Sources

For statewide exit exam programs—assessment and remediation, primarily—appropriations are authorized by program with a statement of purpose. These data are commonly available in state publications, notably a series of budget authorizations. Additional important facts, from statewide enrollment to average teacher salaries, can be gathered from published sources. In general, when the information will be used by the state itself, then the state makes the information publicly available.

Additional information we required is not routinely needed by the state but may be of occasional interest or of routine interest to some interest groups so we went to them for these data. Average salaries of superintendents, for example, are of interest to associations of superintendents. Information is collected for analysis purposes by the federal government and its agencies, though often with a time lag that makes the estimates imprecise as a guide to current behavior.

Interviews

Reliable recorded data may not be available for some portions of exit exam costs, such as the exit exam share of assessment costs. In these situations, we turn to interviews with state-level personnel, from the department of education and legislative officials to associations and individuals interested in education. We request interviews with a variety of individuals: the head of a teacher's union, a former state superintendent of education, the executive director of a superintendent's association, the legislative lobbyist for local governments, the head of an organization for educating special education teachers, the economist providing cost estimates for proposed legislation, and the director of an association of rural school districts are examples. Further, individuals can often provide an overview of what is important and of evolving trends that is helpful in interpreting published material. Early in the week of a visit to a state, we interview informed individuals to get specific judgments and to get a broad understanding that helps to guide our research during the week.

Interviews were satisfactorily conducted in Indiana and in Minnesota. In Massachusetts, however, the range of interviews was restricted, as we could interview no employees of the department of education or of the executive and legislative policymaking bodies. The problem was two court cases in which the department's legal advisors felt that any testimony to APA—or any APA finding—might adversely affect the outcome of the court case; lawyers on the other side felt equally threatened by findings potentially adverse to their case. As the General Counsel to the Massachusetts Department of Education put it:

Were it not for the pending litigation, the Commissioner and others here would be pleased to talk with Augenblick, Palaich and Associates about the issues you are studying. Under these circumstances, though, we have to defer our response to your questions until the litigation is fully resolved.⁸

The interviews would have provided us with an opportunity to check our information about state-level programs with people officially responsible for keeping informed about these issues.

Problems with Local Recorded Data

The absence of reliable recorded data becomes an acute problem at the district and school level. There is not the formality, the detail, or the record keeping that exists statewide. School boards, superintendents, and even principals have some discretion in spending, and records are often inadequate to reflect how the discretion is used regarding exit exams. Further, a large share of district funding comes in the form of state per capita payments, for which scant accounting is required and which permits discretion in spending. Even when state programs designed to aid local exam-related programs are clearly recorded, the actual usage of the funds at the district level is rarely as clear. As a result, records are an inadequate guide to district-level funding of exit exams.

A second problem concerns creating a state composite from district data. Not only are the districts' records often inadequate, the extent of record keeping varies, often in proportion to the resources available to a district. If data are missing from small, rural districts, should they be assumed to be the same as the data from large, urban districts? If not, how is a composite to be constructed?

8 E-mail in the possession of the author, the intended recipient.

No reliable recorded data exist for events that haven't happened and may not happen. Neither states nor districts have records of spending on a "what if" basis, so records are an inadequate guide to costs in hypothetical situations. Our method of addressing these problems is to use the professional judgment approach.

The Professional Judgment Approach

The professional judgment approach relies upon the combined judgment of educators to translate their experience into a shared specification of the resources needed for a task. The particular instructions and tasks will be described when the results are presented for each state, but in this section we review the basic panel composition and process.

Panels

The panels comprise school and district level personnel familiar with the district's programs and resources. Each task group includes a mix of backgrounds, including a variety of current positions germane to estimating resources and a variety of school districts. The aims are to represent a range of experiences and to avoid the extraneous interactions invoked by having panel members who are similar or significant to each other outside the task group. As with a focus group, the panel members may have some defining characteristic in common, such as region or a special focus on the panel topic.

In selecting panelists to invite, we started with in-state advisors to whom we explained our aims, procedures, and criteria. Bob Wedl, former Commissioner of Children, Families, and Learning in Minnesota, and Paul Reville and Jennifer Candon, of the Center for Education Research & Policy at MassINC in Massachusetts, then led efforts to identify the most appropriate district panel members and interviewees. APA then contacted the named individuals, briefly explained the project, and invited participation for a specific date and time, location, and topic. If an insufficient number of acceptances were received for a panel, then additional invitations to fill underrepresented roles were offered. Invitations to be interviewed were offered to a representative selection of individuals well informed about past, current, and prospective exit exam policies. For review panels, the selection criteria differed somewhat, and no state employees and only one district school employee served on the review panels.

Table 2-1 indicates the current positions of the Minnesota panel members. Superintendents and the assistant superintendents were one-third of the panelists. Often the assistant superintendents had special responsibilities in the areas of the next most common position, curriculum, assessment, and research. School-level personnel such as principals and teachers were about one-sixth of the total, with directors of special education and the district's financial director accounting for most of the remaining positions. In addition to their current positions, many members have occupied other jobs or a similar job in a smaller, often distant district. Some had even begun their careers in another state, such as North Dakota.

Table 2-1: Minnesota Panel Member Positions

Superintendent & Assistant Superintendent	15	33%
Curriculum/Assessment/Research	10	22%
Principals & Teachers	8	17%
Directors of Special Education	5	11%
Finance/Business	4	9%
Other	4	9%
TOTAL	46	101%

Note: Table does not total to 100% due to rounding.

The current positions of panel members in Massachusetts broadly resembled those in Minnesota, with more superintendents. In Massachusetts, we had few panelists with a background in finance or special education, and more with a position related to LEP services and in coordinating school or district MCAS responsibilities. In general, the compositions functioned similarly in both states and resembled those in Indiana, with somewhat fewer principals and business managers.

Table 2-2: Massachusetts Panel Member Positions

Superintendent & Assistant Superintendent	18	50%
Curriculum/Assessment/Research	7	19%
Principals & Teachers	4	11%
English as a Second Language / LEP	2	6%
MCAS	2	6%
Other	3	8%
TOTAL	36	100%

Participation on a panel does not necessarily imply an endorsement of the results. Panel members may differ on priorities or expectations. Some may have a good experience with one software package while others prefer a second. On the LEP panel in Massachusetts, for example, one panelist had good experience with parental consent to after school tutoring as one of a menu of tutoring times, while another member had had low participation in an after school tutoring program due to conflicts with students' after school jobs.

At locations outside of Boston and Minneapolis, panels were held in the administrative building of a large city (Rochester, Duluth, St. Cloud, Plymouth, New Bedford, Springfield, Worcester, and Lawrence school districts) in each of the eight regions. One exception to the diversity pattern is that the host district brought three or more participants to four panels. In contrast, visiting districts never brought more than the invited two members. In Boston and Minneapolis, most sessions were held at a headquarters of a school-related group, and an imbalance of attendance was not a problem.

Evaluation

To investigate some of the potential difficulties with the process, we asked members of the district panels to fill out an open-ended questionnaire at the end of the session. The question texts were:

- Q. 1** Do you feel that the instructions and materials presented to you were clear? Did you have an understanding of the tasks that you were being asked to complete?
- Q. 2** During this meeting have you been able to openly express your ideas and concerns? Do you feel these have been used in the final conclusions of the group?
- Q. 3** Do you feel that we have addressed all of the resource needs involved in implementing programs that would yield the specified outcomes?

The overall pattern is that participants felt able to express their ideas and have them make a difference, but felt that the task itself was difficult and required time to understand. Opinion was mixed, especially in Massachusetts, on whether all the resource needs had been identified, less due to some deficiency in the group or moderator than to the inherent difficulty of specifying all the resources.

Table 2-3: Process Evaluation

	MINNESOTA			MASSACHUSETTS		
	YES	MIXED	NO	YES	MIXED	NO
Q. 1	71%	26%	3%	69%	24%	7%
Q. 2	100%	0%	0%	83%	14%	3%
Q. 3	40%	49%	11%	48%	24%	28%

Though comments tend to be individualized, we provide some examples to illustrate the content of the responses in each category for each question. The shortest answers tended to be in the “yes” category.

Table 2-4: Response Samples

	YES	MIXED	NO
Q. 1	The materials were excellent. The instructions were very clear.	I had some idea of what was expected, but was unclear about some of the expectations.	I don't think we completed the tasks that we [were] asked or invited to complete.
Q. 2	Most inclusive process. Allowed for open debate.	It was ok but some were more dominant than others — districts were not evenly represented.	Chair needs to be stronger. When one person does all the talking & gives the impression of disdain, it can dampen discussion.

Q. 3	All resource needs were fully covered.	We addressed the ones we could think of. I am sure there are more.	No, probably not — as we kept delving, clearly we could identify more & more resources that we either now use, or could use to provide better outcomes — but at least it was a beginning.
------	----------------------------------------	--------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Review

One development of the professional judgment approach by John Augenblick and John Myers has been the use of a review panel at the end of the week to review and reconcile the recommendations of the earlier district panels. Any two panels treating the same topic come to at least slightly different results, and the review panel can reconcile differences by simple averaging, by choosing the best supported alternative, by creating a composite program, or by adding programs where neither panel has addressed a need. Review panels do not simply substitute their judgment for that of the district panels; rather they do check to see that the results fit within the task. Additionally, the review panel checks on the APA cost estimates for each resource.

Adaptations

APA changed both the tasks and the process somewhat from the one used in Indiana. We reduced the maximum number of panels on a topic from three to two, based on the lack of distinction in Indiana between the results of having two and three panels on a topic. An additional panel has been added in each state that focuses on an important special topic that tended to get overlooked by the district panels—special education and limited English proficiency. We added an additional day so that no moderator would attempt to conduct several panels in a day, especially at widespread locations, which caused logistic difficulties in Indiana. The duration of each panel was set at three hours to ensure enough consideration.

These adaptations have costs. The use of two panels risks difficulties in reconciling results when they significantly differ, as was the case in Massachusetts with the panels considering a changed exam standard. The review panel felt that the estimates were too far apart to easily reconcile. APA felt more comfortable combining estimates, as we had been at the panels and understood more of the processes and reasons behind the differences in estimates. Even so, we limited ourselves to a simple averaging of the two estimates, rather than a more substantive adjustment. Three panels provide an easier method of reconciliation. Perhaps surprisingly, the extra time did not provide any slack. While panels were often ready to stop after three hours, none stopped early.

The special topic results have been added to the district estimates, at least to the degree that district panels do not overtly account by programs for special needs groups, but discussions actually covered much of the same ground. Indeed, it was somewhat frustrating to the moderator not to be able to use the estimates of the LEP panel in the broader estimate of district costs, as some of the non-LEP members of the panel provided excellent material on the programs being used, much more illuminating material than was provided by at least one of the panels solely focusing on costs of the current system.

Panel Topics

The basic panel topic was what the costs of the current exit examination are. Two panels in each state addressed this topic, called the “What Is” topic. To help panelists recall the range of programs that might be involved with the current exam, we handed out a list of programs used in one or more states, organized into four broad categories. The program categories were testing, remediation, prevention, and professional development. We stressed to panelists the need to consider this as a marginal cost of the exam being a graduation requirement. Hence, costs of the exam simply as an exam were not considered. Instead, testing programs included, for example, the cost of retesting, which tends to be much higher with high stakes examinations. Remedial program participation during summer school, for instance, is often much higher once exit exams are instituted. While preventing failure on exams is a focus of a great deal of educational activity, it becomes a consequence of exit exams when, for one, students who have failed the eighth grade exam, and only that exam, are singled out for special exam-targeted programs. Similarly, when professional development is targeted to helping students with the high stakes test and only that test, it becomes part of the estimated cost of the current examination policy.

A second set of two panels addressed a “What If” topic. If the exam arrangements were changed, what would be the cost of the change? What new or expanded programs would be expected to follow from the policy change? In Minnesota, the change was to make the current 10th grade math and language arts examination the high school exit exam, rather than continuing to use the Basic Skills Tests. In Massachusetts, the change was to raise the standard for graduation from a “needs improvement” score to a “proficient” score. For the four broad categories of programs (testing, remediation, prevention, professional development), panels estimated how many resources would be needed for changed programs.

A fifth panel focused on a hypothetical exam that we created from characteristics averaged across states that conduct high school exit exams. We created this exam in order to be able to compare costs across states. This exam represented a cross-state average (mode) in the exam structure, in the district characteristics (mean), and in student performance on the exam (mean). The focus was again on the marginal cost of such an exam being an exit exam. Panel members were told to assume that unspecified aspects reflected their current state averages. We took a simple average (mean) of the results available for states that had implemented exit exams. The information presented to panelists was as follows:

The examination:

- Is first taken by the 10th grade students
- Covers language arts and mathematics through the 10th grade level
- Is required for high school graduation
- May be retaken multiple times
- Includes multiple choice, writing prompt, and short answer sections
- Has accommodations for limited English proficiency students
- Can be specifically adjusted for students with particular learning disabilities

- Remediation for students who fail is required and partially funded by the state
- Schools and districts are held accountable for results.

The exit exam objectives accomplished by the hypothetical schools and district under the 2003 standards and procedures are:

■ Initial (10th grade) district pass rate on Reading	76%
■ On Mathematics	67%
■ 12th grade students passing all tests by spring	91%
■ Initial mathematics pass rates for students with disabilities	30%
■ Initial mathematics pass rates for free/reduced price lunch students	49%
■ Initial mathematics pass rates for black students	48%
■ Initial math pass rates for Hispanic and Native American students	58%
■ Initial mathematics pass rates for ESL/LEP students.	40%

The characteristics of the hypothetical school district are:

- 5830 students in 12 schools (300 teachers)
- 7 elementary schools (K-5) with 400 students and 22 teachers each
- 2 middle schools (6-8) with 660 students and 35 teachers each
- 2 high schools (9-12) with 825 students and 33 teachers each
- 1 alternative school with 60 students and 10 teachers
- 36% of students (about 2100) are black, Hispanic, or Native American
- 13% of students (about 760) are in special education
- 26% of students (about 1520) are eligible for a free or reduced price lunch
- 6% of students (about 350) are in programs for limited English proficiency.

The final district panel examined a special topic and considered the marginal cost to special education or LEP students of having the current exit exam requirement.

While the task of the review panel was to reconcile findings and match them to the tasks, neither review panel considered reconciling results across the different types of panels.

CHAPTER 3

Minnesota

Minnesota is a mid-sized northern tier state with average numbers of students at risk—usually in rural or central city areas—and LEP students. It has a smaller share of African-American, Asian, and Hispanic students than does the nation as a whole. Traditionally, it is viewed as providing an excellent education at a moderate cost.

Table 2-5: Minnesota Public School Data

	2002
STUDENTS	850,236
Minority	18%
Free/reduced lunch	27%
LEP	5%
Graduation rate	91%
EXPENSE/PUPIL	\$7,664

Salaries and benefits in Minnesota are not far from the national average. Minnesota attracts teachers from nearby, lower-paying North and South Dakota, but in turn loses teachers to better-paying states to the east, Wisconsin and Illinois.

Table 2-6: Estimated Average Public School Salary and Benefits in Minnesota, 2003

CATEGORY	AVERAGE
Teachers	\$44,850
Librarians/Media Specialists	\$48,750
Technology Specialists	\$53,050
Guidance Counselors, Curriculum Specialists	\$57,450
Instructional Aides	\$19,200
Clerical/Data Entry	\$23,450
Principal	\$71,450
Assistant Principal	\$57,150
Superintendent	\$90,500

Assistant Superintendent	\$78,450
Supervisor/Coordinator/Director/Manager	\$71,750
Substitute Teacher (daily rate)	\$120
Benefits Ratio	24.9%

Basic Skills Tests

As part of the Basic Skills Tests, Minnesota administers mathematics and language arts exams in the 8th grade, and then administers a writing test in the 10th grade. Students who fail the exams may retake them with the first-time test-takers in the spring. Retake opportunities are also offered in the summer for all students and in April for seniors. As the name implies, the exams were created to ensure that no student graduated from high school lacking the basic skills in reading, mathematics and writing that are needed in adult life.

Minnesota was one of the first states to institute a state-created exit exam in the wave of 1990s standards-based reforms. Because the exam was created before full reforms were in place, it is not meshed with a state-sponsored standards-based curriculum. While the initial pass rate is relatively low, the yearly opportunities for retakes and maturation across four years yield a final rate of passage above 99% among high school seniors.

The low grade level of the exam—8th grade—compared to other state exams makes it less difficult for high school seniors to pass than most exit exams. However, the initial failure rate among 8th graders suggests that the exam is not easy for them. Further, Minnesota had the highest 8th grade NAEP scores on mathematics, indicating that students in other states might have even more difficulty with the exam than do Minnesota students.⁹ The maturation of students can turn a difficult 8th grade test into a much less demanding high school exit examination for seniors.

Contemplated Changes

The Basic Skills Tests have remained largely unchanged, and no major changes are contemplated, basically because reforms have passed them by. The focus now is on school and district accountability rather than individual requirements. No Child Left Behind has created a buzz, despite Minnesota's relative lack of difficulty on early indicators such as the percentage of schools failing to make adequate yearly progress. NCLB's requirements that come into place in the next few years are the principal focus of current concern.

Costs of Current Exam

District Level Costs: What Is

To estimate the school and district costs of the BST, APA created a hypothetical district that resembled one the average student in the state attends. The hypothetical district has 6,150 stu-

⁹ The NAEP data may be found at <http://nces.ed.gov/nationsreportcard/mathematics/results2003/stateavgscale-g8.asp>

dents in 11 schools, including two high schools (9–12) with 900 students each and an alternative school with 150 students. Of the hypothetical total, 13% of students are in special education, 26% of students are eligible for a free lunch, and 5% of students are in limited English proficiency programs.

In this hypothetical district, the BST tests, test-takers, and results closely followed current statewide figures:

■ Initial (8th grade) district pass rate on Reading	81%
■ On Mathematics	72%
■ Initial (10th grade) district pass rate on Writing	91%
■ 12th grade students passing all tests by spring	99%
■ Special education share of test-takers (68% of them with some type of accommodation)	10%
■ ESL/LEP student share of test-takers (63% of them with some type of accommodation)	4%
■ Initial mathematics pass rates for students with disabilities	30%
■ Initial mathematics pass rates for free/reduced price lunch students	49%
■ Initial mathematics pass rates for black students	33%
■ Initial math pass rates for Hispanic and Native American students	43%
■ Initial mathematics pass rates for ESL/LEP students	34%

Panels of superintendents, principals, specialists, and teachers were asked to specify the programs and resources needed, because the exam was a graduation requirement rather than simply another student assessment, for this average district to produce average results described above. One analogy offered for their task was to advise the superintendent of this hypothetical district in a state that was similar to Minnesota but had just adopted an exit exam what programs and resources he or she should use to achieve the above results. In sum, panelists were provided a state microcosm at the district level, then asked what programs, with which resources, would be required to produce the BST results, given the district characteristics. On a per student basis, these estimates apply statewide as the local portion of costs.

Resources and District Cost

Both panels examining the resources currently needed for local BST programs provided similar overall estimates, and the review panel recommended using an average of the results. **Table 2-7** presents the overall per pupil local cost, along with the types of resources used. Costs are split relatively evenly (48%–52%) between the school and the district level. The costs are primarily (78%) for personnel, especially at the school level (94%). Teachers account for 43% of the local cost and for 90% of the school level costs, with some cost for specialists at both the school and district level. The district’s largest personnel cost, however, is for cross-school personnel, such as for truancy prevention. The most sizable non-personnel

costs are for technology (12% of total cost), notably software acquired by the district, and for such transportation items as busing for summer school.

Table 2-7: Per Pupil Local BST Costs

	SCHOOL	DISTRICT	TOTAL
Personnel	\$ 58	\$ 43	\$ 100
Other	\$ 4	\$ 24	\$ 27
TOTAL	\$ 61	\$ 66	\$ 128

An overall cost of \$128 per student per year—\$1,600 across a student’s career in grades K–12—is not cheap. Across 850,000 students, that totals over \$109 million per year, nearly 1.7% of Minnesota’s total public school expenditures on elementary and secondary education.

Types of Programs

Activities have been grouped into four broad categories of programs.

Testing emphasizes the actual testing of students. Supportive activities include test development, administration, record keeping, creating alternative exams, providing accommodations for disabilities, and arranging for retests.

Remediation emphasizes efforts moving students from an initial failure on the exit examination to passage and graduation. Supportive activities include summer school, remedial classes, tutoring, pre- and after-school programs, emphasis on test-taking skills, creating a remedial curriculum focused on core skills, using remedial software, and alternative routing to demonstration of skills.

Prevention emphasizes changes in instruction so that students pass the examination on their initial attempt. Supportive activities include, targeting programs at students who have failed exams in earlier grades, cumulative record keeping on individual student skills and test diagnoses, emphasizing early reading, adopting teaching technology aimed at problem students, and developing techniques to present standards-based material to special education, limited English, and at-risk students.

Professional development emphasizes staff training to make changes that lead to higher rates of passage on the examination. Supportive activities include training for test administration, learning to translate test evaluation into teaching, and training regular teachers to address the needs of special education students.

The **Other** category primarily includes cases where a panel provided a direct cost estimate of a combination of programs—the purchase price of software to be used for both preventive and remedial programs, for instance.

Table 2-8 reports on the expenses classified by the types of activities described by the panels. These programs are highly targeted, both to students at the margins of success on the test and to the specific problems that impede success. The considerable majority of resources

is devoted to preventing an initial failure, often by targeting students who have had difficulties on earlier exams for summer school, extended day and half-size classes. Due to the high stakes test, programs to increase attendance are beefed up to get problem students into classes. Remediation programs constitute the second largest category of expense. Those students who have failed are encouraged to attend summer school, are provided special math classes focusing on basic material, and are provided highly qualified teachers who would be otherwise occupied if the BST were not a graduation qualifying exam. English language learners are provided writing skills aid to help with the problems of first language interference, in which the grammar of the hearth tongue is used in place of English grammar.

Relatively little is spent on testing itself or on professional development. Testing expenses are deemed not to depend on the graduation requirement, except for the retakes, which increase fees and occupy personnel but are not different from those of other exams. Professional development is largely limited to targeted programs aimed at technical aspects of using test results. Perhaps because the test is not tied into curriculum, professional development is given low priority among programs.

Table 2-8: Local BST Programs

	SHARE OF COSTS
Remediation	29%
Prevention	63%
Professional Development	2%
Testing	7%
Other	0%
TOTAL	101%

Note: Table does not sum to 100% due to rounding.

It is unclear if the lack of alignment to state standards that is characteristic of Minnesota’s BST has an impact on the level of spending. It does, however, appear to influence the pattern of spending, as wholesale reform would be a case of a rather outdated exam causing major changes in a more modern curriculum. One reason for the specific, somewhat ad hoc responses to the exam may be the lack of fit to the curriculum.

The longevity and level of Minnesota’s BST modify our understanding of these results. The tests have been in place a relatively long time. They are part of the operating system, which has already adjusted to them. As a consequence, the causal role of the BST is murky to Minnesota’s current practitioners. In Minnesota, panelists struggled with the notion of costs contingent on an exit exam more than panelists elsewhere, in part because the exams are embedded in current practice. Unlike those with more recent experience with change, they seemed to have difficulty spotting programs that existed because of the high stakes testing and tended to limit themselves to exam-focused programs and resources.

Additionally, an exam given in the 8th grade over 8th (or lower) grade material generates a different type of concern. Relatively few students who regularly attend regular class

in a district for eight years fail an exam, and if those students regularly attend for the next four years, they are nearly certain to graduate. Little out of the ordinary is required for these students. Instead, the focus is on “problem” students—migrants to the district who have never attended a school, special education students, at-risk students, mobile students and those who do not attend school, students with inattentive parents, students just learning English. Exit exam or no exit exam, the school district should—and often would—provide special programs for these students that would help bring their skills in language and mathematics up to an 8th grade level. In contrast to the BST, No Child Left Behind standards have generated a considerable flurry of excitement and concern about performance, despite Minnesota’s comparatively good showing on the latest results. In part, the difference may trace to the higher standards that students are being measured against.

Special Education Costs

In the Indiana panels, we felt that special education resources were underrepresented. Special education students represented a high proportion of the students with initial and ultimate lack of success on the exit exams and, along with students with limited English, were distinguished by strong barriers to success on the exam even for motivated and conscientious students. However, panelists represented a broad cross-section of practitioners, and the special education professionals tended to get marginalized. Special education was almost never targeted for special programs or programmatic change, though special education students were included in programs aimed at those having difficulty with earlier exams.

As a result, APA established a separate panel for special education in Minnesota, bringing together from across the state those with special education experience. We provided the panel with the same hypothetical average district and same task, except they were to concentrate on resources for special education that were contingent on the graduation requirement of the BST. As none of the regular panels in Minnesota mentioned programs for special education, the estimates provided by the special education panel can simply be added to the costs for other students to arrive at a total figure.

To our surprise, the special education panel identified only \$26 per special education student as costs traceable to the BST exit exam. They believed that the education priorities for special education students were largely unchanged by the exam, with some students focused on pre-cognitive life skills, while others were coping with emotional traumas (some compounded by and intertwined with testing), and others already concentrating on basic cognitive skills that most regular classmates had already mastered. In addition, the BST did not serve as a spur to attention and resources, as a number of other, perhaps more powerful forces already served that purpose, including the inclusion movement, the pressure of lawsuits, federal emphasis and funding, and the emergence of No Child Left Behind standards for students with disabilities and IDEA. Hence, while much effort was being expended on exam-level material, this would have happened without the BST being an exit exam. An additional factor may be that the exams have less pressure than they might due to accommodations and different standards, without which there might be more pressure for results and accompanying programmatic change.

Of the special education exam-related resources, 87% are for testing itself, most notably for the time taken in IEP conferences on exams and retaking the exams. Additionally, expenses associated with the multiple retakes caused by the graduation requirement include the following: after 12th grade students have taken the exam for the final time, the pro-

cess of petitioning for reclassification takes specialist time; Physical Health Disability (PHD) accommodations for retakes require time; and emotional behavioral disorder (EBD) test-taking practice uses resources. Some additional programs, such as teaching content teachers to teach language, are of educational benefit beyond the exam context.

The special education costs add about \$3 per student to the total per student cost of the examination, for a total of \$131. These costs are not evenly divided among students but are instead concentrated on students at risk of failure or with demonstrated failures. For purposes of comparison to Indiana, we have not fully included the special education cost estimates into the Minnesota current cost estimates.

Current State-Level Costs

The state provides a variety of types of funds to the districts, some of which are used for exit exam programs and are included in the district-level estimates already presented. For example, 2% of the base per student amount of \$4,601 is specifically for professional development, with one-half of the allocation dependent on teachers' decisions, one-quarter on school boards' choices, and one-quarter for state-determined exemplary programs. Many of these programmatic choices are influenced by the BST. However, we have already included them in our district cost estimates. Additionally, summer school funding, which passes through the districts and increased significantly with the institution of the BST, has been cut recently.

The state of Minnesota also funds one exit exam-related program directly, without funds passing through the hands of the districts. In fiscal years 2004 and 2005, \$9 million was appropriated for statewide testing, of which over one-third was for BST testing. The other departmental activities that might conceivably be linked to the BST, such as publicity about the test, are part of a single appropriation lacking an obvious means of allocation among purposes. With a 15% cut in the department's budget, many supportive activities in assessment have been reduced.

Most of the conduct of the BST may fairly be entirely allocated as an exit exam cost. "There wouldn't be a BST if it were not a high school graduation requirement," stated Reg Allen, director of assessment for Minnesota, in an interview. The BST might be replaced by a standards-based test in 8th grade as required by NCLB (and BST cannot fulfill that role) and, less likely, by some other 10th grade writing component with similar initial costs. However, much of the cost of administering the exams comes with repeat administrations: multiple administrations during the year—with a basic cost independent of the number of students taking the test—and students taking the exam repeatedly, which would not occur if the exam were not a graduation requirement. About \$1.3 million is the cost of the April and summer administrations, which would not occur without the graduation requirement. Another quarter-million dollars is the estimated cost of administering retakes, as one-third of the cost of the \$1.9 million peak-time administrations is size-related, and somewhat less than one-half of the administrations are retakes. The exit exam-related portion of the state administration of the BST averages under \$2 for each of Minnesota's 850,000 students.

Total Current Cost

With the inclusion of state-level costs, the cost per student of Minnesota's BST is \$130. With the addition of the increased special education estimates, the cost per student is \$133.

District-level costs per student	\$128
State-level costs per student	\$ 2
Special education costs per all students	\$ 3
Total per student cost	\$133

Comparison to Indiana

The meaning of the results in Minnesota becomes clearer when they are compared with the prior findings in Indiana. In Indiana, the costs of the Graduation Qualifying Examination are more than three times as high (see **Table 2-9**). Indiana’s costs are concentrated at the school level, and its district (corporation) level costs do not exceed Minnesota’s to anywhere near the extent that the school-level costs do. Indiana makes much more use of specialists in curriculum and assessment than is the case in Minnesota.

Table 2-9: Per Pupil Local GQE Costs

	SCHOOL	CORPORATION	TOTAL
Personnel	\$ 318	\$ 60	\$ 378
Other	\$ 34	\$ 30	\$ 64
TOTAL	\$ 352	\$ 90	\$ 442

Correspondingly, Indiana’s exam spending focus (**Table 2-10**) is more on professional development and testing than is the case in Minnesota. The percentages may be a misleading, as the dollar amounts that Indiana and Minnesota spend on prevention is in an 11:8 ratio, with Indiana spending more. However, because Indiana spends so much more for everything else as well, the percentages do not reflect the dollar amounts spent.

Table 2-10: Local GQE Programs

	SHARE OF COSTS
Remediation	29%
Prevention	28%
Professional Development	25%
Testing	18%
Other	1%
TOTAL	100%

The state-level costs in both states are \$2 per student and are similarly spent primarily on statewide administration of testing. Whatever the difficulty of the test, the administrative costs are similar.

Explanation: Basic Skills

The Minnesota exam is a basic skills test, while Indiana's covers end-of-9th-grade material. More of Indiana's students fail to qualify on their initial attempt at the exams than do Minnesota's, and nearly one-seventh of Indiana's students have not passed the test as they approach graduation, while almost all of Minnesota's students have qualified. Consequently, Minnesota can target relatively fewer students with specific problems over a longer period of time—sometimes, maturation is enough to do the job. Indiana has more students with more problems to be handled more quickly after an initial failure. As a result, Minnesota can target middle school children while Indiana concentrates on high school students. To educate any student up to the 8th grade level demanded by the BST is less costly than to educate a student to the higher level demanded by the GQE, and most students who fail the BST are unlikely to have far to go to raise their scores enough to pass, while the GQE failures are often reflective of more extensive shortfalls of the standards applied.

In addition, Minnesota's costs are somewhat higher, from salaries to benefits, than Indiana's. Correspondingly, the spending gap would be even wider if costs per resource were considered.

Costs of an Exam with Higher Standards

District Level Cost: What If the Standard Were 10th Grade Material?

For the same hypothetical district for an average student, we asked two panels in Minnesota what would be the added costs of the BST if the graduation qualification requirement were moved to the 10th grade exams for math and language, retaining the current 10th grade writing test. We specified that the district would continue to achieve the same level of initial and graduation pass rate success as the state average now enjoys. Some panel members were skeptical that this level of success could be achieved at any price, as enough middle and high school students arrived from outside the country with limited English so that, combined with students with learning and emotional disabilities, mastering 10th grade material might be beyond the reach of more than 1% of the students. Other panel members welcomed resources and incentives to provide better education to marginal students.

Total Costs

The costs of meeting a BST with a higher standard are illustrated in **Table 2-11**. They are roughly triple the costs of the current examination, and they are concentrated on programs involving school-level personnel. More than three-quarters (78%) of the recommended resources are at the school level and five-sixths (83%) of the resources are for personnel. Regular classroom teachers account for 52% of the costs, a somewhat higher percentage than for current exam-dependent expenses. Both panels recommended sizable investment in new school-based personnel outside the existing classroom, such as paraprofessionals for early childhood and kindergarten, as well as community and parental asset-building liaisons. At the district level, the largest expenses were for technology, such as a language software package for grades 2-5, and specialists, such as district-based family counselors and a professional in data tracking to train teachers.

Table 2-11: Added Per Pupil BST Higher Standard Costs

	SCHOOL	DISTRICT	TOTAL
Personnel	\$ 290	\$ 22	\$ 312
Other	\$ 5	\$ 60	\$ 65
TOTAL	\$ 295	\$ 82	\$ 377

Types of Programs

As indicated in **Table 2-12**, with higher standards comes a shift in the relative emphasis on types of programs into more professional development. While much more is being recommended to be spent on prevention than is currently being spent, the largest growth is for professional development. The 10th grade exam is geared to state standards, and professional development for exam-curricular linkages offers more promise than it currently does. Panelists recommended more (or shifted) professional development days, especially for math teachers at all levels. Any statistical summary, of course, underrepresents the diversity of programs recommended, which include individual coaching, summer school, English and math specialists, instruction in parenting, improved computer record keeping on individual students strengths, licenses for software geared to individual work, and diagnostic tests, in addition to programs already mentioned. The review panel deemed some programs suggested by the panels to be too costly to be justified by an increased standard for the BST, and so these programs were deleted.

Table 2-12: Local BST Higher Standard Program Cost

	SHARE OF COSTS
Remediation	5%
Prevention	80%
Professional Development	13%
Testing	3%
Other	0%
TOTAL	100%

Comparison to Indiana What Is

The hypothetical raised standard for the BST—using the current 10th grade tests—is similar to the current standard for Indiana’s GQE, as well as to the exit exams in most states. However, the cost of a raised standard BST is the cost of the current exam plus the added cost of a raised standard—**Table 2-7** and **Table 2-11** need to be combined, as in **Table 2-13**. **Table 2-13** can then be compared with results for Indiana, as in **Table 2-9**.

Table 2-13: Total Per Pupil BST Higher Standard Costs

	SCHOOL	DISTRICT	TOTAL
Personnel	\$ 348	\$ 8	\$ 356
Other	\$ 65	\$ 84	\$ 149
TOTAL	\$ 413	\$ 92	\$ 505

The costs appear similar for the two exams. The cost per student is about \$62 higher in Minnesota. Minnesota panels recommended (**Table 2-14**) a heavy focus on preventive programs—beyond even that shown by current spending—while Indiana (**Table 2-10**) offered a more diverse fare. Perhaps Minnesota has an educational culture oriented to preventive intervention, while Indiana includes stronger elements of individualized remediation. In any case, Minnesota would spend triple Indiana’s amount on prevention, while Indiana would spend substantially more for each other type of program.

Table 2-14: Local BST Higher Standard Program Cost

	SHARE OF COSTS
Remediation	11%
Prevention	76%
Professional Development	10%
Testing	4%
Other	0%
TOTAL	100%

Similar Tests, Similar Costs

The difference of magnitude in the estimates in Minnesota between the costs of the current exam and that of a higher standard exam indicates that the panel task groups do not simply seem to arrive at high estimates no matter what. Further, the similarity between the cost estimates for Minnesota’s higher standard exam and Indiana’s current exam suggests that differences among states and among actual and hypothetical exams are surmountable barriers to arrive at similar cost estimates for similar exams.

A further bit of evidence comes from the Minnesota panel, which analyzed the costs of a third alternative, a prototypical state exit exam in a district that the national average student might attend, with exam results reflecting the average state experience. This was intended to explicitly allow cost comparisons among states. However, because it involves a 10th grade level exam, a student population with more minorities and more limited English students, and a level of initial and ultimate success between that of Minnesota’s BST and Indiana’s GQE, it can serve as a third similar exam. In this case, the Minnesota panel is estimating the entire cost of a 10th grade exit exam.

Rather surprisingly, the cost estimates for the comparative exam are somewhat higher than either of the estimates for the similar exams. It might be due to the increased number of students having difficulties with the exams. The panel targeted funds for teachers and specialists at the school level (Table 2-15).

Table 2-15: Per Pupil Local Average State Exam Costs

	SCHOOL	DISTRICT	TOTAL
Personnel	\$ 610	\$ 17	\$ 627
Other	\$ 0	\$ 12	\$ 12
TOTAL	\$ 610	\$ 29	\$ 639

With a larger share of children from historically lower-performing groups than is the average case in Minnesota, funds were targeted to early prevention (Table 2-16), including Early Childhood Education (ECE) and extra teachers for block scheduling of instructional groups based on performance and ability. Additional programs train teachers to disaggregate scores and to differentiate teaching, put technical staff for computing in the schools, add a system-wide analysis and diagnosis package and mobile computers, and provide summer school for at-risk middle school students, as well as remedial programs such as half-size classes for high school students who have failed.

Table 2-16: Local Average State Exam Cost

	SHARE OF COSTS
Remediation	8%
Prevention	49%
Professional Development	24%
Testing	19%
Other	0%
TOTAL	100%

The Minnesota emphasis on prevention shows up in all panels and may reflect an aspect of state educational culture.

Summary for Minnesota

Minnesota's Basic Skills Tests (BST) require an 8th grade mastery of mathematics and language arts. Compared to tests requiring command of 10th grade material, they are not difficult, though a large share of Minnesota students require remediation and retake at least one part of the exam. Minnesota targets its BST programs to prevent initial failure, focusing school personnel on students who have had difficulty on earlier exams and using district personnel for resources such as truancy programs. Students needing remediation after failing the BST are offered summer school, special math classes, and talented teachers.

The low difficulty of the BST influences the pattern of resource use. Relatively few students who regularly attend regular class in a district for eight years fail an exam, and if those students regularly attend for the next four years, they are nearly certain to graduate. Instead, the focus is on “problem” students. Special education students require an additional \$26 per student because of the BST being a graduation requirement. Exit exam or no exit exam, the school district should—and often would—provide special programs for these students that would help bring their skills in language and mathematics up to an 8th grade level, so the marginal cost of the exit exam is relatively low and concentrated. The special education costs increase the resource requirements to \$131 per student in Minnesota (and, to anticipate, to \$319 per student in Massachusetts).

State-level programs in assessment and oversight add another \$2 per student in Minnesota (and \$7 per student in Massachusetts) to the resources whose use depends on the exit exam. State programs that pass funds to the district level, such as grants for remediation programs, are already included in the district-level spending.

Minnesota's BST involves fewer programs and much less cost than does Indiana's exit exam, the GQE. The most direct explanations of the differences are that Minnesota has a less demanding exam, a longer period in which to remediate, and an exit exam not tied to a standards-based curriculum. However, many of the patterns of spending are the same, beginning with the concentration of expenditures at the school and district level and the concentration on personnel. Indiana spends more on professional development and testing than does Minnesota.

When Minnesota educators estimated the district cost of a 10th grade exit exam at \$377 per student, it was triple the cost of the current BST. Added to the cost of the current BST, the resources needed exceed those required for Indiana's 10th grade GQE, and the resources shifted toward school personnel and professional development, as in Indiana. However, even in the additional programs, Minnesotans kept their emphasis on preventing initial failure on the exams.

For a hypothetical 10th grade, average state exit exam, with more students in difficulty and ultimately failing the exam, Minnesotans estimated a district-level cost of \$639 per student, the most expensive of the state's three estimates, with an emphasis on testing and professional development costs rivaling those of Indiana.

CHAPTER 4

Massachusetts

Massachusetts is a larger than average, northeastern state with a comparatively well-educated and wealthy but otherwise average population. The Boston metropolitan area is one of the nation's largest.

Table 2-17: Massachusetts Public School Data

	2002
STUDENTS	967,853
Minority	24%
Free lunch	25%
LEP	5%
EXPENSE/PUPIL	\$8,005

Massachusetts educators are comparatively well paid, allowing them to nearly afford the local housing costs. Tests results are commonly near or among the top group of states, as befits the home of public education in America.

Table 2-18: Estimated Average Public School Salary and Benefits in Massachusetts, 2003

CATEGORY	AVERAGE
Teachers	\$49,250
Librarians/Media Specialists	\$53,550
Technology Specialists	\$58,250
Guidance Counselors, Curriculum Specialists	\$63,100
Instructional Aides	\$17,250
Clerical/Data Entry	\$25,750
Principal	\$80,300
Assistant Principal	\$64,250
Superintendent	\$111,550
Assistant Superintendent	\$96,700
Supervisor/Coordinator/Director/Manager	\$96,700
Substitute Teacher (daily rate)	\$150
Benefit Ratio	26.2%

MCAS

The Massachusetts Comprehensive Assessment System is the exam used in Massachusetts to determine eligibility for graduation from high school. The system was implemented in response to the Education Reform Act of 1993, which mandates the testing of all students in designated grades who attend a school receiving any public funds. This law also requires that students achieve a competency determination on the 10th grade exam prior to receiving a high school diploma. Students may retake the exam up to four times in the years following 10th grade, or they may be granted a competency determination through an appeals process. The Massachusetts Board of Education defined a competency determination in January 2000 as “a score of 220 on both the English Language Arts and Mathematics MCAS grade 10 tests,” which corresponds to the performance level description of “needs improvement” or better.

The MCAS has a relatively high initial pass rate, in part because it is nested within a statewide standards-based assessment system. Further, individual and aggregate results are closely monitored and corrective action is mandated. For individual students, this includes a plan for improvement for any student failing the 8th grade exam. For districts and schools, this includes a state remediation plan.

Contemplated Changes

Massachusetts graduated its first class under the MCAS system in 2003, and educators are presently celebrating their achievements, having graduating a high proportion of seniors. The next tasks may be linked to NCLB, especially as the 10th grade MCAS exam is also the 10th grade NCLB exam, and the NCLB-required “proficient” performance is a notch above the performance presently required for passing the graduation requirement.

Costs of Current Exam

District Level Costs: What Is

To estimate the school and district costs of the MCAS, APA created a hypothetical district that resembled one the average student in the state attends. The hypothetical district has 4,700 students in 11 schools, including two high schools (9–12) with 900 students each and an alternative school with 150 students. For all panels, the two high schools were agreed to be consolidated into a single school with 1,800 students, more typical of Massachusetts than our original specification. Of the hypothetical total, 14% of students have disabilities and 3% of students are in limited English proficiency programs.

In this hypothetical district, the MCAS tests, test-takers, and results closely followed current statewide figures:

■ Initial (10th grade) district pass rate on Reading	89%
■ On Mathematics	80%
■ 12th grade students with a Competency Determination by spring	95%
■ Students with disabilities share of test-takers	14%
■ LEP student share of test-takers	3%

- Attainment of Competency Determination on the Grade 10 test:

Students with disabilities	46%
Free/reduced price lunch students	51%
African-American students	52%
Hispanic students	44%
LEP students	34%

Panels of superintendents, principals, specialists, business managers, and teachers were asked to specify the programs and resources needed, because the exam was a graduation requirement rather than simply another student assessment, for this hypothetical average district to produce average results as described above. One analogy offered for their task was to advise the superintendent of this hypothetical district in a state that was similar to Massachusetts but had just adopted an exit exam what programs and resources he or she should use to achieve the above results. In sum, panelists were provided a state microcosm at the district level, then asked what programs, with which resources, would be required to produce the MCAS results, given the district characteristics. On a per student basis, these estimates apply statewide as the local portion of costs.

District Cost

The estimated marginal cost of programs due to the MCAS exam graduation requirement is \$316 per student, a total of \$306 million a year, or 3.95% of the state education spending. The programs and resources are concentrated at the school level, primarily (63%) for teachers, and also for specialists at both the school (11%) and district (8%) levels. These are the costs that are directly dependent on the high stakes nature of the exams, such as retakes and remediation, and programs that would not be in place were it not for the behavior-changing incentive of the graduation requirement.

Table 2-19: Per Pupil MCAS Exam Costs

	SCHOOL	DISTRICT	TOTAL
Personnel	\$ 253	\$ 32	\$ 285
Other	\$ 3	\$ 28	\$ 31
TOTAL	\$ 256	\$ 60	\$ 316

Types of Programs

Most of the programs are for prevention of an initial failure on the 10th grade exam and remediation after an initial failure (**Table 2-20**). These programs tend to be specific to the examination, such as tutoring and parent conferences for students who have failed the 8th grade examination, MCAS prep classes for those in danger of failing, after-school remedial tutoring, staff devoted entirely to remediation, guidance counseling for students needing

remediation, MCAS-oriented curricular revisions by 6th-10th grade teachers, substitutes during retakes, a public awareness campaign, and lost class time for retakes.¹⁰ The testing programs are similarly specific, such as personnel to track competency determination. More general program changes spurred by the graduation requirement include a professional developer at each school for professional development and an assessment tracking and analysis software package, including a test item bank. While panelists desired a sophisticated package such as that developed by the Boston Public Schools and believed that the state should provide one, this was not included in the specification of programs actually in use.

Table 2-20: District MCAS Cost Programs

	SHARE OF COSTS
Remediation	27%
Prevention	32%
Professional Development	32%
Testing	9%
Other	0%
TOTAL	100%

Because of their recent experience in bringing large number of students up to the “needs improvement” level so that the students could graduate, the panels in eastern Massachusetts were proud of the programs in place, and these current programs were vividly described by all panels, not just those panels whose task it was to describe the MCAS as it is. If anything, the estimates of the cost of programs may be low, given the testimony at the other panels. Because the review panel limited itself to describing programs in categories that had gone unmentioned by the district panels, the estimates above are those from the two “What Is” panels.

State-Level Costs

The chief state program specific to the MCAS graduation requirement has been a \$50 million fund for remediation made available for district use. This funds the majority of the \$82 million that the district panels indicated was spent on remediation. The state legislature, however, for the 2004 fiscal year has cut this from \$50 million to \$10 million, leaving the districts to their own resources for most of the remediation cost, as they already are for almost all of the testing, prevention, and professional development expenses. Any funding through the districts has already been included in the district-level spending estimates.

The Office of Educational Quality and Accountability and the Department of Education jointly administer a School and District Accountability System that is used to gauge the progress of schools and districts toward getting all students to reach proficiency in English language arts and mathematics by 2014, the principal goal of NCLB. DOE concentrates on data collection, fact-finding, and review panels, while EQA takes the lead for performance evaluation by the use of a “proficiency index” that measures student (and school and district) proximity to

¹⁰ See Appendix II for an example of programs for one Massachusetts high school.

the proficiency level of performance. The EQA budget was \$2.5 million for FY 2003, and the DOE student assessment budget was \$18.7 million, with both increasing slightly for FY 2004.

It is not clear what share of these funds should be considered as exit exam-related. Most of the programs came in as part of the MCAS system and have been recently revised in light of NCLB requirements, especially regarding proficiency as a standard. While an argument could be made that the entire expense would continue if high school graduation requirements were not tied to MCAS, at least given the current presence of No Child Left Behind, the recent legislative session saw cuts in any program deemed non-essential. Remediation was cut because of public, though not educators', perceptions that the programs were not being widely used by students. In such an atmosphere, programs that are both effective and linked to a high school graduation requirement have a much higher likelihood of being funded than ones that are ineffective or are not linked to exit exams. We set the combined direct and indirect (political) exit exam share of EQA and DOE assessment costs at one-third, or about \$7 million, slightly over \$7 per student.

Panelists repeatedly asked for additional state-level resources, notably in the form of computer software to help collect and analyze student test data in a way that could help correct problems. Boston's software was mentioned as a standard that the state could emulate.

Current Cost

The current cost includes the district-level cost and the state-level cost.

District-level per student costs	\$316
State-level per student costs	\$ 7
Total per student costs	\$323

Comparison to Indiana and Minnesota

In order to understand the meaning of the Massachusetts results, comparison to the earlier results is helpful. The cost estimates (per student) for the MCAS are between those for Minnesota's Basic Skills Tests and for Indiana's Graduation Qualifying Examination. Further, the results in Massachusetts show much lower costs than for a hypothetical 10th grade exam in Minnesota, either its existing examination or an average state exam, so the costs seem lower than a 10th grade exam would warrant. Further, the Massachusetts panelists were convinced that their exam was more difficult than the other 10th grade exams that they had heard about, and more difficult tests require more resources to educate students to pass them. So the low costs are puzzling given the estimated toughness of the test. Additionally, Massachusetts has comparatively well-paid teachers and a low student-teacher ratio, which would be expected to elevate Massachusetts' cost-per-program above that for other states. While Massachusetts does not have a notably large share of its students with disabilities, limited English proficiency, free/reduced lunch eligibility, or minority status, neither do Indiana or Minnesota.

Three possible explanations are the Hancock case, the newness of the MCAS graduation requirement, and the close-knit MCAS package.

The Hancock case is a continuation of the case that originally spurred the governor and legislature to reform Massachusetts education financing. The reform package included

MCAS curricular and assessment reforms in return for increased state funding, especially for districts with a weak tax base. The original case was virtually uncontested, as all conceded that financing was inadequate at least in some districts given the limitations imposed by Proposition 2.5 on local property taxation. The current case, however, is highly contested and pits several districts, including Springfield, against the state. The districts believe they need additional funds to provide a mandated education. Traditionally, teacher unions and superintendents take a similar view. The case has had the effect of dividing those involved in Massachusetts education to such a degree that APA could not get interviews with Department of Education personnel because they feared that their remarks would somehow operate against the state in the court case.

In theory, supporters of the state could be providing much lower estimates than they would in the absence of a court case, as arguments and information have become mobilized against “throwing money at the problem.” Certainly, there were numerous comments at and about the panels indicating that people saw the significance of the results in terms of the court case. As a result, at the review panel it was agreed that APA would state the following: the results of these panels do not necessarily reflect the views of any single participant.

In practice, the court case seems unlikely to account for the low estimates. An equally good argument could be made that the estimates should have been high—most of the participants are the sort of people friendly to greater funding for education. Indeed, one of the two panels estimating the costs of the current system was held in Springfield, the most sizable district on the plaintiff’s side, so it would be easier to dismiss high estimates as the consequence of the court case than it would be to dismiss low estimates. More generally, this intended effect model is implausible: 1) only the review panel saw any monetary estimates; the rest merely discussed specific programs, and the review panel made few changes in the programs; 2) the panels provided many programs, and the comparatively low cost comes from the consistent frugality with resources of the programs, whereas a simpler low-cost strategy would simply be to avoid including programs; 3) educators tend not to be strategic in their behavior in general, and low-balling each of two dozen programs requires lawyerly strategic persistence; 4) the marginal cost of an exit exam is far removed from the actual issues in the case, which concern the entire range of education, including some education not yet being provided.

Massachusetts graduated its first class under the MCAS exam last year, which separates it from Indiana and Minnesota, who have experienced a notably longer period of graduation under the requirement of passing an examination. While it is obvious that this difference could be related to the cost estimates, it is not obvious at first consideration why it would lead to lower estimates rather than higher ones. Indeed, when panelists in other states were asked about change to an exam more like Massachusetts’, they provided high estimates. Massachusetts has just been through the period of considerable change and turmoil in curriculum, assessment and remediation, so these costs are clear in panelists’ minds in comparison to Minnesota, where the BST was so long-standing a part of the system that its historic consequences were invisible to panelists.

The Average State Examination

Before considering the third explanation, we turn to the results in Massachusetts for the hypothetical average cross-state examination. Recall that in Minnesota the costs were over \$600 per pupil for an average state exam with average results in a nationally average school district. In Massachusetts, panelists estimated dramatically lower costs for the identical situation (Table 2-21).

Table 2-21: Per Pupil Local Average State Exam Costs

	SCHOOL	DISTRICT	TOTAL
Personnel	\$ 125	\$ 8	\$ 133
Other	\$ 0	\$ 12	\$ 12
TOTAL	\$ 125	\$ 20	\$ 145

The lion's share of the funds (86%) goes for teachers' salaries and benefits at the school level, often for tutoring, summer school, and after school remediation programs. Remediation is the main focus (Table 2-22). The professional development funds are to train teachers in test result analysis and to train teachers outside the fields of math and English how to use their subject matter to bolster students' proficiency in the testing fields.

Table 2-22: Local Average State Exam Cost

	SHARE OF COSTS
Remediation	59%
Prevention	6%
Professional Development	25%
Testing	10%
Other	0%
TOTAL	100%

For Massachusetts panelists, the average state exam would only be one-half as expensive as the MCAS, notably lacking the funds for prevention that the current test requires while maintaining the same funding for remediation. There was no part of the average state high stakes testing situation that impressed Massachusetts panelists as more demanding than their current situation, despite the larger minority population and higher initial failure rate. So the estimates of the costs to produce those results were not high.

Factors Affecting Costs in Massachusetts

The single key indicator of the importance of support is the high initial (10th grade) rate of passage of the MCAS. On a mathematics exit exam, for instance, 32% of the students in Indiana and 33% of the students on the average state exam fail, leading to a large remediation problem and retesting cost and emphasis on prevention. In Minnesota, 28% of the students fail in their initial attempt at the 8th grade exam, leading to sizable numbers of retakes. In Massachusetts, however, only 20% of the students fail in their initial 10th grade attempt, leading to a smaller problem with remediation and testing and also to a narrower group of probable failures to target among 8th grade students, and correspondingly less need for preventive intervention. The relatively high passage rate on the initial attempt produces a less expensive examination because fewer additional programs are needed to educate students to pass the exams. Prevention is cheaper than remediation in education as in health care. Most of Massachusetts' categorical grant programs and funds are targeted for prevention—early childhood, kindergarten development, early literacy, and early intervention tutorials. Prevention may be harder to separate as a specific exit exam cost because it is preventive across a number of later outcomes or programs, rather than being narrowly targeted to exit exams.

A likely explanation for the comparatively low costs in Massachusetts is the support that already exists for success on the examination. Part of this support is the integrated package of reforms of which the exit exam has always been part. The integrated package of standards-based education and assessment starting in elementary school, topped by the high stakes test in high school, makes it difficult to separate out any element and consider its marginal cost. Indeed, Paul Reville of MassINC, who was one of a handful of prime movers for the reform package, argued that it was unwise to attempt to isolate the cost of any component of the MCAS or the accompanying policies. The relatively high cost of state-level programs in Massachusetts due to extensive statewide monitoring of performance is one aspect of this close-knit system. One of the reasons for the high costs of prevention in Minnesota is the lack of integration of the BST to the curriculum and the standards; in Massachusetts, the integration tends to hold down the costs of the programs to support the exams. For example, each failing student gets an individualized success plan well before he or she reaches high school.

If the standards-based curriculum and integrated assessment are working, then it is comparatively inexpensive to administer a high stakes test, at least one with “needs improvement” as its threshold of passage. Support can be highly targeted to problems closely related to the exit exam and to groups with difficulties, rather having to address basic education problems for a considerable share of the student population. That appears, so far, to be the lesson of Massachusetts.

Limited English Proficiency

Learning from Indiana, APA conducted a separate panel in Massachusetts focusing on the exam costs for students with limited proficiency in the English language. This panel was similar to the special education panel in Minnesota, except that participants included not only specialists in LEP but also non-LEP personnel from both average and high LEP districts. The aim was to provide expertise about LEP costs but within the context of the other programs and students of a district. The district provided was the same average Massachusetts district as for the panels examining broader costs, but the focus was on programs for LEP that would not have been provided without the high stakes testing.

Table 2-23 indicates the overall per LEP student costs of the exam-related programs. Massachusetts panelists identified passing the MCAS exit exam as the top educational priority of LEP students, with considerable school-level resources developed to educate students so that they would be successful. The most expensive of the recommended programs focus on teaching for students who have failed in-class MCAS-focused exams and for tutoring and supervision of the tutors.

Table 2-23: Per LEP Pupil MCAS Exam Costs

	SCHOOL	DISTRICT	TOTAL
Personnel	\$ 1,938	\$ 0	\$ 1,938
Other	\$ 0	\$ 77	\$ 77
TOTAL	\$ 1,938	\$ 77	\$ 2,015

The first question raised by the LEP panel was “how is LEP defined?” The problems for the 1% of students who have not been schooled at all before arriving in the state are different from those of the 1% of students who have some disadvantage (disability, poverty) in addition to language, which in turn are more severe than those of most (3% of total students) LEP students. However, students who are no longer officially classified as LEP still have language-based difficulties with the exams. Hence, one of the top programs recommended is to benefit these students: for LEP teachers to coach non-LEP teachers on the characteristic problems of LEP students and what to do about them. This tends to increase the apparent cost of LEP programs, as the cost is calculated for the narrow base of LEP students, though some of the benefit is more widespread. Accordingly, it may be better to calculate the LEP costs on the basis of per student, LEP and non-LEP combined, as in **Table 2-24**. Note that this adds about one-third to the cost of the examination, as the other panels did not discuss programs for LEP students. These costs are more in proportion to the number of official and unofficial LEP students among those retaking the exam.

Table 2-24: Per Massachusetts Student LEP Exam Costs

	SCHOOL	DISTRICT	TOTAL
Personnel	\$ 97	\$ 0	\$ 97
Other	\$ 0	\$ 4	\$ 4
TOTAL	\$ 97	\$ 4	\$ 101

The programs for LEP students focus on remediation and, as mentioned, on LEP-oriented professional development for teachers of non-LEP students. The standard LEP classes are geared toward helping students with English and are relatively unaffected by the MCAS graduation requirement, except when substitutes are needed because the LEP teachers must administer retake examinations or when so many students are taking remedial tests that the class time is educationally ineffective. One difficulty mentioned is that many LEP students have to work, so getting the commitment to remedial programs outside the classroom can

require parental conferences and alternative after-school, Saturday, and summer scheduling so that students can find a feasible time slot.

Table 2-25: LEP Programs for MCAS Exit Exam Success

	SHARE OF COSTS
Remediation	60%
Prevention	15%
Professional Development	24%
Testing	2%
Other	0%
TOTAL	101%

Total does not add up to 100% due to rounding.

Comparison Between Special Education and LEP Programs

While the special education panel in Minnesota indicated little impact of the BST on special education costs, the LEP panel in Massachusetts described an intense focus on the MCAS exams. It is not clear whether these results are generalizable. Special education may be so need-driven, with comparatively abundant resources, that the logic of exit exams does not really apply, while for LEP students the high school diploma may be more valuable and motivating and the exit exam the crucial barrier. One observer suggests that “IDEA legislates for the additional educational needs of students with disabilities, whereas there is nothing equivalent in place for ELL’s.” If so, the results would likely apply in the same way in different states.

Alternatively, there may be peculiarities in these specific states. The Minnesota costs are small, but panels were not reticent about indicating that larger amounts would be committed if the need were greater, as in the case of more stringent exams. The LEP figures are much larger than Massachusetts panels have suggested for other aspects of the MCAS, suggesting that the figures should be taken at face value. Unlike most Massachusetts students, many LEP students transfer into the state well along in their K–12 years, so they miss receiving the support that is available to other students. A similar problem with LEP students was discussed extensively in several Minnesota panels, with a similar diagnosis of a costly solution, especially for an exam geared to 10th grade material. Minnesota has a considerably lower initial pass rate (30%) for its students with disabilities than does the average state (40%) or Massachusetts (46%), arguing at least that the Minnesota panelists’ analysis does not stem from an abnormally high success with an easy exam. These considerations suggest that the special education and LEP results should be taken at face value as descriptive of the costs of exam-related programs for these populations in general.

Revised Current Costs

If the LEP costs are added to those of the more general estimate, then the marginal cost of education for the MCAS graduation requirement is \$424 per student. If the LEP and special education adjustments should be added to the estimates for every state, the MCAS would remain as comparatively inexpensive as before.

Costs of a Higher Standard of Proficiency

For the same district for an average student that was used to estimate current costs, we asked two panels in Massachusetts what would be the added cost of the MCAS if the standard for Competency Determination were raised from the “needs improvement” category of a scaled score of at least 220 to the “proficient” category of a scaled score of 240 or more. What would be the cost if all students had to get to 240 in order to graduate? Under current definitions used for NCLB in Massachusetts, schools will be responsible for students achieving a “proficient” score in the future. We specified that the district would continue to achieve the same level of initial and graduation pass rate success as the current state average. Some panel members were skeptical that this level of success could be achieved at any price, as “proficient” was perceived as average college prep or better. For example, with vocational tech students spending only one-half of their time in academic instruction, what chance would most of them have to reach this level of academic proficiency?

Panels had difficulty with the task. In part, the context is one of perceived recent success at the existing MCAS standard by exerting all possible effort and skill. Hence, panel members often found it difficult to think of what additional to do for the usual student who was marginal on the exam. Indeed, they tended not to clearly focus on the students who currently pass but would have difficulty at the new standard. One panel was reluctant to make any numerical estimates, though it discussed many programs. The other panel focused in large part on special populations, notably special education students. The review panel saw no easy way of combining the two panels’ different results, \$928 and \$139 per student. We present the average of the results of the two panels, recognizing that they are the least reliable of the results.

District Level Cost: What If the Standard for Passing Were “Proficient”?

The increase in standards would, in the average estimate, more than double the costs of the MCAS exam because it would mandate more education, hence more programs and resources. Almost all of the resources discussed were for school-level personnel, two-thirds for teachers and 30% for specialists. This is mostly more of the same: more marginal students, more failures, more remediation, more retakes, more summer school, etc. These panels emphasized for the most part the same types of programs already being used, including teaching teachers how to analyze and use scores, though the scope has broadened. The more extensive problems require an expansion of programs—more MCAS coordination, more substitutes, more problems with students missing class to take exams, more individual success plans for failing students, smaller classes for more students who have failed the exam. Additionally, as the diagnosis of special education partly depends on assessment results and as parents are likely to want to push for the resources available for special education students when their children are failing, the cost of the existing special education programs would be likely to increase.

Table 2-26: Per Massachusetts Student Costs for a Raised MCAS Standard

	SCHOOL	DISTRICT	TOTAL
Personnel	\$ 566	\$ 0	\$ 566
Other	\$ 0	\$ 9	\$ 9
TOTAL	\$ 566	\$ 9	\$ 575

With a higher standard, the emphasis has shifted to prevention, beyond that available with the current support system. The recommendation is to go to more extensive coaching, with in-school specialists in helping fellow teachers learn, to conduct performance analyses by principals on a monthly basis, and to train teachers to conduct performance analyses. For testing itself, as exams become more difficult, as the special education population grows, and as retakes become more frequent, accommodations become more demanded, and students cannot receive exam accommodations that they do not receive on a regular basis, hence the cost of accommodations even outside of testing situations would go up.

Table 2-27: Programs for a More Demanding MCAS Exit Exam

	SHARE OF COSTS
Remediation	6%
Prevention	43%
Professional Development	14%
Testing	37%
Other	0%
TOTAL	100%

Total Costs

We made no estimate for the additional state-level costs or LEP costs associated with an increased standard for the MCAS.

Comparison to Indiana What If

The hypothetical standard increases the difficulty of the MCAS examination in a way similar to the hypothetical standard used in Indiana. The additional cost in Indiana was \$685 per pupil, adding 150% to the cost of the existing exit exam system. The change in Massachusetts is, if anything, more dramatic, and the cost increase is equivalent. The cost estimates for hypothetical changes are less reliable than the costs for the current system. The results, however, are notable enough to suggest that significantly higher levels of performance on high school exit exams will involve the commitment of resources for increased education that is proportional to the increased performance. In the case of Indiana, Minnesota, and Massachusetts the primary focus of increased programs for higher standards are for prevention of initial failure.

Summary

The Massachusetts Comprehensive Assessment System includes a 10th grade test whose passage is a graduation requirement. In 2003, Massachusetts graduated its first class under the MCAS system, and educators graduated a high proportion of seniors. The MCAS has a relatively high initial pass rate, in part because it is nested within a statewide standards-based assessment system. Further, individual and aggregate results are closely monitored and corrective action is mandated. For individual students, this includes a plan for improvement for any student failing the 8th grade exam. For districts and schools, this includes a state remediation plan.

The resources used at the district level because of the exit exam requirement constitute nearly 4% of the Massachusetts budget for K-12 education. The costs are concentrated on school-level personnel and are spread fairly evenly among preventive programs, remedial programs, and professional development. The remedial programs have been chiefly funded by a \$50 million appropriation from the state, which has, however, just cut the program by 80%.

Because MCAS is perceived to be a demanding 10th grade exam and because per student costs for education in general are higher in Massachusetts than in Minnesota or Indiana, the cost estimates for the current MCAS seem low compared to estimates in other states. Further, the estimated cost of the hypothetical average state exam was dramatically lower than the estimate in Minnesota, as well as lower than the cost of the current MCAS, and remarkably few of the resources were to go to prevent an initial failure on the exam.

One key explanation stems from the high initial pass rate in Massachusetts, leading to low remediation and testing costs. Prevention as a strategy is cheaper than remediation, and Massachusetts' strategy of an integrated standards-based curriculum and assessment program may be particularly cost effective. It permits early targeting of students and schools for corrective action that is linked to the regular program of instruction. Support can be targeted to problems closely related to the exit exam and to groups with difficulties, rather having to address basic education problems for a considerable share of the student population. Little of this is identified as a cost of the exit exam, as it is associated with the broad package of reforms of which the exit exam was part.

Not all of the MCAS is low cost. Massachusetts panelists identified passing the MCAS exit exam as the top educational priority of LEP students, with considerable school-level resources developed to educate students so that they would be successful. The LEP emphasis is strongly on remediation. That adds \$101 per student, about one-third, to the total cost of the MCAS exam, as the "What Is" panels did not estimate costs for LEP programs, for a total cost of \$424 per student.

The cost of the MCAS if the standard for Competency Determination were raised from the "needs improvement" category to the "proficient" category would more than double. The additional programs would not focus on remediation of regular students but on prevention, testing, and professional development targeted at students with special needs. This is similar to the estimated impact of raising standards for Indiana's GQE, and not far from the tripling of cost associated with changing Minnesota's BST to a 10th grade exam. Significantly higher levels of performance on high school exit exams will involve the commitment of proportionally increased resources for education.

CHAPTER 5

Comparing Costs, Programs, and Education Across States

In the course of examining the programs and costs for exit exams in Minnesota and Massachusetts, we have compared them across states. The focus, however, was primarily on understanding each state's cost. In this chapter, we attend more directly to the question of cross-state patterns. First, we arrange the results by the grade level tested. Then corrections are applied for the differences in pricing among the states. Discussion of the possible impact of teacher-pupil ratios, state supportive structures, and other state characteristics wraps up this basic presentation.

Subsequently, the results for special education and limited English proficiency students are added, creating a composite cost index. We raise the question of whether these costs can be reliably extrapolated across states, and consider the apparent effects of the recency of the institution or change in the exit exam.

Exam Difficulty and Cost Among States

Table 2-28 presents the per student cost results across states, arranged by the difficulty and failure rate of the examinations. The columns represent the three states. The rows represent different exam situations. In the cells of a combination of a state and an exam situation are the panel estimates of the per student costs of the resources needed in the state for the exam. At the top are the state-level costs, which should be added to each and every exam situation in a state. Below that are the district costs, which vary with exam situation. For purposes of comparison, we have added the cost of a changed examination to the cost of the current exam in order to arrive at an estimated total cost of the exam in the changed format. The state-level costs vary but are of so modest a size compared with district-level costs that they may conveniently be treated as relatively constant in the overall picture.

The costs tend to follow a grid defined first by the difficulty of the examination. Minnesota's Basic Skills Tests is the easiest and the cheapest. In all states, higher standards are associated with higher costs. This implies that the cost is primarily the cost of the education to meet the standards. Additionally, higher initial failure rates are associated with higher costs, with one exception. Initial failure rates are linked to the costs of remediation. Finally, the costs of making the examination more difficult than it is at present appear to be high, while the costs of changing the examination to an easier format seem low. Hence, the Massachusetts panel rated the hypothetical average state exam as less costly than their current, more stringent MCAS, while a Minnesota panel rated the same exam as very costly, viewed from the perspective of the BST. Implicitly, the cost of change is asymmetric: change to an easier format merely removes programs while change to a more demanding format requires additional programs.

Table 2-28: Examination Difficulty and Failure, by State

		PER STUDENT COSTS	MINNESOTA	MASSACHUSETTS	INDIANA
STATE-LEVEL			\$ 2	\$ 7	\$ 2
DISTRICT-LEVEL					
8th grade			\$128		
10th grade	Low* failure			\$316	
	Moderate failure		\$505		
	High** failure		\$639	\$145	\$442
	High standards, low* failure			\$891	\$1,127

* < 20% initial failure on mathematics, < 6% failure to meet graduation requirement by spring of senior year

** > 30% initial failure on mathematics, > 9% failure to meet graduation requirement by spring of senior year

We cannot separate out the different impact of initial and final failure rates. They vary together across the conditions specified for the panelists, except for Minnesota, where the panels considering a 10th grade exam with high initial failure and low final failure two years later were skeptical about the possibility of producing such widespread change in a short period. The higher the initial failure rate, the more the remediation that needs to occur to achieve any given final failure rate. However, the higher the final failure rate, given any initial failure rate, the less the remediation that has occurred. The two failure rates imply opposite remediation costs. Because the panels’ estimates of cost follow the logic of the initial failure rate, we have concluded that it is more significant in accounting for results. This leaves the impact of the final rate unclear.

We will return to a discussion of these results but first will correct them for price differences among the states in order to make more meaningful cross-state comparisons.

Pricing Differences Among States

Differences among exams may account for much of the pattern of costs. The states differ in other ways that may influence cost estimates. For any given resource made available to help students achieve the mastery tested for the graduation requirement, the cost may vary across states. In the abstract, we would prefer to be able to specify the resources, and then calculate the cost, and we would prefer to directly compare the required resources across states, rather than compare costs, which may have an unrelated pricing component. However, panels sometimes provided direct specifications of costs, sometimes resources, so it is not easy to disentangle the cost considerations. Additionally, the distinction between resource and cost can become cloudy. The average principal in one state may be better paid and also better trained. Given the tendency toward relatively equal salaries in education, cost differences within a state might understate resource differences. As resources are ultimately defined in terms of their educational impact, when the impact is unclear, so also is the distinction

between resource and cost: a lower pupil-teacher ratio is costly but whether that cost is considered pricing or resource depends on whether the lower ratio is educationally effective.

For K-12 education, about one-half of the costs are teacher compensation, and other prices may be linked to these salaries and benefits. For practical purposes, a major share of the possible pricing differences can be examined by simply considering teacher salaries.¹¹

Table 2-29: Teacher Salaries in Indiana, Massachusetts, and Minnesota

AVERAGE TEACHER SALARY	
Indiana	\$44,884
Massachusetts	\$49,250
Minnesota	\$44,850

The salaries are higher in Massachusetts than in the other two states. Once we correct for salary differences, as in **Table 2-30**, Massachusetts seems even thrifter.

Table 2-30: Adjusted Costs of Exams, by State

PER STUDENT COSTS		MINNESOTA	MASSACHUSETTS	INDIANA
STATE-LEVEL		\$ 2	\$ 7	\$ 2
DISTRICT-LEVEL				
8th grade		\$132		
10th grade	Low* failure		\$296	
	Moderate failure	\$519		\$466
	High** failure	\$657	\$136	
	High standards, low* failure		\$835	\$1,189

* < 20% initial failure on mathematics, < 6% failure to meet graduation requirement by spring of senior year

** > 30% initial failure on mathematics, > 9% failure to meet graduation requirement by spring of senior year

Beyond teacher salaries, there are numerous differences among states that have not been accounted for, notably pupil-teacher ratios and supporting state structures. Minnesota's BST is notably an orphan, divorced from state support (especially after budget cuts) and the rest of the educational and assessment system. By contrast, Massachusetts' MCAS is integrated into a single, largely statewide system, and performance is closely monitored by a special state assessment agency. Indiana's GQE falls in between, integrated with curricular standards but

11 The cost of benefits, especially those that do not actually leave the coffers of state government, can be more problematic than the cost of salaries and are omitted here.

with less state intervention than in Massachusetts. However, despite the BST status as isolated from the curriculum, Minnesota panels avoided programs smacking of “teaching to the test” and instead emphasized prevention. This might be a difference in educational culture, with Minnesota having an emphasis on collective benefit, which requires actual education.

There are additional considerations that apply to states that are not included in this study. The three included states have population sizes that are average or above average among all states, are eastern–north central, are relatively well educated, with sub–average minority populations, and perform well on national tests. Large and small states, southern and western states, relatively poorly educated states, states with large minority populations, and low–performing states might show quite different patterns or might show similar patterns.

Special Education

The Minnesota panel on special education provides the only direct evidence on the costs of special education that are underestimated and rarely mentioned in the regular panels. The underestimation is presumed to exist in all states for all examinations, so the Minnesota panel provides information that can be used to adjust the estimates for all panels.

The cost per special education student can be translated into a cost per total student by multiplying by the special share of the students—16% in Indiana, 15% in Massachusetts, and 13% in Minnesota. Prices can also be adjusted to reflect pricing differences among the states.

The simplest option is to apply the Minnesota result, adjusted for the special education share of the students and for prices, to all exams in all states. This option provides the lowest estimate of the cost of special education adjustments. The current Minnesota BST is the least expensive of all the tests and, if costs for special education go down in proportion to costs for other students, then the cost for special education might be expected to be higher in other states and for other exams. Indeed, the panel that provided explicit special education estimates had notably higher cost estimates than did the BST special education panel.

A second option is to calculate the special education increment from a ratio of the costs per special education student to costs per regular student. The increment per special education is this ratio minus one times the cost per regular student. The ratio assumes that costs deriving from exam difficulty and failure rate and change from the current exam would affect regular and special education students proportionally.

The Minnesota special education panel, however, suggested a reason why the current BST might have unusually low costs for special education students. Despite the low 30% 8th grade pass rate on the mathematics BST of special education students, little extra was done to remedy these students’ problems because the normal and expected special education procedures focused on these basic skills anyway or on emotional problems that interfered with learning demonstration. As a consequence, special education students had very high rates of passage by the end of their senior year. However, a more difficult exam might pose a more difficult problem because of the disconnect between what special education is currently designed to do and what would be expected in order to pass the examination. The evidence from the Massachusetts panels focusing on an exam with increased difficulty for passage suggests that special education costs could rise much more rapidly than costs for regular students on average.

Lacking adequate data to estimate special education costs for a more difficult exam, we will use the second option, figuring the special education increment from a proportional share of regular education expenses, adjusting for the size of the special education population and pricing differences among the states. The Minnesota estimate is that special education costs are 20% higher on a per student basis, which adds about 3% to the cost of each state exam. However, we suspect that this is an underestimation of the special education costs.

Limited English Proficiency/ English Language Acquisition

The LEP situation is structurally similar to the situation for special education: the costs are underestimated by regular panels, indeed are rarely estimated at all. The sole panel focusing on these costs estimated them for the current MCAS examination. The options are to apply this estimate as a flat amount, to assume proportionality, or to estimate a varying LEP proportion depending on exam difficulty and change from the status quo. We lack any data on which to base a varying proportion.

While the structure is similar, the problem differs somewhat because the LEP panel was composed differently and behaved differently than the special education panel. While the special education panel comprised only special education experts, the LEP panel included regular panelists both from high LEP and low LEP districts. As a consequence, the special education panel focused on the match between special education curriculum and test preparation, while the LEP panel focused on the match between LEP and non-LEP preparation for the exams. Additionally, the special education students and parents do not seem to have a single-minded focus on getting a high school diploma, while the LEP students and parents do. Consequently, the special education curriculum has been less shaped by the exit exam than has the LEP preparation. Finally, because the MCAS exam is already a difficult exam, the cost estimates, to the degree that they vary from proportionality, are as likely to be low as to be high as a basis for inference to other exams.

Using proportionality as the guide, the LEP costs are more than six times greater than the costs for the average student, so the 3-5% of LEP students in these states leads to cost increases of 27% in Minnesota and Massachusetts and 16% in Indiana. These adjustments seem high, probably due to a combination of a low estimate for the costs for regular students and a high estimate of the costs for LEP students. However, they are consonant with the testimony of panelists about the relative importance of the exit exam to LEP students and to its limited importance to regular students in Massachusetts. The estimated total cost of the various exams, including price adjustments and increments for special education and limited English proficiency, are presented in **Table 2-31**.

Table 2-31: Costs by Exam Type and State, Adjusted for Price, Special Education, and Limited English Proficiency

PER STUDENT COSTS		MINNESOTA	MASSACHUSETTS	INDIANA
STATE-LEVEL		\$ 2	\$ 7	\$ 2
DISTRICT-LEVEL				
8th grade		\$171		
10th grade	Low* failure		\$385	
	Moderate failure	\$673		\$557
	High** failure	\$851	\$176	
	High standards, low* failure		\$1,084	\$1,419

* < 20% initial failure on mathematics, < 6% failure to meet graduation requirement by spring of senior year

** > 30% initial failure on mathematics, > 9% failure to meet graduation requirement by spring of senior year

Because the LEP and special education adjustments operate proportionally across exams, they do not change any conclusions about the pattern of differences. Instead, they merely make the estimates larger.

Reanalysis

Analysis of the adjusted figures suggests the most important distinctions among exams. Two of these distinctions are clear from the table: the difficulty (grade level and high standards) of the exam, and the level of the failure rate. Two additional distinctions arise from comparison of the hypothetical results: changes in the difficulty of the exam and change in the failure rate. Within a state, change is measured as the difference between the current exam and a hypothetical exam situation. A multiple regression analysis of the cost against the difficulty of the exam, the failure rate, and the changes in difficulty and failure rate suggests some tentative cost parameters—tentative because with only seven exam formats and four explanatory variables, there are no statistically significant estimates. First, there are no separate state impacts, once pricing has been adjusted for. Second, maintaining an exam of low difficulty and a low initial failure rate would be relatively inexpensive, about \$66 per student, as there would be little extra education or programming required. By contrast, an exam as difficult as the “proficient” MCAS, representing a change from an exam similar to the current MCAS, combined with a high initial failure rate, representing a drop from an astronomical failure rate, could cost about \$2,300 per student. This would be the cost of educational change to produce the results.

The single most important consideration is change in the failure rate: reducing initial failure rates within a state from high to moderate or from moderate to low increases costs by an estimated \$380 per student. These costs are for failure prevention educational programs. Next most important is the difficulty of the exam: more difficult exams cost \$280

more per student than exams of a notch less difficulty. Educating students up to a higher standard costs more. Of similar cost influence is the failure rate: high failure rates mean more remediation to the cost of \$262 per student. Finally, increases in the difficulty of the exam add \$134 per student to the costs. These impacts are additive. In combination, they can produce the wide range of costs from the \$171 for the current BST to the \$1,419 for the GQE with a low failure rate.

Recency

Minnesota has had the BST relatively unchanged since 1996, while Massachusetts graduated its first class under the MCAS exit exam requirement in 2003. Indiana has been graduating students under the GQE since 2000. On the basis of studying Indiana, we hypothesized that the recency of the exam would affect the cost structure, as states moved over time from costs focused on testing and remediation to costs focused on professional development and prevention. That certainly was the tenor of remarks by the panelists, as the Massachusetts panelists presented detailed stories of remediation schemes, while the Minnesota panelists had little interest in short-term responses to the BST. Further, the types of programs bearing the burden of costs follow the hypothesized pattern, with the prevention share of costs much higher in Minnesota than in Massachusetts.

Some caution is in order about the recency effects, however. The response of Minnesotans to proposed changes in the BST and to the hypothetical average state exam was to call for even more prevention, which is not the hypothesized response to change. As the standards increase, the move is to push costs back to the prevention stage. While many of the Massachusetts programs functioned as remediation, they were technically prevention as students who failed 8th grade exams were targeted with success plans and special programs. The emphasis on prevention corresponds to the high standards, but panelists were often at a loss to see how they could meet even higher standards with such tactics.

The most pronounced recency effects were upon the perceptions of costs, rather than the reality. For the BST, the initial costs in response to the new tests were now lost in time and memory, just part of the familiar superstructure, so they were no longer considered associated with the BST. Nowhere was this more pronounced than with the special education panel. Recent packages of reforms skew the perception of costs a different way, by making it difficult to isolate the exit exam portion of the acknowledged recent changes in costs. In Massachusetts, the package was the MCAS, but in every state the effects of NCLB are becoming confused with those of exit exams. More generally, the entire shift toward standards-based education and inclusion tends to be perceived to have similar costs and effects, so the attribution of causation becomes mixed. This undoubtedly reflects a reality of interacting effects.

Program Costs

An exploratory regression analysis was conducted on each type of program cost separately—without adjustments—in order to address some hypothesized differences in programmatic resources. Remedial and professional development costs showed a similar pattern: they were relatively high with the current exams and tended to drop with proposed changes in the

exam structure. Exam increases in failure rate and difficulty focused resources on prevention, and change in the difficulty of the exam also funneled resources into testing itself.

The more difficult exams and those with the highest initial failure rates also were the focus of the most spending on all types of programs—remediation, professional development, prevention, and testing itself. This straightforward pattern is somewhat obscured by the differences among programmatic spending estimates when change is considered.

Of the programmatic results, the clearest—and the only statistically significant ones—are for remediation. Spending on remediation goes up with the difficulty of the exam and with the initial failure rate. However, change has a counter-intuitive effect: increases in the difficulty and the failure rate produce decreases in spending on remediation, while decreases in difficulty and failure lead to increases in remediation costs. These results make sense if remediation is the non-structural response. When problems get easier, resources are concentrated on ad hoc, remedial solutions, while when problems get tougher, resources become shifted to structural change in order to produce an adequate solution.

Perception

We cannot distinguish between “real” and “perceived” costs. Especially for hypothetical exam situations, there are no “real” costs to measure, and even for the current exam structure, the perceived costs at a minimum seem more accurate than any official accounting or classification of costs. Nonetheless, there may be situations in which perception can be systematically led astray. The uncertainty surrounding some programmatic estimates may be such as situation. In general, people tend to project the present and past into the hypothetical future, relying in this case on the recent exam experiences to estimate future costs. However, if the hypothetical is perceived to be too dissimilar to the present, then panelists may be reluctant to rely on recent experience, turning instead to principles and to hopes (or fears) as a basis for prediction. It may be, for instance, that educators hope that prevention programs would be the response to raised standards but that once students began to fail, remediation would become the predominant response.

In a related way, panelists may be underestimating the costs of exams that are easier than their current one and may be overestimating the costs of exam situations that pose greater problems than their current ones. Once the status quo is taken for granted, as in the hypothetical situations, panelists may ignore its cost and focus instead on the marginal cost of change. In the case of an easier exam, this process may underestimate the true cost of providing services and resources, as no new resources are required. In the case of a more challenging exam situation, panelists may ignore the simple solution of extending present programs. These are suggestions about possible sources of misestimate. One future possibility to deal with this problem is to collect systemic cost data in a single district.

Summary

There are enormous differences in the costs of the various exams. The major differences are less across states than they are among exams. We conducted an analysis of the patterns of costs. The costs for special education students and for LEP students are added to all tests by considering them as in the same proportion to total costs, student for student, for each

exam, then weighting them on the basis of the share of the population which is the object of the additional cost. This does not change the cost of the exams proportional to each other within a state. We adjusted costs using teacher salaries to a three-state average salary.

Change in the initial failure rate within a state is the most cost-consequential aspect of the exam. Reducing initial failure rates from “high” to “moderate” or from “moderate” to “low” would cost an estimated \$380 per student. The costs of programs for success on the exit exams tend to follow a grid defined by the difficulty of the exam, from Minnesota’s BST to the “What If” raised standards for Massachusetts’ MCAS and Indiana’s GQE. More education leads to more cost, at about \$280 per student. Higher initial failure rates are linked to higher costs, especially for remediation, of \$266 per student. Finally, increases in the difficulty of the exam add \$134 per student to the costs, while change to a less demanding exam is relatively costless.

These costs are additive, so maintaining an exam of low difficulty and a low initial failure rate would be relatively inexpensive, about \$66 per student, as there would be little extra education or programming required. By contrast, an exam as difficult as the hypothetical “proficient” standard for the MCAS, which represented a change from an exam similar to the current MCAS, combined with a high initial failure rate, which represented a drop from an astronomical failure rate, could cost about \$2,300 per student, about 30% of the Massachusetts per student K-12 education budget. To reiterate, these are the estimated costs of the educational programs to produce the success measured by the exit exam.

The recency of exit exam institution and reform has an impact of focusing narrowly on testing and remediation, while the tendency for an established system is to evolve costs toward prevention and professional development. However, the most notable impact of recency is to change the perception, rather than the reality, of exit exam costs: recent exams have costs that are difficult to separate from the cost of associated reforms while established exams have costs that are difficult to separate from the rest of the ongoing system.

A major finding is the important distinction between exam structure and exam change. Exam structures that are difficult or have high initial failure rates have the expected pattern of costs. However, differences between states in structures do not have the same cost consequences as changes within states in structures. Change produces a different pattern of resource use, notably a surge in prevention programs and a shift away from remediation as the exams become more difficult or the failure rate becomes higher. We suggest that more severe problems call forth structural solutions such as preventive programs.

CHAPTER 6

Expectations and Results, Future Research, and Conclusion

Expectations and Results

The expectations presented in Chapter 1 are summarized in **Table 2-32**, along with the results. Of the nineteen expectations, based on findings in three states, a little under one-half (9) are clearly true; two are partial truths, the truth of four remains unclear, and four are clearly false. If results always matched expectations, we wouldn't need to conduct research. If the initial and final failure rates could have been varied, that would have provided a clear test of expectations about the final rates. In general, the expectations that were incorrect failed to take account of the pattern that increases in the exam's difficulty and failure rate are as important as the difficulty and failure rate themselves in influencing costs. This is perhaps the most noteworthy new idea of the current stage of the study.

Table 2-32: Expectations and Results

	EXPECTATION	RESULT
1	The most sizable expenses would be spent at the school and district level.	True: 98% of spending is at the school-district level.
2	Minnesota's Basic Skills Test would require fewer resources than Indiana's more rigorous GQE.	True: BST: \$128 (\$171 in total standardized costs) per student; GQE: \$442 (\$557 in total standardized costs) per student
3	Massachusetts' rigorous MCAS would be more costly than the BST and GQE.	Half-true: MCAS: \$316 per student (\$385 in total standardized costs)—between BST and GQE
4	High per student exit exam resources for special education students	False: per student costs 20% higher.
5	High per student exit exam resources for limited English proficiency students	True: per student costs more than double and additional costs for students with LEP problems but not in LEP program.
6	A hypothetical, more difficult standard for passage would involve higher costs.	True: 10th grade BST costs an additional \$377 per student; raised MCAS standard costs an additional \$575 per student.
7	Similar exams would have similar costs in all states.	True: costs depend on exam difficulty and failure rate, but also on changes in these.
8	Cost differences among the states for similar types of exams trace primarily to the general cost of education.	True: once salaries are controlled for, there are no state cost impacts.
9	Exit exams integrated with exams in earlier grades and with curriculum have lower marginal cost.	Unclear: even with controls for other factors, the residual costs cannot be used to test the expectation.

Table 6-1: Expectations and Results (continued)

EXPECTATION		RESULT
10	The cost impact of changing an exam traces to the exam structure (e.g., grade level).	False: the major cost of change traces to the implied magnitude of student achievement increase.
11	The task group method uncovers significant differences among types of exams.	True: there is enormous (\$171 v. \$1,419 per student) variation in measured costs of exams, and this variation mainly stems from exam structure change, not from between-state cost differences.
12	Newly implemented exams have high testing and remediation costs, while long-standing exams have high prevention and professional development costs.	Unclear and perhaps not very significant: the distribution of costs among programs does not primarily depend on recency; state culture and exam difficulty appear to be more important factors. Change is associated with higher spending for prevention.
13	Massachusetts invests in pre-exam education, while Minnesota and Indiana focus more on post-exam remediation.	False: while we have no data on the cost of the education itself, the investment of exam-related resources does not fit: Minnesota invests the highest share in prevention. The amounts spent in Massachusetts on remediation and prevention equal the average for Minnesota and Indiana.
14	With a more basic exam, there is less need for remediation.	True: Minnesota spends little (\$38 per student) on remediation, while Massachusetts and Indiana spend twice and three times as much.
15	With a more basic exam, there is less need for early intervention.	Half-true: while Minnesota spends least on prevention for the current exam, the differences (\$83 v. \$95, \$131) are small compared to the prevention costs of making an assessment more difficult (\$322 to \$511 per student).
16	With a more basic exam, there is less demand for professional development.	True: Minnesota spends little (\$3 per student) on professional development, while Massachusetts and Indiana spend about \$100 (\$95, \$117) per student.
17	The higher the initial failure rate, the more that education must occur after the exam (remediation) rather than before-hand (prevention).	False: all types of programmatic expense go up with initial failure rate.
18	The higher the final graduation qualification rate, the more education (and expense) actually has occurred.	Unclear: no separation of effects of final rates from initial rates.
19	A large gap—high initial failure followed by high qualification—suggests a large amount of learning (and expense) focused on remediation.	Unclear: no separation of effects of gaps.

Future Research

The estimates of the costs of types of exams are suggestive, but more states and more variation are required before these estimates can be relied upon. The first task is to add states from the south and west, including at least one state with a larger minority population. If these states included at least one state with an individualistic educational culture, that would increase the variation. Louisiana and Nevada are potential candidates.

Additionally, increased variation in state size might indicate whether there are economies of scale in providing exit exams and supportive structures. New York, Texas, Florida, and perhaps California would be potential large state case studies, and Nevada or another southwestern state would be a potential small state case. While no small set of states can represent all possible variations, three additional states—Louisiana, Nevada, and a large state—might be enough to cover the most important variables. The minimum set might be Nevada and Florida, but this set risks confusing regional, size, and performance effects. Another possibility is to include a state that tests using end-of-course examinations. That might be a useful variation because states are moving in that direction and would indicate how tests tied to specific courses affect costs.

Within states, an effort at estimating the cost of special education for an exam of high or moderate difficulty would help answer questions about the estimate based on the BST. An additional panel focusing on at-risk students, as indicated by free and reduced price lunch participants, would help provide an estimate of their possibly underestimated incremental cost. While other topics such as charter schools are of interest, they are not of sufficient importance to warrant inclusion at this time.

We have additional ideas for development in the methods of estimation. One idea is to add a single all-day panel that considers current costs in the morning and costs for a changed exam in the afternoon. While posing logistic difficulties, such a panel would provide bolstered confidence that the comparison between current and changed costs does not depend on panel composition. A second idea is to conduct a panel aimed solely at an atypical, urban district with a high concentration of students having difficulty passing statewide exams. The cost estimates in districts with low failure rates may reflect slack resources available to be shifted to a relatively few students, a shift not possible when most of the resources are already being devoted to students in difficulty. A third and related idea is to conduct an intense study of a single district, including close examination of its records; while there are undesired interactions that come from having too high a share of members from a single district on the same panel, a firm estimate for a single real district would help test the results for a hypothetical average district. A final idea is to provide panelists with more firm resource parameters, perhaps in the form of estimates from earlier panelists. The most obvious application would be to let the special focus panel build from the district estimates of the more general panel.

One consideration without a current proposed solution is whether the review panel can be used to provide help in reconciling results across panel topics. A second is the desire to allow initial and final failure rates to vary with regard to each other.

Conclusion

If the standards-based curriculum and integrated assessment are working, then it is comparatively inexpensive to administer a high stakes test, at least one with “needs improvement” as its threshold of passage. Support can be highly targeted to problems closely related to the exit exam and to groups with difficulties, rather than having to address basic education problems for a considerable share of the student population. That appears, so far, to be the lesson of Massachusetts’ MCAS.

A basic examination is also relatively inexpensive once it is underway, as is the case in Minnesota. However, an exam such as the BST that is not integrated with a standards-based curriculum and with other assessments can be expensive to change.

The pattern is for costs to increase with exam difficulty and the initial failure rate, most notably for remediation. However, the largest costs are the costs of change, especially the costs for preventive programs resulting from change to an exam with higher difficulty and failure rate. It is cheaper to get exit exams properly aligned from the beginning.

Because the exit exam costs occur primarily at the district level, they may be relatively invisible to state policymakers, except insofar as the districts are spending state money expressly appropriated for exit exams, notably for remediation. Across states, apparently regardless of explicit state authorizations, the spending occurs in accord with the exam structure and difficulty. Using these results, state policymakers can anticipate the educational needs and the accompanying expense.

APPENDIX I

List of Panel Participants and Interviewees

Willingness to participate does not mean that participants necessarily agree with the conclusions of the study.

Part A. Minnesota Panel Participants and Interviewees

Ben Hawkins	Hank Stankiewioz
Ben Perry	Jane C. Holmberg
Bernice Berns	Janet Johnson
Bill Marx	Jerry Williams
Charles Kyte	Judith Coley
Chris McHugh	Judy Schaumbach
Constance Hayes	Julio Almanza
Don Pascoe	Kate Beattie
Dr. Barbara Burke	Lori Rockney
Dr. Beth Sullivan	Marcia Averbook
Dr. Don Helmstedder	Marcia Nelson
Dr. Doug Marston	Mary Pat Cumming
Dr. Jan Witthuhn	Michael Degen
Dr. Kay Worner	Mr. Jim Lee
Dr. Kristen Anderson	Mrs. Bonnie Holme
Dr. Paul A. Gustafson	Randi Johnson
Dr. Phillip Moye	Randy Nelson
Dr. Rick Spicuzza	Reg Allen
Dr. Roger Giroux	Rex Hein
Dr. Wendy Shannon	Robert Meeks
Dr. Zhining Chin	Robert Wedl
Ed Saxton	Sandy Leewandowski
Ellen Wiss	Stan Mack
Eric McDonald	Steve Kelley
Gary Phillips	Steve Niklaus
Greg Crowe	Virginia Richardson

Part B. Massachusetts Panel Participants and Interviewees

Anthony DeMatteo
Barbara Donaghue
Beverly Brown
Cameron Huff
Carla Jentz
Claudia L. Bach
David Danning
David Tobin
Deborah Jose
Donna Scanlon
Dr. Carolyn Wyatt
Dr. Eduardo B. Carballo
Dr. Jo Sullivan
Dr. Ron Souza
Elaine Pisciotoli
James Caradonio
James Keefe
John J. Welch
John Mika
John Robertson
John Schneider
Joseph F. Connors
Joseph Oliver
June Gilch
Kristen Harmon
Marie Ferrari
Mary Elizabeth Beach
Mary Ellen Donahue
Michael Fitzpatrick
Michael N. Dubrule
Michael Sentance
Michael Welch
Ms. Jahn Hart
Paul Madden
Paul Reville
Paul Schlichtman
Peter Finn
Robert Snow
Sandra Guryan
Stephen Pereira
Susan McGilvray-Rivet, EdD

APPENDIX II

New Bedford (MA) High School Services to Members of the Class of 2003 Who Had Not Met Their MCAS Competency Determination (Partial List)

Chairs and faculty from the English and Mathematics departments of New Bedford High School modified the curriculum in light of item analyses of student scores on the MCAS test.

All students who failed a portion of the 8th grade tests were required to participate in an MCAS review course during their freshman year, either a regular course, a summer course, or a 3-hour, once a week, after school, evening, or weekend meeting.

An internet support program was made available for use from home, along with a state-sponsored program (Princeton Review). Additional services for a state-endorsed program were available three days a week, and another state-endorsed program (Smarthinking) was available two of the other days a week. A software-based support program, Plato, was used at the option of staff, with 30 additional licenses and a computer laboratory designed solely for use in MCAS preparations.

MCAS facilitators in English and mathematics provided individual assistance, initially targeted to students who were closest to attaining a passing score. Eventually the facilitators created individual student success plans for all students who had taken but not passed the test. The plans were discussed with students and shared with the appropriate instructors and with parents during individual parental conferences.

A summer program was designed for and offered to the Class of 2003. All students who had not yet passed and their parents were urged to participate.

A supervisor directed all MCAS academic support services and developed an action plan for expanding learning opportunities. A senior year MCAS advisor assisted seniors in passing the MCAS, explored alternative pathways if the MCAS requirement was not met by the end of the senior year, created agreements with local community colleges to serve students not successful on the MCAS exam, developed a program of community involvement, and developed a portfolio for students entering their senior year who have not been successful in a minimum of two retest opportunities.

Appointed four community contact workers (and one interchurch council member) representing the three cultures serviced by the NBHS bilingual program to increase participation and attendance. Created agreements with businesses for student employment while students worked towards competency determination.



Center on Education Policy
1001 Connecticut Avenue, NW, Suite 522
Washington, D.C. 20036