

**Detours on the Road to Reform:
When Standards Take a Back Seat to Testing**



July 2001

Lisa A. Banicky, Ph.D.
Educational Researcher for
Policy Analysis

Audrey J. Noble, Ph.D.
Director

With data collection and analysis contributions from

Andy Augustine, M.S.
Educator in Residence

Robert Hampel, Ph.D.
Faculty Affiliate

Yaron Siach-Bar, Ed. D.
Educational Researcher

**Delaware Education Research & Development Center
University of Delaware
Newark, DE 19716**

Publication T01.022.2

ACKNOWLEDGEMENTS

This study was a collaborative effort of various members of the Accountability Research Team of the Delaware Education Research and Development Center. The authors extend special thanks to :

Ann Case

Beverly Corelle

Eileen Malone

Barbara Threlfall

This research was made possible through the support of the Delaware State Board of Education.

Table of Contents

EXECUTIVE SUMMARY

INTRODUCTION

WHERE ARE WE?

Instruction

Complexities of Teaching to the Test
One Size Fits Most

Curriculum

Assessment Driven Curriculum
Standardization of Curriculum
Why Homogenization not individualization?

Assessment

Data Useful for Administrative Purposes
Data Not Useful for Instructional Improvement
Why the Different Opinions?

Educator Authority

Centralized Decision-making and Educator Disempowerment
Culture of Compliance

HOW DID WE GET HERE?

POTENTIAL COSTS OF THE DETOUR

Compromised Professionalism
Deskilling of Teachers
Decreased Morale
Exacerbating Educator Shortages
Prescriptive Professional Development
Continual Loss of Advocates for Reform
Limited Educational Experiences

WHERE DO WE GO FROM HERE?

Summary and Policy Considerations

APPENDIX A: Methods

APPENDIX B: Participant Check Responses

APPENDIX C: Interview and Focus Group Protocols

ENDNOTES

EXECUTIVE SUMMARY

Like many states in the nation, Delaware has been pursuing a standards-based reform agenda for nearly a decade. While the goal of standards-based reform was to create a balanced system of standards, assessments, and accountability a recent publication by Education week indicates that, in many states, accountability and testing often overshadow the standards within a systemⁱ.

The research that informs this report indicates that, in Delaware, the standards have indeed been overshadowed by accountability and that several “detours on the road to reform” have resulted in a system that is measurement-driven as opposed to standards-based. By placing our major findings in a context that includes comparisons to the ideals posited by standards-based reform more generally and Delaware’s New Directions initiative more specifically, we seek to illustrate how disconnects between the ideals and current realities manifest themselves in Delaware schools and classrooms.

For the purpose of producing a summary of the full report only the standards-based reform ideals are listed below. The full report includes a more complete discussion of Delaware’s interpretation of the reform as reflected in the New Directions plan. The realities presented below are the major findings to emerge from the case study.

The Ideal of Standards Based Reform and Today’s Reality

| System Components | The Ideal | Reality |
|-------------------|--|---|
| Instruction | Standards-based reform puts the focus on <i>what</i> students learn rather than <i>when</i> they learn it. This shift in focus should allow teachers to more easily accommodate various learning styles. It was hoped that standards would guide instructional practice and encourage teachers to use the most effective strategies. ⁱⁱ | Instruction is not focused on the standards but on teaching to the DSTP. Instruction is becoming less deliberate, less individualized, and more homogenized. |

ⁱ Quality Counts 2001: A Better Balance—Standards, Tests and the Tools to Succeed. <http://www.edweek.org/sreports/qc01> (1 July 2001)

ⁱⁱ Mid-continent Research for Education and Learning (2000). Noteworthy perspectives on implementing standards-based education. <http://www.mcrel.org/products/standards/implement.pdf> (20 June 2001)

| System Components | The Ideal | Reality |
|--------------------|--|---|
| Curriculum | Standards-based reform would foster “an alternative vision of curriculum and raised expectations for student learning by defining what students should know and be able to do within that curriculum” . ⁱⁱⁱ | Curriculum changes are the most often cited areas of change in response to the student accountability plan. But like instruction, curriculum is less focused on student needs and is more likely to be driven by the state test. |
| Assessment | Assessment should provide feedback about student performance relative to standards that will promote continuous improvement of teaching and learning. | The value of the DSTP data varies with the needs of the educator. For principals, the data provide direction in their decision-making as educational leaders. Teachers indicate that the DSTP data are of little value to them in their efforts to improve instruction. |
| Educator Authority | Standards-based reform would provide “top-down support for bottom-up reforms in teaching, curriculum, assessment, and professional development” . ^{iv} | Measurement-driven reform in Delaware is creating a culture of compliance among teachers. Decision-making power continues to move further from the classroom and school. |

With each of these realities are hidden costs to the system and to the individuals within the system (i.e. teachers, students, stakeholders) that might not be readily apparent in the short-term but may be more likely to emerge in the long-term. Some of these “costs” could include: compromised professionalism, deskilling of teachers, decreased morale, exacerbated educator shortages, prescribed professional development, continual loss of advocates for reform, and limited educational experiences for students.

ⁱⁱⁱ Wheelock, A. (1995). Standards-Based Reform: What Does It Mean for the Middle Grades? (ERIC Document Reproduction Services No. 389 772)

^{iv} Smith, M., & O’Day, J. (1991). Putting the Pieces Together: Systemic School Reform. CPRE Policy Briefs. CS: Consortium for Policy Research in Education, New Brunswick, NJ.

Introduction

In 1999 the Delaware State Board of Education initiated several studies designed to monitor and evaluate the impact of Delaware's student accountability plan on students and schools in Delaware. In response to this effort we have proposed and are continuing to conduct several studies^v guided by the work of the National Research Council as reported in the book High-stakes testing for tracking, promotion and graduation. In our investigations of the accountability plan we have been particularly vigilant for unintended consequences that might occur with the implementation and operation of a high-stakes accountability system.

The first year of the case study focused on changes in schools that resulted from the introduction of the student accountability system. We chose to frame the major findings from the first year of study in light of the original intentions of the accountability plan as articulated by the people involved in its creation.¹

Our goal in this the second year of our case study, was to extend the findings of last year's study by taking a classroom and school-level view of the changes occurring as a result of the student accountability plan. We also chose to place our findings in a more extended and enriched context by examining the findings in light of the ideals posited by standards-based reform more generally and Delaware's New Directions initiative more specifically.

The Ideal of Standards-Based Reform and Today's Reality

Standards-based reform began to take shape in the mid to late 1980s in response to the perceived decline of American schools. A balanced system of standards, assessments, and accountability was championed as a tool for improving schools with an emphasis on high standards for all students. Calls were made for clearly articulated and challenging expectations for student learning and Delaware answered this call by developing a plan for reform referred to as New Directions. The collaborative effort of Delaware educators, employers, state officials, and business leaders was defined as,

“ a vision for education in Delaware in which all children are able to reach their full potential and are prepared to lead full and productive lives as citizens and workers in the 21st Century.”²

Delaware, like most states in the nation, has been pursuing a standards-based reform agenda for nearly a decade. In the fifth edition of their annual report entitled *Quality Counts 2001*, Education Week examined the progress states have made in the implementation of standards, assessments, and accountability. According to the authors

^v The overall research plan involves three different studies. The first, a policy study, examined the original intentions of the accountability plan as expressed by policymakers. The second, a statewide longitudinal study, is examining the impact of the plan on two cohorts of students followed over a period of 5 years. The third is a three-year case study designed to explore what is changing in schools over time. This report is the second year of reporting on the case study.

of that report “state tests are overshadowing the standards they were designed to measure and could encourage undesirable practices in schools”.

We believe that the challenge facing other states, that of finding a better balance between standards, assessments, and accountability, is the same challenge currently facing Delaware. The research that informs this report indicates that the standards have indeed been overshadowed by accountability, that on the road to reform several detours have resulted in a system that is measurement-driven as opposed to standards-based. Our study reveals some of the disconnects between the assumptions of standards-based reform and the current realities of measurement-driven reform and how some of those disconnects manifest themselves in Delaware schools and classrooms. We believe that it may be time to take stock of where we are with respect to reform, how we got where we are, and what if any steps should be taken to change the current course.

Where Are We?

Given that standards-based reform was intended to provide learning goals to students and useful measures that could be used to align various components of the education system (i.e. instruction, curriculum, assessment, etc.),³ we chose to organize the major findings around these components. Our goal in the following sections is to highlight the intended destination, the ideals of the reform as expressed by writing on the standards-based movement as well as the early work on New Directions. The intended destination is then contrasted with our current location with respect to the areas of instruction, curriculum, assessment, and educator authority.



Instruction

The Ideal: Standards-based reform puts the focus on *what* students learn rather than *when* they learn it. This shift in focus should allow teachers to more easily accommodate various learning styles. It was hoped that standards would guide instructional practice and encourage teachers to use the most effective strategies.⁴

Delaware’s Interpretation of the Ideal: New Directions “is about using the most effective teaching practices so that all students can achieve to higher levels...What matters is that they (teachers) be effective, that they engage and challenge students.”... “Achievement against the standards will be the continual unit of progress; time spent on a subject will become the variable.”⁵ New Directions will promote the best teaching practices that will help students achieve the new standards.”⁶

Reality

◆ Instruction is not focused on the standards but on teaching to the DSTP.

◆ Instruction is becoming less deliberate, less individualized, and more homogenized.

Rod Paige, the U.S. Secretary of Education had the following to say about the role of testing in education:

“anyone who opposes annual testing of children is an apologist for a broken system of education that dismisses certain children and classes of children as unteachable. The time has come for an end to the excuses, for the sake of the system and the children trapped inside. Both the system and the children need reform...The centerpiece of the president's ‘No Child Left Behind’ plan is a system of high standards, annual testing against those standards of every child, and a system of accountability that makes schools responsible for results... Those who say this will result in a system in which teachers simply teach to the test don't understand the plan. A good test—the kind the president and I support—is aligned with the curriculum so the schools know whether children are actually learning the material that the states have decided a child should know. In such an aligned system, testing is a part of teaching... The time has come for meaningful change across the country, and that means annual assessment and accountability for every child in every school. The opponents of testing would have us cling to the status quo, and I have yet to hear an argument in favor of that.”⁷

Finding ourselves among those who obviously do not fully ‘understand the plan’, we argue that the current program of measurement-driven reform supported by the White House and evident in the state of Delaware is not what was intended when standards-based reform was originally conceived. Furthermore, we believe that teaching to a test, regardless of the quality or the form of that test, is not necessarily good practice. Moreover, teaching to the test is not a singular practice—it involves modifications that range from direct test taking skill development to overall program change.

Complexities of Teaching to the Test

In Delaware we have found that teaching to the test in a high stakes accountability environment triggers a range of instructional changes and distortions not limited to the simplistic perspective forwarded by the U.S. Secretary of Education's Office. Moreover, we find that teachers are not convinced that the changes prompted by high-stakes assessment are for the better.

“Across the whole country we’re trying to improve test scores, and not necessarily improving their learning.”

“This test has totally ripped apart education. It’s not a learning environment anymore.”

According to Quality Counts 2001, 67% of teachers polled felt their teaching had become “somewhat” or “far too much” focused on state tests.

“We’re teaching to a specific test. We’re not teaching for skills. We’re not teaching these children to be good learners... We’re teaching them to take that test.”

We have found that classroom instruction is becoming more test-like in nature and more focused on improving students' scores on the state test. Also, what many call “teaching to the test” proved not to be a single practice but was multifaceted in character. It also varied across a range of activities. At its lowest level, direct testing preparation activities are designed to help students be better at taking tests. Teachers functioning at this level of test preparation tend to spend time on test-related activities designed to increase students' familiarity with *test-taking strategies*.

“One of the things we have to do is teach how to take a test, which we really haven't had to do before.”

“I see that, in a way, as a good thing. Because they're going to be taking standardized tests for the rest of their lives. And they need to know, if you don't know a question, you skip and go on. You go back to it later. If you're not sure you can eliminate A or B, well, it could be C or D. Can you make a guess? I mean, that, to me, is really important.”

At the next level teachers seem to structure their teaching in ways that *replicate the format of the questions* as they appear on the state test.

“There is not too much surprise about what's on the test, just generally we can know what to expect and how the questions are to be asked. And you feel like you need to teach them.”

“We spend our professional development trying to structure our reading program questions like the DSTP.”

“You can't just teach reading. You have to teach specific ways that they're going to be given examples. Certain ways that they're going to be tested on a test. It's not like we can just go, ‘okay you know, let's explore creativity.’ They have to have a certain structure that we need to follow.”

Some teachers are teaching to the test by including different items in their classroom-based assessments that are similar to items on the DSTP. That is, they're focusing their instruction on *specific test items* that they've seen before on the state test.

“I look more at what was on the test than what they got. Like I know what was on the test, so I know what I want to cover based on that.”

“I look at the questions.”

“Obviously, you better know the questions that are on the tests. That is the guide.”

Some teachers see focusing their instruction on *topics* that appear on the state test as the best way to prepare the students. Not necessarily as a part of their regular classroom instruction or specifically linked to their schools' curriculum, teachers are addressing specific topics that they know have been covered on past DSTPs.

“We're teaching material that they're going to see on the test, in order for them to have the opportunity to do as well as they can.”

“There are many topics that I've given up that I would have loved to teach about with the kids. But I don't do it anymore. Why? It's not on the test.”

At the next level teachers are not only altering the topics that they are teaching, they're more likely to make *changes in overall content*. In their efforts to prepare the students for the state test and due to the time constraints of the school year, they are making choices about which content is most appropriate, that is, would be most helpful in helping the students score better on the state test.

“We took out science and social studies for two months. Not totally, but basically they told us to replace it with math and reading.”

“You can't really teach poetry, because it's not on the tests, so why waste a month teaching poetry before the state test.”

“I just went to this National Council of Teachers of Mathematics conference. I learned a lot of great ideas. ‘Hey you can try this with your class. You can try this project. (But) we can't do that. I'm looking at my watch and saying ‘hey it's January already and I have to get through a lot more information.’”

Quality Counts 2001 reports that 66% of teachers felt they were concentrating on tested information to the detriment of other important areas of learning.

Teachers are not only adjusting their content focus, they are also *changing the instructional strategies* that they use in the classroom. And not all of these changes are for the better. Some instructional models that they have used in the past as well as some that are supported as “best practice” are being sacrificed in the name of preparing students for the test. Some teachers feel that they do not have enough time to engage in the range of instructional strategies that they've used before because of the pressures that they're feeling. Others, not convinced that recently adopted instructional strategies will yield good test scores, are falling back to old methods of instruction.

“We used to do centers. When we go to do free choice in math, that's the first thing to go. And that's the first thing that needs to be there for mathematical understanding and fluency, is for the children to have free time playing to explore numbers. And when we hit free choice in TERC we go, 'oh, no, no more center time. We're going to skip that because it's no more time for them to play.’”

“Teaching by themes. And now that's kind of gone away.”

“We no longer do whole language.”

Teaching to the test has brought about complete *program changes* in some schools. Programs and activities that do not seem to be directly connected to helping students perform better on the state test are being sacrificed to make way for those that seem more directly connected. Also, personal development activities that are not seen as directly contributing to students' skills in reading or mathematics are becoming less and less important.

“Far fewer field trips. Obviously you can't be in two places at one time. So before, where we went out and observed and did things, now if we do any, we need to do them after the test. But they're still far fewer.”

“Assemblies are basically nonexistent. We used to enjoy all the day concerts. And Christmas time, no, that's gone.”

“I have a basic exploratory reading class...Not every 8th grader has to do it. Some of them may take Spanish or French.

The overall goal of all these teaching to the test strategies is to improve students' test scores. The teachers and principals readily admit that they are not teaching to the standards nor are they preparing students for the next grade. Their goal, whether they are pleased with it or not, is to improve the students' scores on the DSTP.

“It all comes down to how well kids do on the test.”

“If the teachers are so focused and so worried about accountability, I question how much learning will take place. Or will this just be a place where we look at the Delaware state test results and try to prepare kids for the test.”

“I mean the whole thing about the state test is that it's taking the focus of what we should be doing, which is educating the children and it places it on force feeding the state test material.”

“I think we're all kidding ourselves if we think a standards-based curriculum is going to get a kid to pass a math test. Because that won't do it. You need to know what's on the test and your whole curriculum, the whole unit has to be encompassing. But it has to cover the material that's on the test. Obviously.”

“My goal used to be to get them ready so they can be successful at high school. Now my goal is to get them ready so they can be successful on the DSTP.”

This range of teaching-to-the test activities clearly demonstrates that teaching to the test is not as straightforward or simple as some at the national level might like us to believe. The all too often used phrase, ‘you get what you assess’, falters in light of the complexity that has been revealed here. What you get when you assess is not so obvious. Considering the dimensions of teaching-to-the-test activities, one might argue that some outcomes are more acceptable than others. Perhaps some would consider programmatic or instructional changes that are made to help better prepare students as improvements. Perhaps it is better that content is becoming more focused. Perhaps some would define the use of test items in instruction as a violation of test security. Some might claim that the loss of creativity and the lack of focus on personal development of children come at a price that is too high to pay. Nonetheless, it is clear that teaching to the test is not necessarily a simple practice nor one that should be benignly accepted. The U.S. Secretary of Education claims that “in such an aligned system, testing is a part of teaching” True, testing is a part of teaching; but this research and other studies that have been conducted before clearly indicate that when high-stakes testing becomes a part of teaching, it usually distorts it.⁸

One Size Fits Most

The distortions in teaching that we have observed so far indicate that instruction is becoming less deliberate, less individualized, and more homogenized. We believe that time plays a role, but is not the sole factor, in creating conditions that are conducive to such distortions. The ideal of standards-based reform was to place the focus on what students know and not when they know it. Consequently, testing at set points in time, i.e., specific grades levels, and consequences based on those performances, is creating a great deal of pressure within the system.

“There is simply so much more to do...and so much less time to prepare for it.”

This statement could have been made by any number of people-- staff at the Department of Education, a member of the Curriculum Cadre, someone at DSEA, a school district superintendent, or a teacher. What we have found is that teachers are finding ways to deal with the pressures of too much to do and too little time to do it. As would be expected, the sheer weight of the new demands being made of teachers has begun to overwhelm even the most committed among them. It is clear that many do not feel that

they have enough time to cover all of the information students are expected to know for the DSTP.

Quality Counts 2001 revealed that 7 out of 10 teachers felt there was not enough time to cover everything in their state standards.

“But where the problem and the stress comes in is, I want to expose them to every different type of writing that I possibly can before March...I’m buried.”

“There’s almost too many balls to watch right now, too much up in the air.”

One of the effects of this time pressure is that instruction, particularly in the weeks before the testing, becomes less deliberate. In their desire to help their children do as well as possible on the DSTP, many teachers have begun “throwing” everything they can at their students in the hope that something will have a positive effect on their test scores. In their words, they are hoping something “sticks.”

“I feel like I need to expose them to as much as possible...and hope that some of it sticks with some of them.”

“We haven’t gotten there in our process. But I know that’s on the test so I have to hurry up and cram that in. And it doesn’t do them any good. A month down the road they won’t remember.”

“I taught some real slipshod lessons to try to cram some real simple basics in those last few days.”

“Right before the test, I was just throwing things at these kids.”

“It’s almost like you’re just shoving it at the kids and hoping that they consume some of it because you have to pack in so much.”

In addition, instruction is becoming less and less individualized in many classrooms. Educators are aware of the various instructional needs of students, but because of the pressures of getting the students ready for the test, individualization of instruction has become a luxury, not a necessity.

“I think it has changed me a bit because I’m a person that believes a child grows over time. And I can’t teach them everything in one year. So I would look at their individual growth...And now it’s, ‘regardless of who you are, you’re going to get this. I’m going to keep right on plugging through at it because we have to get through this.’”

“Instead of having five levels and giving them some of the things they truly need, we’ve gotten rid of our below grade level remedial groups and they’re put in regular pre-algebra. And in some cases, the kids swim, are surviving. But in many cases they’re going under.”

“I truly regret being forced to teach all, or treat all students the same, when kids aren’t the same. Everybody can learn but not everybody can learn at the same rate.”

“Now I’m basically doing the same lesson plan with everybody. I never did that...I was trying to meet them where they were...I would treat them all differently. Now each class gets the same books, whether you can read it or not. Everybody reads the same stories whether you can understand them or not.”

The exception is the special education classroom where the penalties for not achieving on the DSTP are not seen as so damaging.

“In special ed our kids get promoted regardless of how they do on the state test. Not that we don’t want them to do well but I don’t have the same accountability issue with kids.”

“Especially for resource level kids who can’t quite do what the regular ed kids can do...And they need to know when they take a test of this nature, it’s not going to prevent them from graduating.”

One of the criticisms that standards-based reform was designed to address was that the educational system failed to meet the needs of all students. Subsequently, the standards-based system promised to hold high expectations of all students. Recognizing that the needs of students differ because of their abilities and prior knowledge, the new standards-based classrooms would require teachers to incorporate a variety of teaching strategies to meet the children’s specific learning needs. But in an environment pressurized by high-stakes testing that includes fixed time and performance expectations, we found the opposite to be occurring. Instruction has become less individualized and more homogenized.

Curriculum

The Ideal: Standards-based reform would foster “an alternative vision of curriculum and raised expectations for student learning by defining what students should know and be able to do within that curriculum”.⁹

Delaware’s Interpretation of the Ideal: New Directions will add a more rigorous curriculum with high expectations for all students to succeed; classroom exercises that actively engage children in learning...a variety of teaching strategies to appeal to the various ways that children learn.”¹⁰

Reality: Curriculum changes are the most often cited areas of change in response to the student accountability plan. But like instruction, curriculum is less focused on student needs and is more likely to be driven by the state test.

Assessment-driven Curriculum

Teachers and principals believe that the major change caused by the reform is in the curriculum. However, many state that the curriculum is more aligned with the test than it is with the standards. Apparent in their comments is their lack of trust in the alignment of the test with the standards, and consequently, they feel that it is important to be sure that their curriculum aligns with the end product, the test.

“We put in a new Language Arts program that is standards-based...and it closely aligns with the state test.”

“When we looked at all the different series that were available, we looked for the one that most closely aligned with what we’re asking our students to do on the DSTP.”

“You take that test, and unfortunately, that is your curriculum.”

“And the things that we’re going to use from this Math in Context are things that mirror the DSTP...We went through it and I see a part that I know is exactly from the DSTP and I say “hey this book’s good. There are some questions in here that are just like the DSTP.”

“the only negative point...is that we threw the curriculum together before we were told the students were going to held accountable for the state test. So now we’re saying ‘wait a minute’, Do we need to go back and redo it again?”

Standardization of Curriculum

In addition, as the teachers become more focused on the DSTP, the curriculum being used across classrooms is becoming more comparable. In many districts, curriculum models and textbook series are being mandated.

“We didn’t have math curriculum when I started 4 years ago. And (now) everything is you either use TERC or you use TRAILBLAZERS.”

“They [district] said, ‘This is what you must teach.’”

“It’s a state curriculum now, and pretty soon it will be a national curriculum.”

“we’re setting up this district-wide curriculum. This is what you have to do. Everyone has to follow the same curriculum.”

At the elementary level teachers tended to describe the curriculum as becoming more lockstep in nature.

“We went from no curriculum to firmly-guided, are-we-all-on-the-same page curriculum.”

“It looks like the same thing every week. Monday and Tuesday, it’s reading the story and introducing vocabulary...It’s something I’ve never done in my life until this year.”

Why homogenization not individualization?

We found evidence in all of the schools studied that curriculum and instruction are becoming increasingly more homogenous, that is, teachers are teaching the same thing in the same ways to most students. Most teachers were critical of the distortions in their teaching but felt powerless to do anything else.

We believe that this homogenization of curriculum and instruction is not simply a path of least resistance. We believe that there are numerous factors at work that have contributed to this:

- First, in a system that is measurement-driven, expectations of student performance (a “3” or better) and time (grades 3, 5, 8, and 10) are set, not variable. The system creates a pressure to prepare all of the students, regardless of their capacity and prior knowledge, to meet the same benchmarks within the same amount of time. Consequently, this causes teachers to be wary of taking too much time for any one child, minimizing their willingness to individualize instruction at the expense of the other children.
- Second, curriculum control at the district and/or state level forces a limited focus and restricts access to off-grade materials that would enable and encourage teachers to individualize their instruction.
- Third, educators’ concentrate on the test and not on the individual. As a result their energies are focused on how students perform on the test and what needs to be done to get the ‘2s’ to ‘3s’. Subsequently this minimizes their sensitivities to students’ diversity and individual needs.
- Fourth, the broad categorical nature of the DSTP data may be making it even more difficult for teachers to individualize instruction. The data that teachers receive from the DSTP are not diagnostic in nature. Rather they present a broad-brush view of how student groups perform. The data speak to group not individual needs and achievements.

Assessment

The Ideal: Assessment should provide feedback about student performance relative to standards that will promote the continuous improvement of teaching and learning.

Delaware’s Interpretation of the Ideal: The Comprehensive Assessment System should focus on the improvement of teaching and learning; help the public (students, teachers, parents, general public) to examine the content and performance standards...continue the opportunity to engage in the revision of standards, always focusing on continuous improvement,...”¹¹

Reality: The value of DSTP data varies with the needs of the educator. For principals, the data provide direction in their decision-making as educational leaders. However, teachers indicate that the DSTP data are of little value to them in their efforts to improve instruction.

One of the original intentions of Delaware’s student accountability plan, as articulated by the original designers, was to create a system for measuring student performance against the state content standards that would focus on continuous improvement. The assumption was that the data generated by such a system would play a critical role in decision making at the school level and within the individual classrooms.

Data Useful for Administrative Purposes

When we spoke to principals and teachers about the utility of the DSTP data for various purposes we found what initially appeared to be a disconnect between the response of principals and teachers. Principals reported that the data were useful for informing a variety of areas including curriculum, content coverage, professional development, school-level planning, and teacher evaluations. Principals also believed that teachers found the data informative as well.

“We’re looking at the data that our students across the districts are producing, so that we can make sure we’re covering everything that needs to be covered.”

“Find it useful so that it motivates them to make professional decisions that they have to make. I mean, it’s data driven, but it’s not data dominated. We take the data, and then on that we use our professional training and insights to see what we can do.”

“Our writing scores continue to be a disappointment to us so we are focusing very heavily on the area of writing this year, more so than we did the past two years.”

“...teachers, I see looking at their data, and using their data to improve their instruction...”

Principals further indicated that the more useful part of the data generated by the DSTP were the needs reports associated with student performance. Instructional needs reports are comments generated on the basis of the number and types of items students answered incorrectly on the DSTP. Much of the conversation around the data focuses on these reports.

“I said ‘Look at the comments that are most frequently...’ and the English Department, ‘Look at your comments that are happening most often’...when you are instructing, obviously you want to tell the kids, you have to do this, this, and this.”

“I do share teacher results specific to that teacher, so they know how their kids did, but also I can talk about it with them, ‘Well, globally it looks like your kids all did not do well in this, say geometry and math. So we need to look at how you’re approaching geometry or spatial, or with your writing’...So that globally... our team looks at where we are weak...within our building.”

Data Not Useful for Instructional Improvement for Individual Students

Unlike the principals we interviewed, teachers did not find the DSTP data useful. Most teachers indicated that knowing what the specific items were on the DSTP was more informative than students’ scores or the comments on students’ instructional needs report.

“It’s numbers on paper.”

“If somebody thinks that we change... based on data collected from that, that’s totally wrong...I think we all update what we’re teaching based upon what we see on the test.”

“I look more at what was on the test, than what they got. Like I know what was on the test, so I know what I want to cover based on that, as opposed to how they score.”

Teachers cite several reasons for the limited utility of the DSTP data. Those that were most frequently mentioned included the timing of the return of the results and the vague feedback generated by the instructional needs report.

“We don’t get them in enough time to make them useful in the classroom. We don’t get them until after schools starts.”

“But then, even then the results that you do get, the kids are gone. So I have a new set of students this year, so last year’s scores won’t really...I don’t base what I teach in language arts on last year’s scores...”

“I don’t think the data is detailed enough to give truly useful feedback to the instructor.”

“...when we did get those results...they were hard to interpret...and I’m saying ‘Okay, well where does that mean that they’re lacking?’ I’m saying, ‘What does this mean?’ It was such a vague term.”

Why the Different Opinions?

The discrepancy between principals and teachers in their perceptions of the utility of the DSTP data can be explained by their respective roles as educators. For principals, the data provide direction in their decision-making as educational leaders. Principals perceive the data as specific enough to address issues related to curriculum, professional development, and teacher evaluations. Teachers do not find the data timely or specific enough to allow them to individualize or improve instruction. Also, teachers typically want data to serve diagnostic purposes, that is, they want the data to help them better understand why a particular student is not performing well and help them better understand how to remedy the learning problems of individual students in their classes. The DSTP was not designed for this purpose, and subsequently, does not provide data at this level. As a result, teachers are less likely to view the data as helpful as they do not meet teachers’ needs in this way.

Educator Authority

The Ideal: Standards-based reform would provide “top-down support for bottom-up reforms in teaching, curriculum, assessment, and professional development.”¹²

Delaware’s Interpretation of the Ideal: New Directions: Three themes, i.e., defining content and performance standards, assessing student outcomes and restructuring schools, emerge as the central foci of Delaware’s educational reform vision.... The third, how to organize and operate schools or restructuring “can be best achieved in public education by empowering those professionals closest to the classrooms (i.e., teachers, principals, superintendents, and local boards) with decision-making authority and responsibility.”¹³

Reality: Measurement-driven reform in Delaware is creating a culture of compliance among teachers. Decision-making power continues to move further from the classroom and school.

One of the tenets of standards-based reform was that with increased responsibility for student outcomes came increased authority over the decisions made at the school-level. Providing authority to school personnel would provide those responsible with more leeway to find creative ways to improve learning while at the same time recognizing them as professionals. Our research finds that internal and external pressures resulting from the accountability system are creating a culture of compliance among teachers. At the same time teachers report that decision-making power continues to move further from their reach.

Centralized Decision-making and Educator Disempowerment

The pressure reported by teachers comes from a variety of sources, some internal to the teachers and others from external sources. In the instances cited earlier the pressure appears to result from the expectation based on the timing of the test. In other instances, mandates about what teachers are expected to do come from above.

“Years ago we used to be able to pilot programs, and curriculums, and decide what ...for all students, and what curriculum we could use...now it is district mandated curriculum saying that we must teach on grade level, this specific curriculum, because we want everybody teaching the same thing.”

“not every child learns the same way, and most of us still supplement from other places, even though they say ‘Oh, you shouldn’t, this is all you need’. But we know that is not true...they are telling us to do this, but we know we have to use our own sense to make it work”

Some teachers interpret their diminished involvement in decision-making as an indication that they are not trusted to make decisions about students or that their professional judgment is not important in regards to teaching and learning.

“We are not being trusted to teach these children.”

“We are not being given the credit for being professionals”

“The district, some time during the summer, chose a series. Once they had chosen that series, people came into our classrooms during the summer and removed every reading book, manual, everything that had anything to do with any series we’d ever taught before. Because they didn’t want us to be dependent on those and go back to them, because they were easier. And everybody was told, everybody must teach this particular program, at grade level.”

The feeling of powerlessness that teachers report may also stem from issues related to communication or the perceived lack thereof. Teachers expressed concern over the lack of communication between and among the school, district, and state.

“So there’s this disjointed sense that we’re communicating here in our little spot, but...it isn’t being heard or it isn’t being looked upon seriously”

“Everything that we’re doing, I believe, is some sort of way...knee-jerking to address what’s being handed down by the state. And I think that we, as a result, see, there’s a complete lack of positive communication between teachers and the district administration. I think the teachers feel rather alienated.”

“the messages are always so ambiguous, of what the test is going to mean, and when it’s going to mean something.”

“we were given a four point rubric to use to prepare students, and the state came back with scores on a five point rubric. We didn’t know there was a five. Nothing they ever gave us indicated there was a five...and the kids had the results before the teachers even knew that there was an additional level there.”

Culture of Compliance

When faced with these internal and external pressures we found that the predominant response of teachers was compliance. Much of the language teachers used to describe the changes brought on by the student accountability plan reflected a sense of powerlessness and coercion. Teachers used words like “forced,” “railroaded,” “have to,” “can’t,” and “hammered” to describe their experiences under the current system. Some of the comments most reflective of this compliant attitude included:

“Teachers will do anything that you tell them to do. Isn’t that sad?”

“I had to go to another thing, they sent me. No questions asked.”

“They don’t allow us...we were told you can’t do it.”

“Well this is what the state has mandated. This is what you have, do it.”

“This year they threw in a whole new reading series, we had to follow.”

Teachers may also be responding with compliance because they feel overwhelmed and powerless to change what is occurring. This powerlessness may stem from teachers’ perceptions that they have little say in decisions that are being made at their school.

Even more disturbing to them is that they feel that their professional judgment is not valued.

“I’m not saying that they’re right or wrong, I’m just saying that there’s not a lot of input by teachers”

“But we are not having a say in it. We are the ones that are assessing the children, and deciding what is good for them. But as a teacher my voice is not being heard.”

“But then they’re not allowing you to be the judge of your room, and the students in your room.”

The idea that teachers are no longer the judge of student work was recently, albeit inadvertently, illustrated by an article that appeared in the Delaware News Journal. The article described the manner in which DSTP scores would be used to make summer school decisions but the headline read: **State Test Chooses Summer School Students.** Teachers spoke of their diminished decision-making role because of the state test and its influence on student’s attitudes.

“It makes teachers grades...and these are things they said in class, when we’ve talked about this...it makes grades obsolete if one test is going to keep me from passing the 8th grade...”

“...the big thing I heard {from students}, when we first finished with the tests, when we finally finished with them, was, ‘Nothing else is important...we’re going to coast for the year’”

{students say} *“Hey, the DSTP is going to decide whether I pass or fail, why bother doing the work?”*

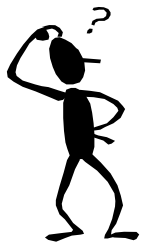
Research in the area of compliance indicates that people sometimes comply because they fear the negative social consequences of appearing deviant.¹⁴ Comments like those expressed by Rod Paige exemplifies this notion as he accuses those opposed to testing as being in favor of mediocrity and the status quo. Teachers may be responding to the pressures of accountability with compliance because they fear that resistance signals an unwillingness to be held accountable or to hold high standards for students.

“You would be a louse if you didn’t go along with them. You know. There’s a little bit of that, I think. I mean, I listen to the national news. And I think this is...this is the trend across most states.”

On the surface, compliance may appear to result in the desired changes if improved test scores is the ultimate goal. But if the goal is to make authentic changes that improve student learning, then compliance may be a less than desirable response. At a most basic level, compliance involves changes in behavior that result from a direct request.¹⁵ But

classic research in the area of social influence indicates that compliance often involves superficial changes in behavior in which dictates are followed but not personally accepted.¹⁶ Consequently, compliant behavior will likely be short-lived and disingenuous.

Since the beginning of Delaware's reform effort, teachers have supported the standards-based reform movement. To speak out against the reform now would seem to be a compromise of their earlier commitment to the reform. Research in the area of social influence indicates that most people desire to see themselves as consistent,¹⁷ therefore teachers may feel a need to comply because it verifies their initial support of standards based reform.



How did we get here?

Given the old ideals and the current realities of the reform initiative in Delaware, how did we get here? Delaware educators have no argument with standards-based reform. That is not at all surprising considering the efforts made by the state in the early years to foster buy-in. The initial processes were highly participatory and involved every constituency group in the partnership. While educators are typically critical of educational reform initiatives, seeing most as passing fads, most educators believe standards-based reform is here to stay. With the focus of standards-based reform on the genuine and continuous improvement of curriculum and instruction, how could educators argue? And they didn't; they strongly supported and continue to support the notion. However, their arguments surfaced when discussions of accountability came to light. More specifically, when accountability was equated with high-stakes testing, discontent blossomed.

So how did this change from standards-based to measurement-driven reform come about? In the late 1990s, the primary emphasis shifted from improving curriculum and instruction by focusing on standards to holding students and educators accountable by focusing on student test scores and rewarding/sanctioning educators and students for these scores. We provide some illustrations of what we see as key events that may have brought about this change. This is not intended as a complete explanation but rather an exploration of what happened that shifted the focus.

Change in the Form of the State Testing Program

New Directions called for standards and a comprehensive assessment system¹⁸:

- one portion of Comprehensive Assessment System was to be an 'on-demand' statewide assessment program to be administered annually in grades 3, 5, 8, and 10 in reading, writing, and mathematics "in order to institute school-level accountability in the core essential skill areas;

- the other portion was to be an ‘embedded statewide common core of learning performance assessment program’ that was to cover all grade levels and incorporate a broader range of subject areas (e.g., science, social studies, ELA, and math) and was to be incorporated into the ongoing classroom teaching.

The ‘on demand’ component of the Comprehensive Assessment System was primarily intended to serve system and accountability needs; the embedded component was to serve instructional improvement needs. When the expansive size and the related costs of the entire system were fully realized, the embedded portion of the CAS was abandoned and the ‘on-demand’ assessment remained. What was left was expected to serve both purposes, even though it was designed for school-level accountability not instructional improvement. In 1997, the change of form was validated by the change of the assessment program’s name, from the Comprehensive Assessment System to the Delaware State Testing program. However, the buy-in that was established among most educators was based on the assessment’s capacity for instructional improvement. And in many cases, this continues to be the case even though many teachers find little value in the DSTP for improving their teaching.

Change in the Function of the State Testing Program

After the standards had been written and the development of the DSTP was well underway, discussion among policymakers turned to accountability. As deliberations progressed, the role that the Delaware State Testing Program was to play within the accountability system changed. Initial expectations were that it would serve as one of many indicators in a system whose primary intent was to improve teaching and learning; later, it came to be seen as the sole indicator in a system that had accountability as its primary purpose with instructional improvement as secondary. The shift of DSTP’s primary emphasis and role is clearly evidenced here.

“Given the great public interest in education and the need to ensure that the accountability system and the measures within that system are an accurate portrayal of performance, a view of accountability is being advocated that extends beyond traditional definitions tied solely to single point in time achievement data. This is critical. DSTP has been designed to support the improvement of teaching and learning....”¹⁹

“Originally, our primary focus was to develop a test, unique to Delaware, that accurately measures our students’ mastery of the Standards and curriculum effectiveness. When we began our educational reform effort five years ago we designed our path forward based on certain facts that are no longer true today... *Five years ago our primary concern in designing the state testing system was to improve instruction and not as the foundation of a student accountability system. Today we know that the test must be designed to support its use for accountability at several levels, which will also serve to drive improvements in instruction. One*

factor remains constant, resources are limited.”²⁰

- *Development of the Performance Indicators*

Performance indicators were never a part of the original standards-based reform plan in Delaware. The plan involved state level development of content and performance standards, assessing student outcomes, and restructuring schools and schooling. Local control of instruction was valued as districts and schools were to maintain responsibility for teaching, learning, and professional development. One of the New Directions goals was to “give local school districts the freedom to decide how their students reach the standards.”²¹ But after the standards had been written and the primary emphasis of the assessment began to shift from instructional improvement to accountability, districts began to feel discomfort about whether their curriculum was adequately aligned to the standards and whether this alignment would result in adequate test scores on the DSTP.

“(We could) adjust some of the emphasis on things in the curriculum. That’s what scares me about testing. You do that because that’s obviously what the state values. Then every interim they changed the test.”

They demanded and Steven Adamowski, then an associate superintendent in the Department of Education, gladly provided what was intended to be a clearer pathway between the standards and the DSTP--the performance indicators (PI’s). However, response in the schools was mixed as to their value. Some criticized the performance indicators as being too discrete and destroying the holistic nature of the standards. Some welcomed them as a means to focus the curriculum development efforts and the classroom lesson planning. Even now we find some principals who believe the PI’s should focus what goes on in the classroom.

“We look at their planning. Are they identifying specific parts of the content standards? Performance indicators? A lot of them like to use the performance indicators, to even become more specific. That is something we want them to do, and they do.”

Nonetheless, for some, the performance indicators did create a pathway between the standards and the DSTP. However, in conjunction with other pressures in the system to have educators focus on test scores, some might say that the PI’s provided a detour to that end--- at the expense of having educators focus on standards.

- *Numerous Changes in Leadership & Participation*

Since the early 1990’s when standards-based reform began in the state, there have been many changes in leadership and participation. To name a few---the leadership at the Department of Public Instruction/Department of Education changed many times; there was a major shift in education policy from the State Board of Education to the Delaware General Assembly; and, the role of the business community declined from its highly active position in the early years. Most evident is the change in the level of participation

among the education stakeholder groups. In the early 1990s, the New Directions partnership was highly participatory; it included all 19 school districts, the business community, government, higher education, community organizations and education associations, including the State Board of Education, the Delaware State Education association, the Delaware School Boards Association, the Delaware Association of School Administrators, and the Delaware Congress of Parents and Teachers. Today the group actively forwarding the reform initiative has reduced significantly; the responsibility for the reform now lies primarily with the Department of Education. Over the past decade, tensions have splintered the original partnership and now many members have retreated to their respective corners to protect the interests of their particular constituency.

The Potential Costs of the Detour on the Road to Reform

Some may believe that the changes described to this point are necessary for changing the status quo and improving the education system. Therefore, the ramifications of the change from a standards-based to a measurement-driven reform may not be significant. We believe, however, that teaching to the test, homogenizing instruction and curriculum, and fostering a culture of compliance comes with a large price tag. These costs could include but are not limited to:

§ Compromised Professionalism

Compromised professionalism occurs when educators make changes in their school or their classrooms even though they question the worth of those changes. We believe that teachers are struggling to continue to do the work they are prepared to do in an environment that has begun to compromise their professionalism.

“as teachers, we’ve changed a lot, not for the better, as I said, but we’re all, across the whole country, we’re teaching, we’re trying to improve test scores...and not necessarily improving their learning...”

There is a perception on the part of teachers that these practices (i.e. teaching to the test, homogenization of instruction) are necessary to meet the demands of the accountability system. Teachers report a great deal of pressure to engage in these practices even though doing so sometimes creates a conflict between what teachers think they should be doing and what they feel they must do.

“...some days...when your kids have had it because you’ve been hammering them with it, you say ‘Oh, let’s stop. Let’s do something fun.’ Uh-uh, I’ve got to keep going on through this, because I want them to understand it. And at the same time my mind’s telling me, you know, you really shouldn’t do this.”

“I’d like to get deeper involved into something where the kids might actually learn it better. They might be able to intrinsically learn this material. But I know I can only spend so much time”

“We have many students who read on a 3rd to 4th grade level. And yet I’m expected to teach them a curriculum written, gauged to a 7th grade reading level, and that’s the way it will be presented to them, at that level. So there are things that I would have spent more time on...providing a base of prior knowledge, and things like that. But in the interest of time, we felt that we had to get them through more, a greater part of the curriculum.”

Teachers and principals are struggling in this high-stakes testing culture to acknowledge and incorporate what they know about teaching and learning in their day-to-day activities. Even though they may have readily accepted and believe that “all children can learn”, they cannot deny that they believe that some students have more innate ability than others. They recognize that students’ prior knowledge along with their level of intelligence affects how well they will score on the DSTP.

“Everyone talks about raising the bar. But for some kids, you can raise the bar an inch and they’re still not going to be able to jump over that. I mean, you have to have some natural ability...Look at it literally. You can raise it to 7 foot and I’m not going to be able to jump over that.”

“I’ve often said, short of a brain transplant, there are many kids that will never, ever pass the test. And it’s genetic. They don’t have it. They never will. We’re not all going to be rocket scientists.”

“Kids are kids, but the way they come to us, they’ve been imprinted differently.”

These types of statements about students’ innate ability are not a popular stance for teachers who have committed themselves to a basic premise of standards-based reform, that is, ‘high standards for all’. But the concept of holding all students to high standards also triggers a conflict for some who try to reconcile it with what they as professionals know—that students do have different strengths and weaknesses, different talents and aptitudes, and different intelligence levels.

Another way that teachers are struggling to make sense of what is expected of them as it conflicts with what they know about learning surfaces in regard to the notion of developmentally-appropriate instruction. Teachers have been educated to understand and are expected to recognize the developmental needs of children and how those needs should be addressed instructionally. However, as curriculum is mandated and expectations of performance are static, teachers feel less and less able to “take kids where they are”.

“We have to do what’s right for the kids. And we’re missing that... And right is building their skills day to day. Not this test.”

“And we’re being told, take the child where they are at, and move them forward. But if you’re taking kids from here to there, and you’ve got to move them all up to here, it can’t be done.”

Other research concurs that the more pressure teachers felt to raise test scores, the lower their professional self-images.²² In addition, other researchers found that the dictates of externally-mandated tests reduced both teachers’ perceived levels of professional knowledge and status.²³

§ Deskillling of Teachers

The stress created in high-stakes testing systems has been shown to take its toll on teachers in various ways.²⁴ Teachers in Delaware are clearly experiencing stress in response to the state’s reform agenda.

“I feel like I have been stressed out by the, being told what I have to do as opposed to what I know I should be doing...”

“teachers would be feeling a little better about this, and maybe a little less stressed if they didn’t feel like they were being bashed all the time”

“And it’s not that I think, that people so much are resistant. They just, teachers are feeling terribly overwhelmed...”

§ Decreased Morale

Not only are teachers being deskilled they are also being demoralized by the measurement-driven reform. Teachers cited low morale as a by-product of the high-stakes system. Teachers also report that teaching is not as enjoyable as a result of the pressures associated with the accountability plan.

“I think if you’re any kind of a professional, that’s the way you’re going to feel. Its just the way...I’m certainly glad that I’m on my way out. I mean as much as I love teaching, I’ve got maybe six or seven more years. I don’t know if I could do this for 25 years, getting them ready for a test...I couldn’t have lasted as long as I did.”

“I know I can say, I think it took a lot of fun out of my teaching.”

“I remember looking at the older teachers and thinking, ‘Oh, they’re just jaded. They’re just tired.’ Well, I’m jaded and tired now.

“...we do feel the pressure of getting through the curriculum and so do the kids—it trickles down. It’s taking the enjoyment out of education all the way around, for the teachers and for the kids...”

§ Exacerbating Educator Shortages

Earlier research on the impact of high-stakes testing on teachers has clearly shown that it has the potential to drive out good teachers and ‘deskill’ those who remain. Good teachers either found a means to resist the deskilling process or left teaching.²⁵ This is a possible outcome in Delaware, and one that the state can ill-afford in light of the need for quality teachers in every classroom.

According to a recent report examining teacher supply and demand in the Mid-Atlantic region, there is a serious prospect of a teacher shortage in the near future. The Delaware Department of Education has reported critical shortages in the areas of bilingual education, foreign language, mathematics, science, technology education, and special education. In addition, five of Delaware’s 19 districts reported that they had to limit or discontinue student course offerings because of staffing shortages.²⁶ We found that many teachers and principals expressed concern that teacher and administrator attrition would increase as a result of many of the costs discussed above.

“They’re going to be facing an even bigger teacher shortage in three to five years when people like me will be leaving. More people are leaving. What teacher wants to go work in a school...if you had your choice, working in a school where kids are on level, and doing a good job, and you want to go to a really positive accredited school, why would anybody want to come to teach in a school that was below the standards—knowing all the pressure that is going to be on you to bring those kids up to standard?”

In addition, we found that teachers in one historically high performing school indicated a hesitancy to work with student teachers. According to the teachers, they did not want the added responsibility of overseeing a student teacher and felt that a student teacher would not be able to adequately prepare the students for the state test.

§ Prescriptive Professional Development

Pressure to have students perform well on the DSTP has begun to take its toll on the capacity of administrators to effectively support a full range of professional development. Even though the state has professional development opportunities, predominantly in mathematics, science, and writing, teachers are beginning to be less able to participate in what they might, as individuals, need or want to pursue. In addition, as educators, teachers and principals, realize that quality professional development is not achieved through one-shot workshops, their decisions become more complicated. In addition, principals see the need to have professional development “match the test”. As a result,

educators express concerns that professional development is becoming more narrow with potentially negative impact on areas.

"My fear is that math is going to end up taking a backseat at the moment, because I harp, and I know every administrator probably is harping about, 'you need to do something about writing.'"

As always, limited time and financial resources exacerbate the dilemma.

"I feel like I'm always robbing Peter to pay Paul, and I'm not sure what pot of money I'm going to find."

"We've tried occasionally to schedule professional development during a regular student day and provide substitutes. But we're always short on substitutes. It becomes a logistical nightmare."

"Inservice after school is not a good idea. Teachers are tired at the end of the day and they have other responsibilities that take away from their ability to attend."

As many principals try to make due, they utilize resources within their school to attempt to meet the varied needs of their teachers "because we don't have resources to bring in (experts), we've had our own staff train our own people."

It is uncertain what the ultimate cost of this narrowing of professional development may be. But what does appear clear is that with the current demands on the time and resources of educators at the school level, something will suffer. Extensive investments have been made to date by the state, especially in NSF-supported mathematics and science, as well as the Delaware Writing Project. As time and local funding become increasingly more restricted, competition among these professional development efforts could easily result.

§ Continual Loss of Advocates for Reform

It is evident that the level of collaboration in the early years of the reform has dwindled leaving the Department of Education with the primary responsibility for the reform. This has occurred for numerous reasons. Regardless, this poses a major dilemma that speaks beyond the level of capacity of the Department. It has already been seen that as implementation moves forward, the work gets more complex. Moreover, the onus of responsibility must be at the level where the change is desired to take place, that is, at the level of the individual schools and classrooms. It is doubtful that disempowered teachers and administrators have the will or the capacity to effect the lasting improvements. Without supports beyond the Department, the chances of the reform having the sweeping impact that was originally proposed are highly unlikely.

§ Limited Educational Experiences

The compromises that teachers are experiencing combined with the drive to increase student test scores has many teachers concerned about the quality of instruction. Many teachers worry that students are not receiving a balanced, well-rounded education. We find it ironic that a system designed to raise the bar on student learning may instead be providing students with more limited educational experiences.

‘this is handed down by the district, so...with the science kits, and the mandated four blocks, and this and that, I’m just feeling like, well, if I have to do the four blocks and everything else, I’m going to end up short-changing something...’

“And we’ve been putting so much emphasis on writing this year for instance, that I feel that I’m not doing math as well as I did last year, or as well as I could do”

“...if I could spend a day or a couple of days and take apart, and look at the inside of a battery, let’s say, that’s science that matters to them. I wasn’t able to take advantage of that teaching moment”

“I feel like we’re not giving the children a full education. A well-rounded education...”

Limited educational experiences may be just one cost to students; we do not yet know what the full price to students will be. With the consequences associated with the DSTP set to begin next spring we plan to focus the third year of the case study report on students’ responses to the accountability plan.

Where do we go from here?

Summary

Over time, and for a variety of reasons, measurement-driven reform has taken the place of standards-based reform in Delaware. This overemphasis on testing has caused standards to take a backseat. This is clearly supported in our current research and is succinctly summarized by one teacher in a school that has had considerable success on DSTP scores.



“You know, you hear that, if you teach the standards, the kids will do well on the test...that is a total lie. I would look anybody, and teacher in the eye and say, ‘You’re lying’.”

National experts also speak of the emphasis on testing over standards. Anthony S. Bryk, a professor of education and sociology at the University of Chicago says,

“When we talk about standards-based reform in Chicago, and it’s actually true everywhere, don’t show me the standards documents. Show me what you test, because the load-bearing wall in all of this is not the standards documents, it’s the assessments.”

Consequently, we believe that much of the distortion that we currently see in Delaware schools is attributable to this shift of focus. For example, the conflict that educators struggle with as they try to reconcile the tenet “all children can learn” with what is known about intelligence, aptitude, and developmentally-appropriate instruction does not derive from the ideals of standards-based reform. Rather, it stems from the disconnect between standards-based and measurement-driven reform. Clearly stated in all of the Delaware Content Standards documents is “the principle that all students can learn and consequently will be held to high academic expectations of knowledge and performance.” The operative phrases here are “can” and “will be held.” When viewed within the standards-based system, ‘can’ translates as ‘has the capacity to.’ Moreover, “will be held,” from the standards-based perspective, translates to insisting that teachers hold high expectations of all students, helping each of them strive toward higher levels of learning, while individualizing their instruction so as to maximize their potential to learn. In the measurement-driven system, the translation is that all students must achieve a certain score on a test. In the case of measurement-driven reform, so much is “lost in the translation”.

Policy Considerations:

- How can the Board and the Department get the focus back on standards?
- If all schools refocus their instructional programs on high standards, what board policies will sustain that in the future?
- How can this be done in light of the fact that teaching to the test and homogenization of curriculum and instruction, in some places, are actually yielding high test scores?
- How can the state continue to encourage the professional development of teachers in areas that may be beyond the immediate scope of the state testing program?
- How can the state help foster professionalism in an environment that is beginning to have the opposite effect?

Appendix A

METHODS

Site Selection

While some may argue that it is not possible to “see the whole picture” without involving all 19 districts, we believe that the sample selected for this study provides an adequate representation of the school districts in the state. This is particularly relevant in that the study is formative, not evaluative, in nature and intent. This sample of Delaware school districts was based on the following selection criteria:

- Grade configuration: elementary school (2), middle school (2), high school (2)
- Districts and counties (rural and urban areas)
- DSTP: low, average, and high
- Size of school: small, medium, and large
- Minorities: Percent of low, average, and high
- LEP (Limited English Proficiency)
- Income Level (SES): low, average, and high
- Special Education

The selection of the schools (site sampling) was made among several schools that initially agreed to participate in the study. Due to the difficulty of finding schools that fit all of these criteria and then matching the selection to achieve a perfect representation, the best effort was made to select the optimal combination.

Verification of Year 1 Findings: Participant Checks

One means of ensuring the validity of qualitative research findings is through the use of participant checks. This process involves having study participants critically examine the findings and give feedback as to their credibility and trustworthiness. At the beginning of the 2000-01 academic year, study participants at the five sites, principals and teachers, were asked to critique Year 1 findings (**see Appendix B**). The strength of the participants' concurrence with the findings supports the validity of the analysis. This is particularly important since the Year 2 data collection and analytic activities were designed to elaborate the results of Year 1.

Data Collection

The data collection process included two primary sources of data, interviews and documents, gathered from school level personnel. Structured interview protocols (see **Appendix C**) were developed and used to guide the interviewing and subsequent analysis of data collected. Members of the Research & Development Center staff conducted the focus groups and individual interviews at each of the six (6) participating sites. All study participants were assured that their identities and the identity of their district would be held strictly confidential.

Data Analysis

Multiple forms of iterative data analysis were used for both the interview and document data collected. Preliminary analyses were conducted by members of the interviewing team through the writing of analytic memos. The initial analysis activities generated case data specific to each of the schools. Cross-case data analyses that followed were designed to enhance the generalizability of the findings. Coding of interview and document data was accomplished using a computerized data analysis program, Ethnograph 5.0. The analysis was focused by categories were developed based on last year's findings. A cross-site coding procedure was then used to generate the findings as delineated in the report.^v Triangulation, including the use of multiple data collection methods (interviews & documents), multiple researchers, and multiple sources (teachers & principals) contributed to the overall validity of the study.

^v Specific information regarding process of within-case and cross-case data analysis is further described in "Qualitative Data Analysis" by Matthew B. Miles and A. Michael Huberman.

Appendix B

Participant Check

Responses to *Navigating Accountability: The Year 1 Case Study Report*

Please indicate where you work (Circle one):

Elementary School ($n=8$)

Middle School ($n=7$)

High School ($n=3$)

How well does our report reflect your view about the student accountability plan?

Not at all

Very well

1

2

3

4

(average response=3.47)

How well does our report reflect your peers' thoughts about the student accountability plan?

Not at all

Very well

1

2

3

4

(average response=3.47)

To what extent did the questions asked in the interview adequately address your thoughts and concerns about the student accountability plan?

Not at all

Very well

1

2

3

4

(average response=3.61)

To what extent do you feel that the interview process provided you with an opportunity to express your thoughts?

Not at all

Very well

1

2

3

4

(average response=3.72)

What issues were not addressed in the report that you think should be raised?
(open-ended responses)

1. I thought the report was very thorough.
2. Parent accountability touched upon but not a major focus as it should be.
3. Students are being tested too often, students are being tested at too young of an age (3rd grade), high stakes testing make school “test prep” schools—learning comes second.

Participant Check (cont.)

What issues were not addressed in the report that you think should be raised?

(open-ended responses)

1. How will the needs of “non-academic” students be met? Where will future mechanics, electricians, etc. be educated? Sussex Tech is not a vocational school in the true sense of the word. All students are not college bound. Even the vocational world is more exacting.
2. Needs more days to teach before DSTP. Begin school earlier! 1/4th of year still left after test
3. How should special education scores be reported?
4. Teaching to the test, test taking practice taking away from instructional time.
5. I think simply a greater analysis should be made on the deleterious effects of poor press and politics in this process.

Are there additional issues that have emerged since the time of our interview that you want the policymakers to know? *(open ended responses)*

1. Additional accommodations for special education students have been added which will make the DSTP more appropriate for these students.
2. I think we have made too many changes too quickly. We have a new reading series, 4 blocks, TERC/TRAILBLAZERS and science kit training, etc. My concern is this: How will we know what worked or did not work? What if the scores drop due to implementation year “jinx”?
3. The high stakes testing is more a sample of the haves and have nots (socioeconomic issues). Although raising the bar sounds good, we are increasing the gap by sorting kids.
4. Issue clear expectations—what is graduation requirement—what holds 11th and 12th graders accountable after they pass the test?

Appendix C

Principal Interview Protocol

Introduction: This is our second year of studying the student accountability plan. As you know we are interested learning about how schools, teachers, and students are being affected by and are reacting to the state’s student accountability plan. I would like to ask you a few questions about your thoughts. Everything that you share with us will be held strictly confidential and none of the reports released by the R&D Center will name yourself, your school, or your district. I have a few questions that I would like to explore with you.

| <i>Focus</i> | Questions | Probes |
|--------------------|--|--|
| Changes in Schools | 1. To what extent do you think the student accountability plan is changing your school? | <ul style="list-style-type: none"> a. How do these <u>changes</u> relate to <u>student outcomes</u>? b. How have these <u>changes</u> affected your <u>staff’s attitudes</u> (e.g. teachers’ instruction methods)? c. One of the issues to emerge in the first year of our study was that principals and teachers felt that the <u>decision-making power was moving away from the school</u>. Do you think that has changed this year? [Examples] d. Does your school have a <u>reward system</u> based on student performance? [Please, describe it]. |
| Data Use | <p><i>One of the main goals of the Accountability plan is to provide schools with data about their students. We would like to learn about how your school uses test data.</i></p> <p>2. How do you see the DSTP testing data being used in your school?</p> | <p>In regard to:</p> <ul style="list-style-type: none"> 2. <u>Curriculum development</u>? - 3. <u>Instructional emphasis</u>? 4. <u>Professional development</u>? 5. <u>Test development</u>? 6. <u>School policy</u>? 7. Other? [Examples] |

| | | |
|---------------------------------|---|--|
| <p>Resource Allocation</p> | <p><i>Delaware policymakers have indicated that one of the original goals of the Delaware student accountability initiative was to create a system that expects more and provides more. We want to learn more about this issue from your perspective.</i></p> <p>3. Have they're been or do you anticipate changes in resource allocation in relation to your school budget?</p> <p>4. How do you think this (the district's) resource allocation will affect student outcomes?</p> | <p>8. In what areas? [Examples] 9. How would you describe the funding priorities or your district?</p> |
| <p>Professional Development</p> | <p><i>Delaware policymakers also indicated that one of the goals of the Accountability plan was to motivate educators toward continuous improvement as professionals. We would like to know more about the profession development offered to teachers.</i></p> <p>5. What do you see as your teachers' greatest needs in terms of professional development?</p> <p>6. What types of professional development do you see your teachers pursuing?</p> | <p>10. How do you see your teachers benefiting from their professional development? [Examples]</p> |
| <p>Other</p> | <p>7. Is there anything else you want to tell us about the changes that have or are occurring in your school?</p> | |

Teacher Focus Group Protocol

INTRODUCTION: We would like to get a better understanding of teachers’ opinions and experiences as they relate to the student accountability plan and how the plan is changing their school, affecting their classrooms, and changing school-level decision making. We would like to encourage you to speak on behalf of not only yourself, but to the extent possible, your colleagues as well. As we talk today, I ask that only one person speak at a time, that there be no other conversations going on while someone is talking, and that everyone feel free to participate. Since we will be using what we learn from this session for research purposes, I want to assure you that none of you will be identified by name in any of the reporting that is released. Also so that I am sure that I accurately record what you say, I am using a tape recorder to record your responses.

| Focus | Questions | Probes |
|-----------------------|--|---|
| <i>School Changes</i> | <p>We are trying to get an understanding of how teachers are experiencing the student accountability plan. One of the major assumptions of accountability is that it can serve as a powerful tool for change. I would like to introduce a metaphor if I could, and the metaphor deals with “the process of construction” like you often see on Delaware roads or various building sites. Construction can involve remodeling (slight changes to the appearance of the site), renovation (extensive updates and improvements to pre-existing structures), making additions to existing structures or constructing entirely new structures.</p> <p style="text-align: center;">1. Which of these (remodeling, renovation, new construction) best represent the changes you have seen come about as a result of the student accountability plan?</p> | <p>Why did you choose this type of construction?</p> <p>Can you share a specific example that illustrates this?</p> |

| | | |
|------------------------------------|---|---|
| <p><i>Data Use</i></p> | <p>When we spoke to the policymakers involved with the creation of the student accountability plan, they indicated that one of the goals of the plan was to improve student achievement by providing a system with clear criteria of performance and measuring student performance against the state content standards.</p> <ol style="list-style-type: none"> 1. <i>How has the data provided by the student accountability system changed the way teachers approach their jobs?</i> 2. <i>How has the data proved by the student accountability system changed the way teachers deal with students?</i> 3. <i>How has the data provided by the student accountability plan changed the way teachers interact with one another?</i> 4. <i>In keeping with our construction metaphor, think about a construction site and building tools, what tool do you think best represents the DSTP data?</i> | <p>With respect to instruction, planning, timing of content covered?</p> <p>Can you provide specific examples to illustrate this?</p> <p>Please provide an example of a time when you saw the data used as this type of tool?</p> |
| <p><i>Decision-making Loci</i></p> | <p>In keeping with our “construction” metaphor, I would like you to imagine that your school is a “construction site”.</p> <ol style="list-style-type: none"> 5. <i>Are the architects of your construction site (the people making the decisions about what goes on in your school) more likely to be from the state, district, or your school?</i> | <p>Do you see that as good or bad or both? In what ways is it good and in what ways is it bad?</p> |

To conclude, I would like to ask each of you to give your general reactions to what you have heard today. Were there any surprises or concerns?

Teacher Interview Protocol

INTRODUCTION: Thank you for meeting with me today. As you know we are interested in learning about how schools, teachers, and students are being affected by and are reacting to the state’s student accountability plan. I’d like to ask you a few questions about your thoughts on this topic. Everything that you share with us will be held strictly confidential and none of the reports released by the R&D center will name yourself, your school, or your district.

| Focus | Questions | Probes |
|----------------------------|--|--|
| <i>Stress and Conflict</i> | <p>As you know, we had several focus groups with teachers like the one we had here at your school, focused on Delaware’s student accountability plan. One of the clear messages that came out of those meetings was that teachers are under a great deal of stress.</p> <p>Some teachers indicated that they experienced stress because they felt like there was sometimes a conflict between what they thought they should be doing (based on their professional experience) and what they were required to do to meet the requirements of the accountability plan.</p> <p style="text-align: center;">1. To what extent have you felt this way?</p> | <i>Can you give a specific example of a time when you felt this way?</i> |
| <i>Teacher Motivation</i> | <p>In addition to reporting greater levels of stress teachers said that they felt that all of their hard work was unappreciated.</p> <p style="text-align: center;">2. Do you feel this way, and if so, what keeps you going?</p> | |

| Focus | Questions | Probes |
|-------------------------------------|---|---|
| <i>Promoting Improvement</i> | <p>In our focus groups teachers also told us that they did not find the data provided by the DSTP to be all that useful for improving how they teach. Providing data on student performance was supposed to be one way of promoting continuous improvement of classroom instruction. Another piece of the student accountability plan that was designed to promote continuous improvement was the development of IIPs for low performing students but this didn't come up in our focus groups.</p> <p>1. <i>To what extent do you find the process of developing IIPs beneficial for students or for the following year's classroom teacher?</i></p> | |
| <i>Prior Experience with Reform</i> | <p>The student accountability plan is an attempt to reform education but it probably isn't the first time you have experienced a reform initiative.</p> <p>2. <i>Prior to the current push for accountability, can you think of other reform initiatives that were implemented in your school or district?</i></p> | <p>What were they and how were they received by the teachers in your school?</p> <p>To what extent did they have lasting effects? Why do you say this?</p> <p>Of these reforms, were there any that you thought could/would have a lasting effect on student learning and if so what were they?</p> |

| | | |
|-----------------------------------|--|---|
| <p><i>Reform Expectations</i></p> | <p>We would like to know more about what you think might be the lasting effects of the student accountability plan.</p> <p>3. Imagine <i>that the date is <u>Today's date</u>, 2011 (ten years from the date), what will student accountability look like?</i></p> | <p>How will things have changed for students?</p> <p>For teachers?</p> <p>For Delaware's economy?</p> |
| <p><i>Other</i></p> | <p>4. Imagine that you were asked to speak to teachers from other states who were about to embark on an accountability system like the one here in Delaware. What advice would you give them? What is the most important thing they should know?</p> | |

Endnotes

1. Noble, A.J., & Banicky, L. (2000). Synchronizing the Accountability Clocks: A Policy Study of Delaware's Student Accountability Plan. Banicky, L., Noble, A.J., & Siach-Bar, Y. (2000). Navigating Accountability: Delaware Schools' Response to the State's Student Accountability Plan. <http://www.rdc.udel.edu>
2. New Directions, Why New Directions?, published by the Delaware Department of Public Instruction, undated, prior to 4th quarter. 1992
3. Mid-continent Research for Education and Learning (2000). Noteworthy perspectives on implementing standards-based education. <http://www.mcrel.org/products/standards/implement.pdf> (20 June 2001)
4. See previous footnote.
5. New Directions for Education in Delaware Vision/Mission Statement. DE DPI.
6. New Directions, Most Often Asked Questions and Answers, undated, prior to 4th quarter. 1992.
7. Washington Post, May 14, 2001 Why we must have testing. Written by Rod Paige, U.S. Secretary of Education.
8. Corbett, H.D., & Wilson, B.L. (1991). Testing, reform, and rebellion. Norwood, NJ: Ablex Publishing Corporation. Mathison, S.M. (1991). Implementing Curricular change through state-mandated testing: Ethical issues. Journal of Curriculum and Supervision, 6 (3), 201-212. Noble, A.J., & Smith, M.L. (1994). Old and new beliefs about measurement-driven reform. Education Policy, 8(2), 111-136. McNeil, L. (2000). Contradictions of school reform: Educational costs of standardized testing. New York: Routledge.
9. Wheelock, A. (1995). Standards-Based Reform: What Does It Mean for the Middle Grades? (ERIC Document Reproduction Services No. 389 772)
10. New Directions , Most Often Asked Questions and Answers, DE DPI
11. New Directions, Background on the Comprehensive Assessment System, DPI, undated, prior to 4th quarter. 1992
12. Smith, M., & O'Day, J. (1991). Putting the Pieces Together: Systemic School Reform. CPRE Policy Briefs. CS: Consortium for Policy Research in Education, New Brunswick, NJ.

13. March 1992, New Directions for Education in Delaware: A Proposed Game Plan, SBE & DPI.
14. Brehm, S., & Kassim, S. (1993). *Social Psychology*. Boston: Houghton Mifflin Company
15. See previous footnote.
16. Allen, V. (1965). Situational factors in conformity. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology, 2, 133-175. Kelman, H. (1961). Processes of opinion change. Public Opinion Quarterly, 25, 57-78.
17. Cialdini, R., Trost, M., & Newsom, J. (1995). Preference for consistency. The development of a valid measure and the discovery of surprising behavioral implications. *Journal of Personality and Social Psychology*, 69, 318-328.
18. Source: New Directions for Education in Delaware: A proposed game plan. Prepared by the State Board of Education and the State Superintendent of Public Instruction. March 25, 1992
19. Source: Standards-Based Accountability, A White Paper written by Marsha DeLain, Associate State Supt, DPI, and Paul LeMahieu, University of Delaware and DPI.
20. Metts, I. (1998). The Secretary of Education's Vision for State Assessment and Educational Accountability.
21. What are the Elements of New Directions? DPI. 12/94
22. Fish, J. (1988). Response to mandated standardized testing (unpublished doctoral dissertation, University of California). Dissertation Abstracts International, 50, 122.
23. Shepard, L.A., & Dougherty, K.C. (1991, April). Effects of high-stakes testing on instruction. Paper presented at the annual meetings of the American Educational Research Association and national Council on Measurement in Education, Chicago, IL.
24. Hatch, A. & Freeman, E.B. (1988) Who's pushing whom? Stress and kindergarten. Phi Delta Kappan, 70(2), 145-147.
25. McNeil, L. (1986). Contradictions of control: school structure and school knowledge. New York: Routledge & K. Paul.
26. Vance, J. (2000). Steps Toward Data-Driven Policy: A profile of teacher supply and demand data in the Mid-Atlantic region. A report of the Mid-Atlantic

Regional Teacher Project sponsored by the Laboratory for Student Success, with the Council for Basic Education and the Maryland State Department of Education.