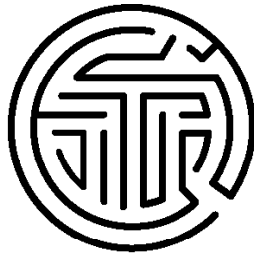




**LINKING DATA AND
LEARNING — THE GROW
NETWORK STUDY**

SUMMARY REPORT



C C T R E P O R T S
| A N U A R Y 2 0 0 5

**LINKING DATA AND
LEARNING — THE GROW
NETWORK STUDY**

SUMMARY REPORT

PREPARED BY

DANIEL LIGHT, PH.D.
MARGARET HONEY, PH.D.
JULIET HEINZE
CORNELIA BRUNNER, PH.D.
DARA WEXLER, PH.D.
ELLEN MANDINACH, PH.D.
CHAD FASCA

CENTER FOR CHILDREN & TECHNOLOGY

Copyright © 2005 by Education Development Center Inc.

This report was written and produced by EDC's Center for Children and Technology

All rights reserved. No part of this report may be reproduced or transmitted in any form or by any means without permission in writing from the publisher, except where permitted by law.

To inquire about permission, write to:

ATTN: REPRINTS

Center for Children and Technology

96 Morton Street, 7th Floor

New York, NY 10017

Or email:

cct_reprints@edc.org

EXECUTIVE SUMMARY

During the last decade and in the wake of the No Child Left Behind legislation, standards, assessments, and accountability have emerged as three prongs of a national education reform movement that has asked district and school administrators to think very differently about educational decision-making and the use of data. However, research about data-driven decision-making is limited. With funding from the Carnegie Corporation, in the spring of 2002, EDC's Center for Children and Technology began a two-year exploratory study that examined how educators and administrators within the New York City public school system are using data—made available to them through the print and web-based reporting system of the Grow Network— to inform decisions about teaching and learning and about educational practices.

As a tool, the Grow Report® brings what are often disparate parts of the educational system into conversation and alignment with each other, reinforcing the notion that standards are meant to inform instructional practices not just serve as a means to accountability while also identifying the standards the tests are highlighting (because every standard cannot be tested). By design, it provides a format that builds a bridge between standards, testing results, and instructional strategies and appears to be highly successful in creating a navigational framework for educators.

Using a mix of qualitative and quantitative methodologies, the two-year study unfolded in three phases. Phase One focused on understanding the ways in which central office personnel, along with district superintendents and their education teams, thought about using data to inform decision-making. Phase Two emphasized ethnographic research in 15 schools across four school districts in New York City that represented various neighborhoods, student populations, and overall performance levels. Phase Three involved the development and administration of two separate surveys across the New York City public school system that asked teachers and administrators about how they interpret data and conceptualize the use of the Grow Reports® for instructional planning.

Teachers' Attitudes Toward Data, Decision-making and the Grow Reports

What we have learned from this study is that teachers, in particular, are open to data, but also scrupulous about its use. Rather than accept a narrowing interpretation of their students' strengths and weaknesses based on a single test, they instead rely on multiple sources of data – impressionistic, anecdotal, and experiential – accrued over the long term and based on many experiences with their students to make most instructional decisions. Yet having been introduced to the possibilities of using systematic data to make instructional decisions, teachers are eager for more and better data. The Grow Report® is an important first step in this process and shows the promise of making the transition to data-driven decision-making in our schools.

No matter how teachers viewed state-mandated, standardized testing, whether skeptically or acceptingly, they recognized, across the board, that part of their job is to prepare students to take

the test. But they have questions about the test. In interviews and survey responses, they expressed concern about the accountability environment's impact on instruction, with three quarters of teachers feeling that the tests lead them to teach in ways that contradict their own ideas of good teaching. The majority of teachers also questioned the test's accuracy in measuring students' academic abilities and in measuring the "life skills" that students need to succeed as well as the test's cultural sensitivity or developmentally appropriateness to all students.

Based on these concerns, teachers frequently seek to monitor student learning and triangulate assessment data in a variety of different ways. All of the teachers surveyed reported using multiple assessment strategies, either always, often, or sometimes. When interviewed, the teachers discussed mixing various assessment strategies to provide a fuller picture of student understanding and learning. Overall, the teachers felt that teacher-made assessments and strategies were more useful than the external assessments. Only a handful of teachers said that they rely solely on formal assessments.

Throughout the survey and in the interviews, teachers across New York City reported using the Grow Reports® in various ways to meet their overall classes' as well as their diverse students' academic needs. Grow-using teachers discussed making decisions within several specific areas of their instructional practice: (1) targeting instruction, with decisions about class priorities, lesson plans, and the academic year; (2) meeting the needs of diverse learners, seen in strategies such as grouping, creating Individualized Education Plans (IEPs), and giving individualized assignments and materials appropriate to the students' levels; (3) supporting conversations with parents, students, fellow teachers and administrators about students' learning; (4) shaping teachers' professional development by reflecting on their own practice; and , (5) encouraging self-directed learning by giving the data to students.

Teachers noted that the Grow Reports® provided them with more information about students than what they had access to previously, both as a class and as individuals. When asked how they use the Grow Reports®, several teachers responded that they use the data when deciding where to target their instruction from creating a more extensive yearlong pacing calendar to planning mini-lessons that review and reinforce certain concepts or skills. Most teachers also felt that having timely data that showed individual student performances helped support their decisions to differentiate instruction according to individual students' needs or to group students, based on how they performed on specific skills to target certain skills.

Despite the high-stakes climate in which they teach, Grow-using teachers felt less forced to exclude anything not on the test from their teaching and estimated that they spent less time on explicit test preparation activities than was true for a national sample of teachers. Compared to the national sample, these New York City teachers can be considered "data-friendly." In other words, the skepticism many expressed about the value of standardized test data in educational decision-making was often more about its incompleteness and the stakes involved, than a rejection of the whole idea of using standardized tests to measure individual achievement. Many teach-

ers were careful to note that the Grow data was one “small piece” in a wide array of assessments they use, including observations, in-class assignments, daily quizzes, unit pre- and post-tests. Moreover, teachers added that the data on the Grow Reports® has “declining value” as the year progresses, in part because students take the exam roughly six months before the reports are distributed.

Administrators’ Attitudes Toward Data, Decision-making and the Grow Reports

Administrators’ attitudes about high-stakes testing are not markedly different from teachers. Administrators clearly feel pressure to improve test scores; they discussed how they support teachers’ testing needs and how the growing accountability culture has influenced their schools and their own decisions. With its adoption becoming more widespread through NCLB, standardized testing is not only being used to directly measure students’ academic progress but also to indirectly assess administrators’ leadership.

However, administrators, like teachers, have some reservations about what the tests are measuring. Administrators questioned the validity and reliability of the test in their responses to our survey; the majority of administrators do not consider the test as accurate as a teachers’ judgment of what students know and can do. Administrators were divided over most other issues related to state-mandated tests, such as whether or not the state-mandated test were an accurate measure of what students know, or whether test pressure narrowed the scope of the curriculum. Despite this ambivalence about the test’s effect on curriculum scope and teaching practice, administrators did feel that the state-mandated test is aligned to what teachers teach in the classroom. And though many expressed a desire to assess student progress from different angles in an ongoing fashion, a vast majority of the administrators surveyed reported that, in order to prepare students for the test, they encourage teachers to “teach the students test-taking skills.”

New York City district- and building-level administrators reported using the Grow Reports® to gain a greater understanding of the educational and instructional concerns particular to their level of the education system. Administrators explained that the Grow Reports® helped them to identify class-, grade-, and school-wide strengths and weaknesses that could then be used to make decisions about planning, shaping professional development activities, and determining student performance and demographics. In interviews, many of the administrators spoke about how the Grow Reports® helped frame conversations they had with teachers, parents, or other administrators related to student learning, professional development for teachers, or addressing school or district challenges. For example, since the Grow Reports® and associated instructional resources are approved by the New York City Department of Education and aligned to state standards, administrators found that the reports were often a good fit for shaping professional development activities.

Technology's Role in Balancing Tension of Practice and Policy

While policymakers have embraced the notion that a single assessment can measure students, educators in this study acknowledge that high-stakes testing communicates only a piece of what they need to know about the complex repertoire of skills and talents that children need to succeed. Herein lies the gap between what national policymakers and local practitioners see as important. Digital technology has already played an important role in bridging this gap by giving teachers access to high-stakes test data. We believe that digital technologies will expand on this role by helping teachers engage in data-gathering and data-analysis processes inclusive of the multifaceted ways in which children show evidence of learning, allowing teachers and administrators to track performance data, observational data, informal conversations, portfolios of student work, self-assessment and reflection, the stuff of daily instructional decision-making in the classroom. The inclusion of these organized, diagnostic, and authentic, performance-based data in the decision-making process would also go a long way toward refining and supporting the practices that educators routinely engage in.

The Grow Report™ represents an important initial step on this path. By creating a lens through which the relationship between standards, assessments, and instruction can be explored, it helps educators to find reason in and navigate the tensions that prevail in the high-stakes contexts in which they work.

Introduction

In the wake of the No Child Left Behind legislation (NCLB, 2001), increasing attention is being given to accountability and data-driven decision-making in public and professional arenas. While urban districts have faced intense external scrutiny for some time (Fullan, 2000), the shift in funding and the regulatory requirements occasioned by NCLB are prompting educators to think differently about the potential of data to inform instruction and decision-making aimed at improving school achievement. School personnel are working hard to develop strategies that support the cohesive use of data across different levels of a school system, and the exploration of how data can inform educational decision-making is becoming a main topic of educational policy (Salpeter, 2004; Secada, 2001). Currently, however, research about data-driven decision-making is limited. We have only a cursory understanding of educators' existing practices and we know little about how these practices are informed by the influx of data-driven tools.

With funding from the Carnegie Corporation, in the spring of 2002, EDC's Center for Children and Technology (CCT) began a two-year exploratory research study to examine how educators within the New York City public school system are using data to inform decisions about teaching and learning. The opportunity that occasioned this research was a decision by New York City's Board of Education to establish a five-year contract with an assessment reporting company, the Grow Network (<http://info.grow.net>), whose mission is to "transform assessment results into instructional tools for teachers, principals, and parents" (Grow Network, 2004) using a mix of print and web-based reporting systems. The print materials, called Grow Reports®, deliver customized print reports to teachers, principals, and parents. Grow Reports® for teachers give a concise, balanced overview of standards-based, class-wide priorities; group students in accordance with their learning needs; and enable teachers to focus on the strengths and weaknesses of individual students. The Reports for principals provide an overview of the school, presenting grade, class, and student-level data. The Reports for parents provide easy-to-interpret information that explains the goals of the test, how their child performed, and what they can do to help. The Reports for the Web provide teachers with much more detailed information about their students, and also make available links to "teaching tools" that help explain the standards and are solidly grounded in cognitive research about effective math and literacy learning (see Appendix A for copies of Grow Reports®), thereby providing data and instructional tools through the same reporting system.

New York contracted with the Grow Network to provide reports on the third-grade through eighth-grade English Language Arts and Math assessments, where there are 30,000 teachers, 5,000 district and school instructional leaders, and 1,200 schools serving approximately 500,000 students. This represented an unprecedented effort to use city-mandated assessment data, coupled with supporting teaching resources and professional development, to improve the quality of educational decision-making across multiple levels of the school system.

Importance of this Study

During the last decade, standards, assessments, and accountability have emerged as the three prongs of the national education reform movement and are now broadly embraced by the education policy community. Over two decades ago, with the release in of *A Nation at Risk* (1983), the need for higher academic standards became a national issue. Following on the heels of this report, governors, business leaders, and education policymakers have gathered at three National Education Summits, in 1989, 1996, 1999, to create a bipartisan agenda in support of instructional standards and learning environments that promote high expectations for all students.

The culmination in the move toward increasing accountability and the delineation of achievement standards for all students has been the NCLB legislation. The attention that NCLB has brought to educational accountability has been unprecedented, a presence felt in everything from professional educational publications to national and local public media. NCLB holds districts, individual schools, and teachers accountable for student performance – a standard that implies that decision-makers have access to data at the appropriate level of aggregation (district, school, teacher, individual student) and that they are able to interpret them. NCLB requires that disaggregated data be examined for subgroup performance and that all subgroups meet mandated adequate yearly progress (AYP) standards.

While there is little doubt that accountability measures have risen in popularity among both politicians and the public, the debate about the efficacy of standardized measures of achievement and “high-stakes” testing continues to loom large in education circles (Elliot, 1993; Herman & Golan, 1990; Lemann, 1999; Meyers & O’Connell, 2000; Newman, King, & Rigdon, 1998; Pellegrino, Chudowsky, & Glaser, 2001; Sacks, 1999; Stiggins, 2002; Viadero, 2000). Under NCLB, many traditionally defined high-achieving schools are indeed failing to meet AYP (Robelen, 2003). The consequences are real and significant for educational decision-makers and stakeholders, and the pressures on practitioners are ever increasing to improve student performance (Linn, 1998, 2001a, 2001b, 2003a, 2003b).

One potentially powerful consequence of the standards and accountability movement is that district and school administrators are being asked to think very differently about educational decision-making and are beginning to use data to inform everything from resource allocation to instructional practice. As researchers at the UCLA Center for Research on Evaluation, Standards, and Student Testing (CRESST) note, “Data-based decision-making and use of data for continuous improvement are the operating concepts of the day. School leaders are expected to chart the effectiveness of their strategies and use complex and often conflicting state, district, and local assessments to monitor and assure progress. These new expectations that schools monitor their efforts to enable all students to achieve assume that school leaders and teachers are ready and able to use data to understand where students are academically and why, and to establish improvement plans that are targeted, responsive, and flexible” (Mitchell, Lee, & Herman, 2000).

Schools and districts grappling with accountability strategies at the local level are the first to acknowledge that this is a time-consuming process that requires extensive, ongoing professional development work with teachers and school leaders (Council for Basic Education, 2000). It is not surprising, then, that a Council for Basic Education report concludes: “States and districts will need to think about moving from the three-legged strategy of standards, assessments, and accountability to a model that has a fourth leg – support” (p. 9). As pressures for accountability continue to increase, effective alignment across multiple levels of the school system will become more important in the use of data (Elmore & Abelman, 1999; Fullan, 2001).

There is no question that data-driven decision-making is a complex undertaking, even for the trained educator who understands statistical concepts. As Secada (2001) notes, data should be used to inform decisions, not replace them, and this process requires time and effort. Yet, we lack a sufficient knowledge base to understand the kinds of data-driven practices and strategies that teachers and administrators can use to improve student performance (Schafer & Lissitz, 1987; Wise, Lukin, & Roos, 1991). While efforts that focus on helping principals and teachers improve their schools have grown dramatically over the last several decades (Kearns & Harvey, 2000), it is rare to find school communities in which teachers and administrators routinely engage in thinking critically about the relationship between instructional practices and student outcomes. We are equally hard pressed to find substantial numbers of educators who have adequate training and knowledge and are prepared to make appropriate use of data and transform it into useable information and practice (Cizek, 2001; Herman & Gribbons, 2001).

Complexity in the data-driven decision-making process also stems from the systemic structure of school districts. Schools are multi-level organizations composed of dynamically interacting components (Cromey, 2001; Mandinach & Cline, 1994; Senge et al., 2000). The systemic perspective recognizes that because of the cross-level interactions within school systems, data, information, and decision-making affect many components of the system at the same time (Fuqua, Newman, & Dickman, 1999). Thus, the need to consider the issue of how data and information flows through a learning organization such as a school district is critical.

While a number of technical advancements enabling innovative reporting mechanisms have made data-supported decision-making a much more realistic undertaking (Wayman, Stringfield, & Yakimowski, 2003)¹, questions about how educators who are working at different levels of the school system use data to inform decision-making remain largely unanswered. Preliminary work on the experiences of different data-systems is underway at handful of sites around the country. These include research on the Quality School Portfolio (QSP) developed at CRESST (Mitchell & Lee, 1998) and IBM Reinventing Education data project in Broward County, Florida (Spielvogel et al., 2001), the Texas Education Agency, and the South Carolina Department of Education (Spielvogel & Pasnik, 1999). Research on the role of data systems and applications in practice is also being

¹The John's Hopkins Center for Research on the Education of Students Placed at Risk maintains a web site that contains updated information and reviews on technology-based data-driven decision-making tools. It can be found at: <http://www.csos.jhu.edu/systemics/datause.htm>

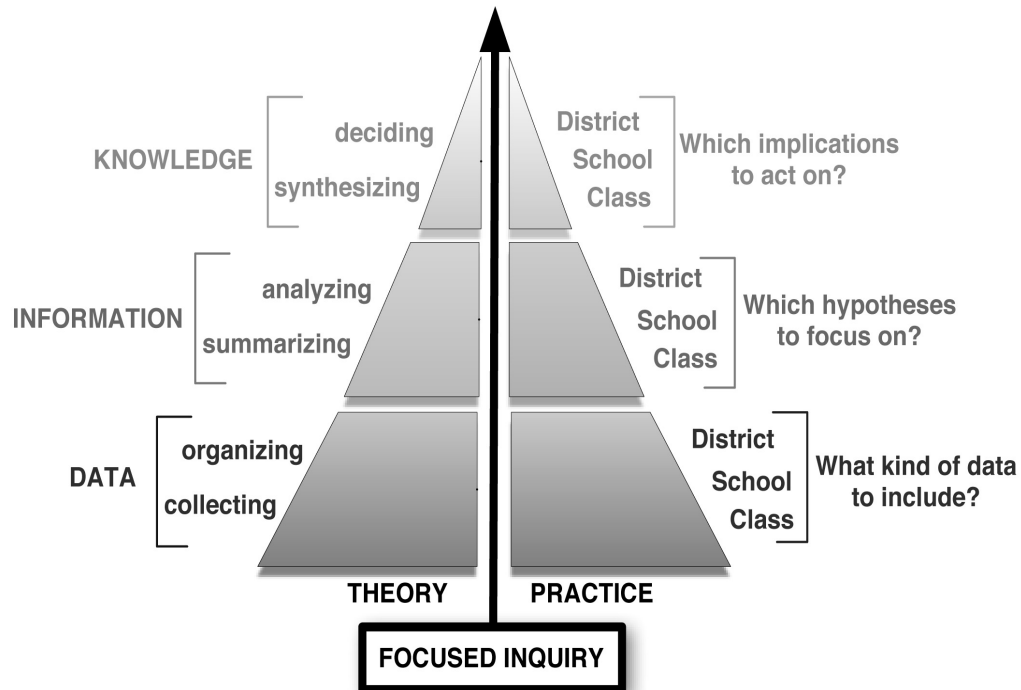
done in Minneapolis (Heistad & Spicuzza, 2003), Boston (Sharkey & Murnane, 2003), and on the implementation of QSP in Milwaukee (Thorn, 2002; Webb, 2002). Nonetheless, the New York City school system's partnership with the Grow Network is possibly the largest project of its kind. The study we report on here investigates how data is used and thought about in the classrooms and schools of our nation's largest school district, and, thus, speaks to important policy concerns about the role of standardized testing and data-driven decision-making in education in general.

A Theoretical Framework for Data-Driven Decision-Making

One of the six guiding principles in a recent National Research Council report on *Scientific Research in Education* (Shavelson & Towne, 2002) is linking relevant theory to research. In developing a theoretical framework that would guide our research in this project, we drew upon organization and management theory, in particular the work of Ackoff (1989), Drucker (1989), and a former CCT visiting scholar (Breiter, 2003). According to Ackoff (1989), data, information, and knowledge form a continuum that can be applied to make decisions. **Data** exist in a raw state. They do not have meaning in and of itself, and therefore, can exist in any form, usable or not. Whether or not data become information depends on the understanding of the person looking at the data.

Information is data that is given meaning when connected to a context. It is data used to comprehend and organize our environment, unveiling an understanding of relations between data and context. Alone, however, it does not carry any implications for future action. **Knowledge** is the collection of information deemed useful, and eventually used to guide action. In relation to test information, the teacher's ability to see connections between students' scores on different item-skills analyses and classroom instruction, and then act on them, represents knowledge.

From these distinctions, we constructed a model of how data are transformed into knowledge, based on a sequence of six steps identified by Ackoff (1989) and Drucker (1989). These steps include collecting, organizing, summarizing, analyzing, synthesizing, and decision-making. The six steps form a continuum of cognitive complexity. During the sequential set of steps, raw data are always made meaningful by a process of contextualizing within the situation that produced the data.



A focused inquiry always guides the processes educators engage in as they move from thinking about relevant data (collecting, organizing), to formulating hypotheses (summarizing, analyzing), to taking actions (synthesizing, deciding). Most importantly, this model suggests that educators' understandings and interpretations of data are likely to vary depending on their role in the school system (teacher, instructional leader, administrator).

The Design of the Grow Reports

By virtue of how they have been designed, the Grow Reports® make a unique and distinct contribution to this theoretical framework. The Reports are intended to serve as a bridge between assessment and instruction, turning raw state and city testing data into information that is organized and aligned with New York State standards.

For teachers, Grow Reports® provide overviews of class-wide priorities, group students in accordance with the state performance standards, and enable teachers to focus on the strengths and weaknesses of individual students. For example, a sixth-grade math teacher teaching during the 2003-2004 school year would have access to a customized report that is grouped according to three questions: (1) How did my students do? (2) What do they need to learn? And, (3) what tools are on the Web? This report performs a summarization and analysis of the data and identifies "class priorities." In response to "How did my student do?" the teacher would see the overall scores for all of her sixth-grade students grouped according to the New York State standards across

four levels, ranging from *Far Below Standards* (Level 1) to *Far Above Standards* (Level 4), along with the students' scale score. On "What do they need to learn?" the teacher would see how her students did in each standard on the Grade 5 test according to New York State's key ideas. She would also see her students as a group compared to all New York City students. In addition, she would have an overview of class priorities, based on last year's test results on the sub-skills. The priorities are divided into three levels – *need help with fundamentals*, *need additional instruction*, and *likely to benefit from advanced work*. These levels are calculated by Grow Network through a complex algorithm comparing each student's sub-skill results to the performance profile of all Level 4 students on that skill, and so on, for each level.

In addition to the paper-based reports, the Grow Network website supports teachers' analysis of the information and instructional decision-making with two additional features. The online tools contain instructional material that define each skill or standard and explain challenges for students in mastering this skill. The class priorities are also linked to resources for teachers and administrators that suggest activities and teaching strategies to promote standards-based learning in the classroom. "Flexible Groupings" of students, which group students by performance in each standard into three categories of performance are also provided. Finally, the reports also link to external resources approved by the New York City Department of Education.

The design underlying the Grow Reports® acknowledges that stakeholders across the school system need different kinds of information, different cuts through and across data, and different levels of focus because the kinds of decisions they make and the kinds of information they need vary considerably. In this respect the Reports are translating raw data into information for different audiences. While the teacher version focuses on students who are at various skill levels, so their instructional needs can be addressed, the parent version reports on individual students and their strengths and weaknesses, explaining the goals of the test, how their child performed, and what parents can do to help their child in Reading and Math. The administrator version examines larger units, including whole classes, grades, and schools, providing an overview of the school, and presenting grade, class, and student-level data. In important respects then, the Grow Reports® represent an ideal artifact through which to understand how interpretations vary depending upon one's position in the school system, and to investigate how the information that the Reports present can serve as a bridge between assessment and instruction.

RESEARCH DESIGN

To understand the ways in which data is thought about and used across multiple levels of the school system and to further illuminate the strategies and practices that characterize data-driven decision-making, we focused the research on three main questions:

1. How does a system-wide adoption of an assessment and instructional innovation unfold in a large, complex urban school system?
2. What kinds of beliefs and understandings do educators have about data, and how do they view its potential to inform teaching and learning?
3. What kinds of resources are needed to support educators in thinking about the relationships among assessment data, instructional practice, and student learning?

Using a mixture of qualitative and quantitative methodologies, we structured the research into three phases. Phase One focused on understanding the ways in which central office personnel, along with district superintendents and their education teams, thought about using data to inform decision-making. During Phase One, we conducted structured interviews with 47 educational leaders, including: central office stakeholders, superintendents, deputy superintendents, math coordinators, ELA coordinators, staff developers, district liaisons, technology coordinators, directors of research and curriculum, and individuals who work with the United Federation of Teachers. We also spoke with several people representing non-governmental organizations who are working closely with the New York City schools on issues such as educational reform and professional development.

During Phase Two we carried out ethnographic research in 15 schools across four school districts in New York City that represented various neighborhoods, student populations, and overall performance levels. Each district identified four schools where we conducted 45 semi-structured and open-ended interviews with principals, assistant principals, staff developers, and teachers. In addition, we observed ten grade-wide meetings and/or professional development workshops. To further explore the ways in which teachers think about using assessment information, we also designed a structured interview protocol, using sample Grow Reports® as projectives. We conducted 31 projective interviews with teachers in the two “high-stakes” testing grades – fourth and eighth - in New York City, as well as with a sample of sixth-grade teachers. The sample of the upper grade teachers was equally divided between those who teach math and those who teach language arts. Phase Two of the research helped us to develop a deeper understanding of how classroom teachers think about using data in relation to their everyday practices.

Phase Three of the study involved the development and administration of two separate surveys across the New York City public school system – one for teachers and one for administrators – and further explored the hypotheses we developed in the previous two phases of work. The surveys asked teachers and administrators about how they interpret data and conceptualize the use of the Grow Reports® for instructional planning. We also inquired about the types of supports needed to

fully leverage the use of data to improve instruction. The survey contained a number of items drawn from a national, random sample survey study about teachers' perceptions of the state-mandated testing program in their state (Pedulla et al., 2003). That study divided teachers according to the level of consequences for students and teachers attached to test results. States were ranked high, medium and low on each dimension. New York was categorized as "HH" – indicating that there were high stakes for teachers and for students. Therefore, on items that we drew from the national study, we compared New York City results (referred to as "Grow teachers" or the "Grow sample") to the national sample of teachers who also worked in "HH" states (referred to as "national sample," "national survey" or "high-stakes states teachers"). The national sample of high-stakes teachers numbered 2,200 responses.

We disseminated more than 750 surveys to 17 schools across the city as well as sent an online survey sent to more than 1,400 teachers and administrators. We received 146 administrator responses and 213 teacher responses from eight schools.² We limited administrator responses to those who use Grow yielding a final sample size of 137 administrator responses. With respect to the teacher responses we restricted the database to language arts and mathematics teachers who taught fourth-grade through eighth-grade students. Because these are the teachers who use the Grow Reports®, this was done to improve reliability. Furthermore, the restriction of our data to a specific sub-population of respondents allows us to triangulate with the qualitative research we conducted which was also targeted on fourth through eighth grades. The final sample size of reported teacher responses is 96. We compared the experience and education levels of our teachers to New York City teachers as a whole, based on data from the New York City Department of Education's 2001-2002 Annual Report Card (Barr, 2004)³.

TABLE 1: COMPARISON OF GROW SURVEY RESPONDENTS AND NEW YORK CITY TEACHERS ON YEARS TEACHING AND MASTER'S DEGREE EARNED

	Survey Respondents (N=96)	NYC Teachers (N=947)
% Teaching >5 years (mean)	51%	51.4%
% of Teachers with a Master's Degree (mean)	71%	71.5%

The data suggest that despite the limited number of teachers who responded to the survey, they appear to be very similar in terms of education and teaching experience to New York City's overall teaching population. Although by no means a representative sample of the City's teaching population, we have confidence that with respect to these three demographic variables they are not a dramatically different group.

² We attribute the low response rate to a change in administrative priorities in the New York City public schools (see section on "Conducting Research in a Climate of Change").

³ The authors would like to thank the Institute for Education and Social Policy at New York University, Steinhardt School of Education for their aid in tracking down this report.

We also compared our teachers to National Center for Educational Statistics (NCES) data (2002) representing the national population of teachers, and our respondents differ from the national population in the following aspects.

- The respondents are ethnically more diverse: 61.4% of Grow teachers are white compared to 87.3% nationally; 8.6% are African –American compared to 6.7% nationally; 19% are Hispanic compared to 4.1% nationally; 2.9% are of Asian/Pacific Islands descent compared to 1.1% nationally; and 2.9% are Native American compared to 0.7% nationally.
- The respondents are less experienced: 15% of the respondents have with less than three years' experience compared to 12.9% of teachers nationally, and 51% of the sample have been teaching for 3-9 years compared with only 28.8% of teachers nationally.
- The respondents are more educated: 71% of the Grow using teachers have a Master's degree compared to 41.9% nationally.

Conducting Research in a Climate of Change

It is important to note that in the course of carrying out this research the New York City school system underwent a complete reorganization and restructuring. Ultimate control over the school system changed hands from the Board of Education to the Mayor. With a new Chancellor and Deputy at the helm, the newly renamed Department of Education restructured the previous 32 school districts into 10 regions, overseen by a regional superintendent and deputy and led by a cadre of local instructional superintendents. Each region encompasses between 100-120 schools that now adhere to a uniform curriculum for language arts and mathematics at the elementary level.

Our original intent—to observe how a tool designed to make the data interpretation process more accessible to teachers and administrators impacts on decision-making at all levels of the NYC school system—had to be altered somewhat during the course of this study. Not only did sweeping administrative and instructional changes take place, the new leadership introduced another accountability resource, supplied by The Princeton Review, that provided teachers and administrators with testing results based on assessments administered three times during the current academic year. While educators still had access to the Grow Reports®, many of the teachers and administrators we spoke to were under the impression that their new priority was to work with the Princeton Review resources.⁴ This was a challenging time, therefore, to observe and chart the movement of an innovation through a system, since the system itself had changed dramatically. Nevertheless, the data we collected can be interpreted to shed a good deal of useful light on data-driven decision making in K-12 schools. This report, therefore, pays attention to both the ideas

⁴ As of the writing of this report, the Grow Reports® are once again being embraced by the current administration and are being used in schools throughout the city.

and issues K-12 educators have about data-driven decision-making, in general, as well as the specific ways in which educators were making use of the Grow Reports®. As it turned out, Grow was a good “conversation piece” for our interviews and surveys. It allowed us to focus both deeply on the actual data produced with this tool and widely on why such innovations are needed and how to make the best use of them.

This report is organized into three sections, which draw on the multiple sources of data we collected in the project. *Educator Perspectives on High-Stakes Testing* investigates teachers’ and administrators’ attitudes toward high-stakes testing and the accountability pressures that surround, influence, and inform their work. *Educator Perspectives on Assessment Practices* examines the kinds of assessment practices that teachers and administrators routinely engage in. And *Educator Perspectives on the Use of Grow* examines the specific ways in which they discuss using the Grow Reports®.

EDUCATOR PERSPECTIVES ON HIGH-STAKES TESTING

Teachers' Attitudes Towards High-Stakes Testing

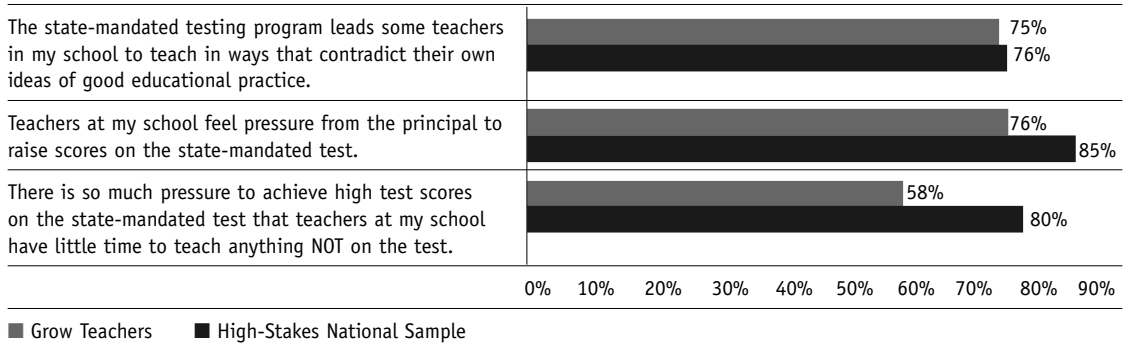
While urban educators have faced accountability pressures for some time now, the implementation of NCLB has ramped up these pressures and raised the profile of standardized testing in teaching and learning to a more prominent role. Given the magnitude of these pressures, we sought to examine as a component of this research educators' attitudes toward standardized testing, their perceptions of test pressure, and their feelings about the appropriate uses of standardized test data. Drawing on the data from the survey and the interviews, we identified three themes: 1) teachers' attitudes toward standardized testing 2) teachers' attitudes about the test as a measure of students' abilities, 3) teachers' attitudes about the use of test results. We describe and discuss each of these themes below.

Teachers' Attitudes Toward Standardized Testing

No matter how teachers approach standardized testing, whether skeptically or acceptingly, they felt, across the board, that part of their job is to prepare students to take the test. All of the teachers we spoke with said they prepare their students for the test in some form or another. The survey results confirm this finding. Ninety-five percent of the New York City teachers surveyed reported that they engage in test preparation activities with their students, and only five percent said that they "do not do any specific test preparation."

Teachers cited the pressure to raise test scores and for students to do well as reasons for dedicating classroom time to test preparation activities. Three-quarters of Grow survey respondents reported feeling test pressure from their principals (75%). The percentage was even higher among teachers in the national survey (85%). A majority of teachers in both surveys reported that test pressure restricted their teaching to likely test content. However, while 80% of national sample reported feeling that they have little time to teach anything not on the test, only 58% of the Grow sample said that the pressure precluded them from teaching other content.

Figure 1. Percent of Grow teachers and the high-stakes national sample agreeing with the following statements (Grow sample, N=96).



Note. The high-stakes national sample represents approximately 2,200 teacher responses.

The interviews helped to expand our understanding of how the accountability climate in which teachers work affects them and their students. Teachers, of course, want their students to succeed, and yet, the intense focus on students' performance on high-stakes tests often makes teachers feel like they are under the microscope. A Bronx middle school language arts teacher expressed this sentiment when stating, "Yes, I feel pressure as a teacher, more because I want my students to do well. I don't want them to fall behind." [TSC4]⁵ Reflecting on the test results, another sixth-grade language arts teacher from the same Bronx middle school in a separate interview questioned his own teaching, "They bombed in this area that we worked so hard on! What does that say about me?" [TSC2]

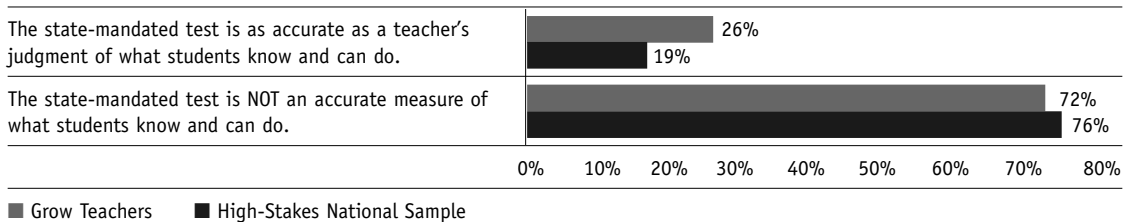
Many teachers shared that their students are aware of the importance of state-mandated tests and feel the pressure as well. "The kids know that there is a lot riding on the test," [TINT13] said a fourth-grade teacher in Harlem. Hence, teachers feel responsible not only for preparing students to take the test but also for helping students to manage any anxiety that they might feel. As a fifth-grade teacher in Canarsie commented, "[My students] do not want to fail. I think that they are really nervous, I try to comfort them." [TINT16] A fourth-grade teacher on Manhattan's Upper Westside explained the pressure on her students to score well on the state-mandated test so they will be accepted into a good middle school. Even though she felt nine-year-old students are too young to experience this kind of pressure, she had resolved that supporting them was part of her job. "If I do not do everything in my power, then the students cannot compete," she said. In the same breath, however, this teacher lamented the weight given to a single point in the academic year and students' careers. "[The test] tells you how kids do for a moment in time, and that moment counts for a lot. [Therefore] do as much [test prep] as you can because it counts for a lot. It makes me sad." [TINT19]

⁵ These codes represent internal references to interview data sets.

Teachers' Attitudes Toward The Test as a Measure of Students' Abilities

Despite teachers' acceptance of testing and their responsibility in terms of test preparation, they voiced skepticism about the test's educational merits and concern about its growing role in education. The majority of Grow and national study teachers questioned the test's accuracy in measuring students' academic abilities. Seventy-six percent of the national sample, and 72% of the Grow sample agreed with the statement: "The state mandated test is not an accurate measure of what my students know and can do." Though more Grow teachers questioned the accuracy of the state test than their national counterparts, more Grow teachers (26%) also agreed with the contrary statement: "The state-mandated test is as accurate as a teacher's judgment of what students know and can do," than the national sample of teachers (19%).

Figure 2. Percent of Grow teachers and the high-stakes national sample agreeing with the following statements (Grow sample, N=96).



To understand more fully teachers' attitudes regarding testing, we probed in interviews teachers' beliefs about the test's reliability and validity. Since teachers used a wide variety of terms and definitions when referring to these issues, we grouped them based on the following: statements about student performance were categorized as reliability concerns and statements regarding the test, test items, and the information covered by the test were categorized as validity concerns. Most of the teachers touched on one or both of these concerns. Despite these reservations, the majority of teachers said that they were generally not surprised by their class' or individual student's results, and when they were, it was typically only in a few, particular cases, as described by a fourth-grade teacher in Harlem:

There weren't that many discrepancies. I mean, usually they were where I thought they would be. It's been one or two people that I kind of thought would be a little higher or something, but they scored poorly. It kind of brought into question, for a couple of kids, whether they were bad test takers...Or [some] did better than I thought. [TSC24]

When teachers talked about reliability issues, they primarily spoke about the kind of "test taker" a student was, or whether a student was having a "bad day," which would affect their performance. Several teachers also expressed concern that some students "test poorly" and would, consequently, be placed in activities below their level, while other students "test well," scoring higher than their in-class understanding of concepts indicates and making them ineligible for the academic supports needed to adequately master the material. A few educators pointed to examples of students who

were so overtaken by stress when taking the test that they broke down in tears or even became ill, obviously affecting their performance. Most teachers explained that when they doubted the reliability of the data, they employed their own strategies to assess a student's actual ability.

Teachers had a wider range of concerns regarding the test's validity. Teachers voiced apprehension about the test's face validity, questioning whether the test measures the "life skills" that students need to succeed. A fifth-grade teacher in Harlem understood the issue in the following way:

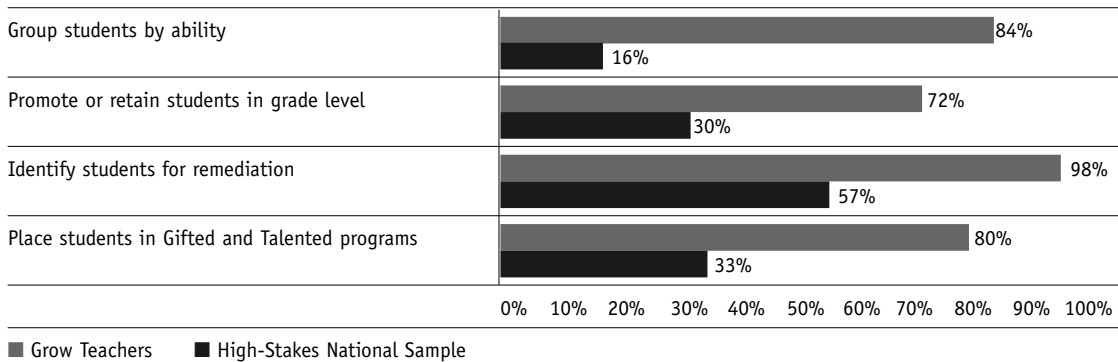
Now if you're saying that the test is supposed to measure what that child could do given those conditions, then the test obviously is reliable. It's the same test going to everyone and it's valid. Everybody is being graded the same way. So be it. However, is that really the indicator that we want to have to measure a student's level of success and progress? [TINT12]

A few teachers questioned whether all of the test items were "culturally sensitive" to all students or "developmentally appropriate." A fourth-grade teacher in Harlem claimed that the state-mandated, fourth-grade test was "too high" for fourth-grade students, asking them to answer questions and demonstrate skills for which they are not yet developmentally prepared. However, a significant portion of the Grow teachers said that they believe that the tests are aligned not only to the New York State Standards (86%) but also to what they teach in their classrooms (79%).

Teachers' Attitudes About the Use of Test Results

With increasing pressure on teachers to not only raise test scores but also utilize test score data to make instructional decisions, we chose to investigate teachers' attitudes about the different ways in which test data is used. In accordance with the view that accountability is one of their responsibilities, the Grow respondents were more comfortable than the national sample using standardized test data to hold students accountable. Ninety-eight percent of the Grow sample saw the use of the data to identify students for remediation as appropriate, as compared to only 57% of the national sample. The majority of the Grow teachers also considered it appropriate to use test data to place students in Gifted and Talented programs (80%), to promote or retain students (72%), and to group students by ability appropriate (84%), while national survey teachers were less comfortable using data for such placement (33%), promotion/retention (30%) and grouping (16%).

Figure 3. Student Accountability Items: Percent of Grow teachers and the high-stakes national sample who feel the test results are appropriate for evaluating students in the following ways (percent answering “moderately/very appropriate”) (Grow sample, N=96).



Overall, the majority of Grow and high-stakes states teachers are not comfortable using test data to evaluate teacher and administrator performance; however, Grow teachers are slightly more comfortable than their national counterparts. While two-thirds of the national sample do not consider standardized test results an appropriate way to evaluate teachers' instruction, Grow teachers were less dismissive, only 53% of Grow middle school teachers and 60% of Grow elementary school teachers shared this view. Grow teachers were essentially split on the issue of whether the test data can be used to hold schools accountable as only 49% of elementary teachers and 52% of middle school teachers agreed. Among the national sample, elementary teachers (60%) and middle school teachers (56%) both felt that holding schools accountable was appropriate. Grow and national sample teachers differed substantially on whether test data could be used to rank schools publicly. Two-thirds of elementary school teachers and 60 percent of the middle school teachers in the national sample felt that ranking schools publicly according to students' scores is an appropriate use of the data, compared to one-third of the elementary school teachers and nearly half of the middle school teachers in the Grow sample. The range of school settings captured in the national sample could explain this difference. While the Grow sample draws upon one large urban school system, the national sample includes schools in urban, suburban and rural settings. The national sample's diverse mix of schools no doubt captures data from high-performing suburban school districts where rankings are an important part of maintaining the school's high-achievement profile, and, therefore, a use of data that teachers are more accustomed to seeing.

Table 2: Percent of Grow teachers and the high-stakes national sample who feel tests are appropriate for evaluating educators and schools in the following ways (Grow sample, N=96).

EVALUATE TEACHER OR ADMINISTRATOR PERFORMANCE			
	Appropriate	Inappropriate	Total
Grow elementary teachers	40	60	100
High-stakes elementary teachers	37	63	100
Grow middle school teachers	47	53	100
High-stakes middle school teachers	32	68	100
HOLD SCHOOLS ACCOUNTABLE			
	Appropriate	Inappropriate	Total
Grow elementary teachers	49	51	100
High-stakes elementary teachers	60	40	100
Grow middle school teachers	52	48	100
High-stakes middle school teachers	56	44	100
RANK SCHOOLS PUBLICLY			
	Appropriate	Inappropriate	Total
Grow elementary teachers	33	77	100
High-stakes elementary teachers	62	38	100
Grow middle school teachers	47	53	100
High-stakes middle school teachers	58	42	100

Administrators' Attitudes Toward High-Stakes Testing

With its adoption becoming more widespread through NCLB, standardized testing is not only being used to directly measure students' academic progress but also to indirectly assess administrators' leadership. Administrators now operate under intense pressures to either ensure that their schools or districts make adequate yearly progress or face their own set of career-challenging consequences. Since these pressures are likely to remain a constant influence on their work, we sought to investigate their attitudes toward high-stakes testing, their perceptions of the pressures, and their perspectives on the appropriate uses of standardized test data. Based on data collected in the survey and the interviews, we identified three themes (similar to ones identified with teachers): 1) administrators' attitudes toward standardized testing 2) administrators' attitudes about the test as a measure of students' abilities, and 3) administrators' attitudes about the use of test results. We describe and discuss each of these themes below, making note of compelling differences between teachers and administrators when they exist.

Administrators' Attitudes Toward Standardized Testing

As with the teachers, our research sought to understand administrators' attitudes toward testing and accountability. Similar to the teachers, administrators acknowledged that standardized tests and accountability pressures are a constant part of education. As one principal at an elementary school in Canarsie commented, "I like the test, and I like the idea of the test, but I don't like the amount of pressure on the test results" [SADINT19]. Across the board, administrators say that accountability pressures have grown in popularity and are changing the ways in which they think about school performance.

At an elementary school on the Upper West Side of Manhattan, the principal remarked, "This year there is a greater emphasis on the culture of testing and data. This year is different. Everywhere you look, there's testing" [SADINT1]. As part of this "culture of testing and data" she gave each teacher a black binder in which they were asked to keep assessments and data regarding the academic performance of each student. The principal acknowledged that there are pros and cons to a focus on high-stakes testing: "What's good about it? You know more about the test. What's bad about it? You are teaching to the test" [SADINT1].

Another principal in a Harlem elementary school, commented that the tests help focus teaching and learning: "It forces my teachers to focus on what needs to be done within a time frame. If you don't know where you are going, then how are you going to get there?" [SADINT5]. Statements such as these underscore the degree to which testing has become an integral part of administrators' everyday worlds.

Building-level administrators acknowledged that testing holds everyone accountable and test data is the lens through which educators at every level of the system are judged, as one principal in a Harlem elementary school explained:

I do say to the teachers that this is also my basis for evaluating the success of their instructional program. And I do say that my superintendent also looks at data. When I sit with her I am also evaluated. There is a review. I have goals I have given her. My goals are to get ninety percent of my kids both in reading and math to meet those standards, and I've written that in black and white and that's the basis that I am [evaluated] on. The pressure is on you. It's on me too. [SADINT7]

Thus, from the perspective of New York City administrators, everyone – students, teachers, and administrators – is being held accountable to the same test outcomes.

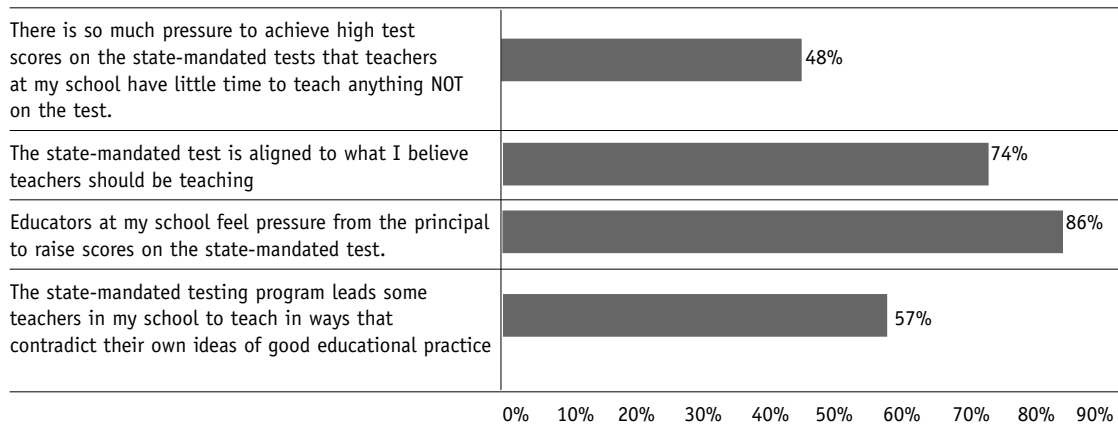
Administrators' Perceptions of the Test as a Measure of Students' Abilities

Although testing is a reality they must live with, administrators, like teachers, have some reservations about what the tests are measuring. One principal in a Harlem middle school observed that tests "are a reality that you can't get away from. It's nerve racking. Some teachers have great classroom environments, and that's not shown on a test" [SADINT4].

Administrators questioned the validity and reliability of the test in their responses to the survey; close to two-thirds (64%) of administrators do not consider the test as accurate as a teachers’ judgment of what students know and can do. For the most part, however, administrators were divided over most issues related to state-mandated tests. Administrators split over whether or not the state-mandated test were an accurate measure of what students know, a little under half (49%) of administrators perceived the state-mandated test as not being an accurate measure of what students know and can do. Close to half (48%) of administrators felt that test pressure narrowed the scope of the curriculum. Fifty-seven percent of administrators went so far as to report that the tests encouraged teachers to teach in ways that contradict their practice.

Despite this ambivalence about the test’s effect on curriculum scope and teaching practice, administrators did feel that the state-mandated test is aligned to what teachers teach in the classroom (74%). Administrators also widely acknowledged that they place pressure on the teachers to raise test scores – 86% agreed that teachers in their building felt pressured. Interestingly, this percentage was larger than the percentage of teachers (76%) who acknowledged this same pressure. We can only speculate that the difference reflects the teachers’ immersion in the day-to-day activities of classroom instruction, which may provide some insulation from the test pressure felt on an administrative level every day.

Figure 4: Percent of Grow administrators agreeing with the following statements (N=137).

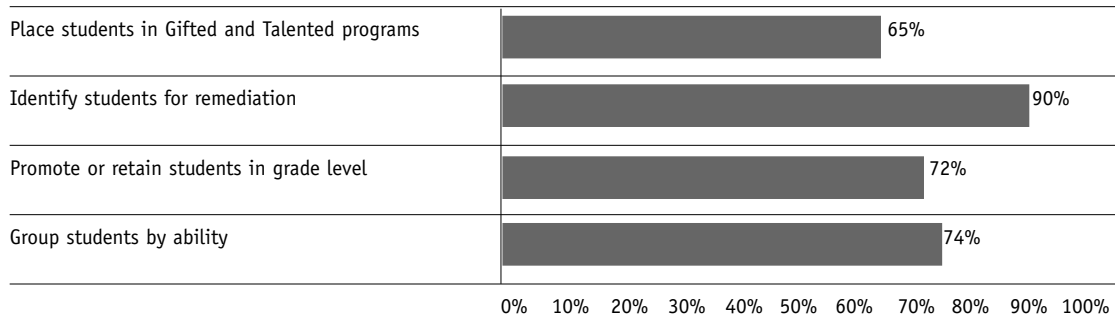


Administrators’ Perceptions about the Use of the Test Results

Administrators were more circumspect than teachers about using the results for making decisions about how to track students academically as slightly smaller majorities of administrators supported all but one use of test results (promoting and retaining). Ninety percent of administrators feel test results are appropriate for identifying students for remediation and 65% agree with its use for identifying gifted students. Close to three-quarters (74%) of administrators are comfortable with the use of test data for ability grouping. Nearly three-quarters of administrators (73%) also feel that testing data can be used for purposes of promoting or retaining students.

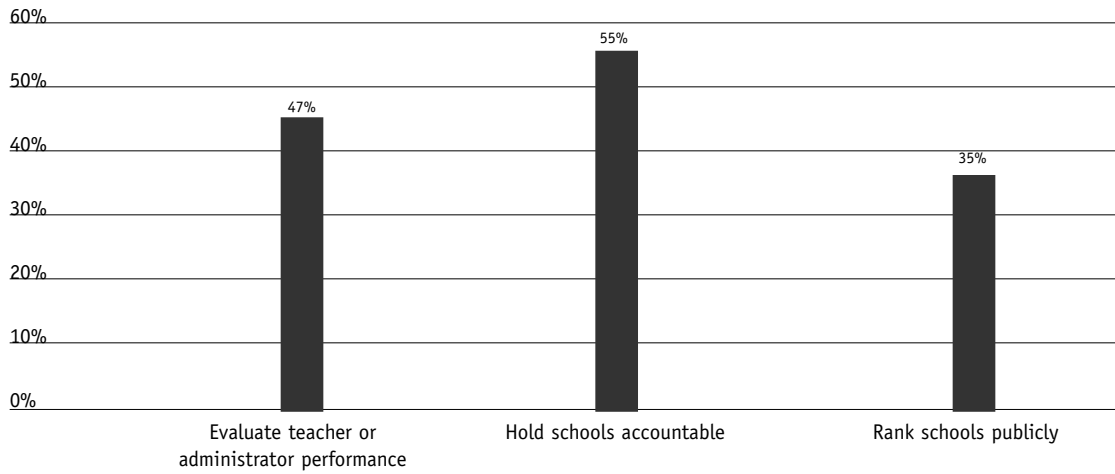
Figure 5. Percent of Grow administrators who feel the test results are appropriate for evaluating students in the following ways (combined percent answering “moderately/very appropriate”) (N=137).

Administrators split on whether or not the tests are an acceptable way to judge teachers or admin-



istrators with less than half (47%) of administrators believing they are acceptable. They also split on school accountability with 55% of administrators considering this appropriate. Nearly two-thirds of administrators were uncomfortable with publicly ranking schools by test results as only 35% considered this an appropriate activity.

Figure 6. Percent of Grow administrators who feel the test results are appropriate for evaluating educators and schools in the following ways (combined percent answering “moderately/very appropriate”) (N=137).



EDUCATOR PERSPECTIVES ON ASSESSMENT PRACTICES

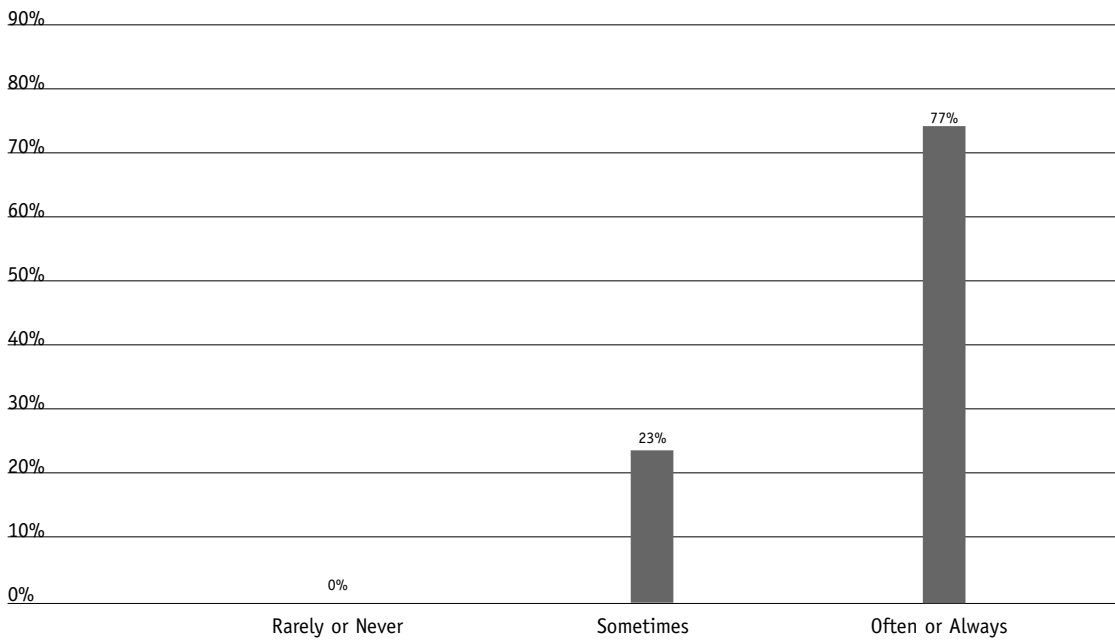
Teacher Classroom Assessment Practices

To further situate the research about Grow, we sought to understand the strategies teachers use to assess what their students know and can do, and how they use the results to inform their classroom practice. We also wanted to understand whether the accountability pressures they feel influence classroom practice. We further explored issues of test-preparation and instructional strategies that related to accountability. Sections of the interviews, scenarios and surveys, therefore, focused on how teachers know their students, what they do with that knowledge, and how they handle assessment in a high-stakes environment. We used a broad definition of assessment to refer to all teacher or student activities that seek to provide information that can be used as feedback to modify teaching and learning activities (Black, 1998). We also distinguished between classroom assessments, which are teacher-initiated, and those external assessments provided or mandated by a district, state, or external agency.

Using multiple classroom assessment strategies

All of the teachers reported using multiple assessment strategies: 77% “always” or “often” use multiple strategies to assess their students’ learning of a single topic or unit, while 23% “sometimes” use multiple assessments. These results mirror findings from the interviews, during which, approximately three-quarters of the teachers interviewed discussed using multiple assessment strategies, while a quarter discussed using only formal assessments.

Figure 7. Frequency with which teachers use multiple measures to assess student learning of a unit or topic (N=95).



The substantive issue around classroom assessment, however, is not frequency of use but how teachers utilize this information. An early review of the literature on assessment (Black 1998) indicates that even though teachers may assess students frequently, they do not always use the results to support learning. The interview responses provide a fuller understanding of both which types of assessments teachers are using in New York and how they are relating classroom assessments to addressing students' learning needs. In the interviews, teachers were asked how they assessed student progress throughout the year and which assessment strategies they use to verify student's test results. In analyzing the interviews we distinguished between those teachers who only discussed using formal assessments, whether teacher made or commercially produced, and those teachers who discussed using a broad range of both formal and informal assessment strategies.

Only a handful of teacher said that they rely solely on formal assessments. The experience of one fifth-grade teacher in Harlem illustrates this approach. She uses a structured assessment that mimics the standardized tests. "After we do our novels, I make up my own test," she said. "I do multiple choice simply because I want them to get used to multiple choice sections because that's what the tests do, and, I time them." [TINT12] The majority of teachers that we spoke with, however, use multiple assessment strategies to monitor their students' current knowledge and abilities. Overall, these teachers discussed using assessments linked generally to their teaching, either in terms of identifying students' needs or progress. According to these teachers, mixing various assessment strategies provides a fuller picture of student understanding and learning than the snapshot provided by relying solely on external assessments.

I use a lot of kinds - from quantitative assessments like multiple-choice questions to writing [...] For the standardized tests, I use the same rubrics on the test and grade them from 1-6 and that's basically it. The most valuable to me is the open-ended writing. You can really see their thinking and where they need the most help, where they're lacking in skills.

Eighth-grade language arts teacher in Bronx [TINT2]

I do a combination of project-based learning and other types of authentic assessments where they do experiments...working together in groups to create something, as well as doing multiple choice and long and short responses based on reading time... I do a great deal of assessment, ongoing, on a daily basis. So there's a number of different assignments that I do during the day that gives you a good idea of how students progress over time, both as an independent student and as a cooperative student.

Fourth-grade teacher in Harlem [TINT13]

After reviewing the wide variety of assessment strategies that teachers identified during the interviews, we grouped them into the following categories:

- Open-ended performance tasks, including student writing samples, projects, presentations, experiments and portfolios
- Observation-based assessments, including classroom participation, student discussions and read alouds
- Formal (structured) assessments, including teacher-made quizzes and tests
- Interviews and teacher-student interactions, including one-on-one conferences and questioning strategies
- External assessments, including simulations, end-of-chapter tests and commercially-packaged exams

All of the major forms of classroom assessment mentioned in the literature were covered by this group of interviewed teachers (Shepard, 2000). Many of these assessments were designed and implemented by the teachers, who then collected and analyzed the data.

Instructional Practice in a High-Stakes Context

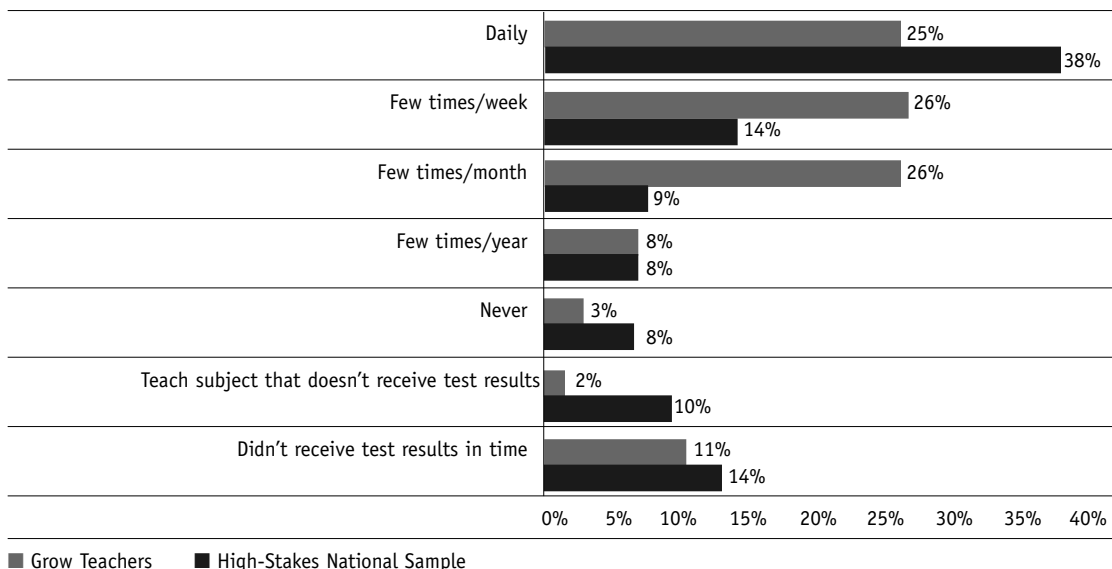
While most teachers employ a variety of classroom assessments to gauge their students learning, they all recognized that one assessment—the state test—holds greater weight and influence than the other assessments. In interviews and survey responses, they expressed concern about the accountability environment's impact on instruction. A sixth-grade teacher in Harlem stated, "You really do start teaching to the test. You try to teach them [the students] those life skills but you do begin to focus on the test." Another eighth-grade language arts teacher in the Bronx commented:

The most the [data] will tell you is that the state test, the standardized test and multiple choice questions are not very valuable in the long run but every school, on the other hand, is basically evaluated and graded by those tests, so it seems like the only thing you are doing is just preparing them for this isolated test to make your school look good, which is a shame because it's not really a very good assessment of the kids' skills and abilities. [TINT2]

Another teacher who works in an elementary school only recently removed from the Schools Under Registration Review (SURR) list, where the pressure to raise test scores pervades, remarked, “It [teaching] is not what I envisioned. I thought they [the students] would be learning to know rather than to take the test” [TINT15].

Three quarters of both the national sample and the New York City teachers reported feeling that the tests lead them to teach in ways that contradict their own ideas of good teaching (75% and 76%). In the interviews, several teachers highlighted the challenges in balancing test preparation with the curriculum. Approximately three-quarters of the teachers in both samples said that the test results have an actual influence on their teaching on more than a monthly basis, with 38% of the national sample and 25% of the Grow sample saying that the test affects their teaching on a daily basis.

Figure 8. Frequency with which teachers feel test results influence their teaching (Grow sample, N=93).

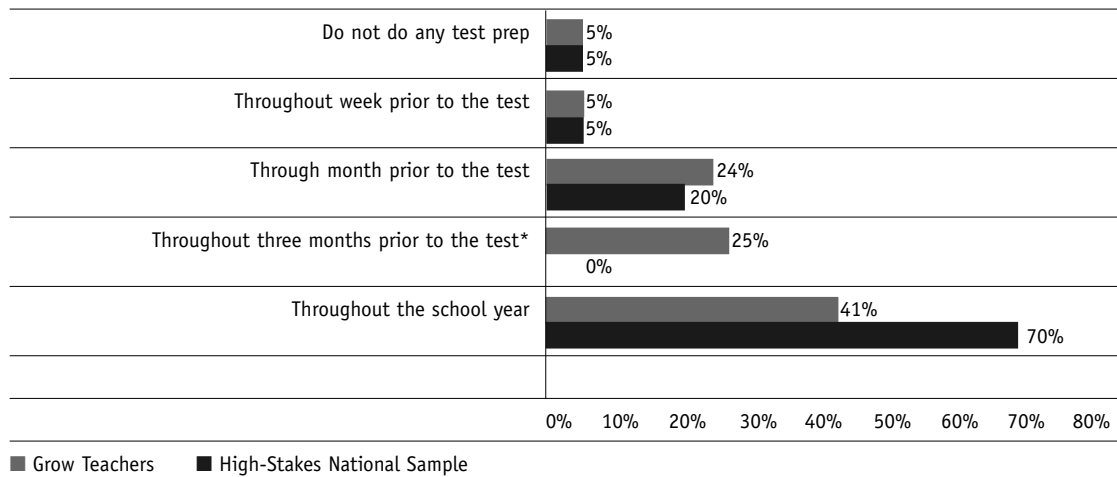


Both studies surveyed teachers on how and when they prepare their students for the test. The items about the length of time dedicated to preparing for the test are not fully comparable, since the national study did not offer the option of “three months prior to the test.”⁶ However, if we combine the last two categories (“throughout three months” and “throughout year”), the two

⁶ The “Three months prior” category was added to the Grow Survey because New York City begins testing in January – which only five months into the school year, and many teacher stop explicit test preparation after the tests.

populations are similar in the amount of time they dedicate to explicit test preparation – 66% of the Grow sample engaged in test prep with their students for three months or throughout the year, while 70% of the national sample engaged in test prep throughout the year.

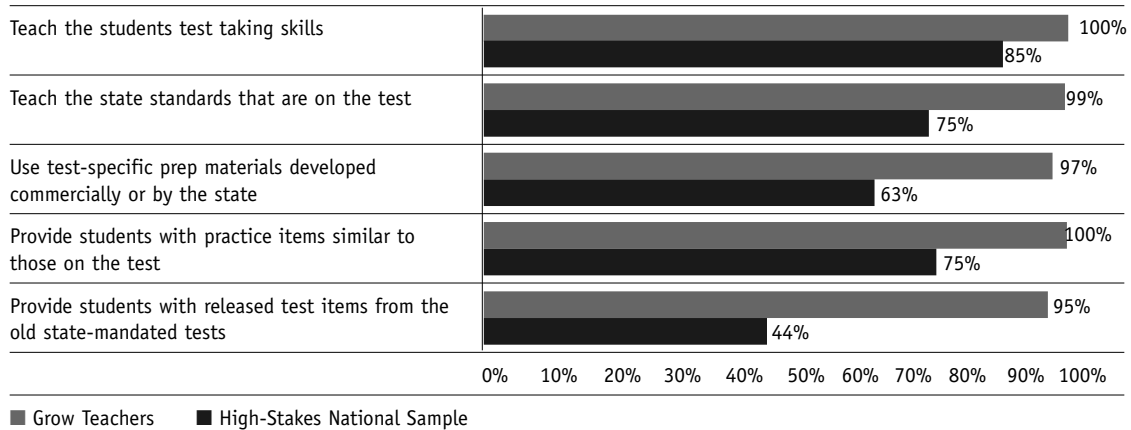
Figure 9. Amount of time Grow teachers and the high-stakes national sample engage in explicit test preparation (Grow sample, N=96).



* “Throughout three months” category was not a category box available to high-stakes national sample

The Grow teachers use a wider range of methods to prepare their students for the state-mandated test. Of the five methods offered, more than 90% of the Grow respondents used all of them. Notably, 97% of the Grow respondents use commercially prepared test preparation materials, compared to 63% of the national sample of teachers. It is important to consider that some of these strategies (like the use of simulation tests) are the result of district decisions and are not under a teacher’s control.

Figure 10. Percent of Grow teachers and the high-stakes national sample who use the following methods to prepare students for the test (Grow sample, N=96).



From the interviews, we expanded our understanding of what “test prep” means and how it varies from classroom to classroom, school to school, and even, district to district. Some teachers give their students weekly quizzes with timed reading and writing sections and math problems; others on a daily basis assign their students skill-based activities found in commercially published test preparation booklets. At some schools, the most struggling students – students below the proficient level – attend after-school classes, which focus on remediation and test prep activities; at other schools, administrators encourage teachers to provide individualized instruction to students who have been identified as those on the cusp of achieving proficiency, with the hope that they will move up.

Perceived utility of the assessments

To gain a full picture of how teachers prioritize between different assessment strategies, we surveyed teachers about the utility and the value of a variety of classroom assessments. Overall, the teachers felt that teacher-made assessments and strategies were more useful than the external assessments. For example, 65% percent of respondents felt that authentic assessments, teacher-made quizzes, and classroom participation were very useful strategies, while 35% and 31%, respectively, felt that diagnostic tests and published exams were very useful assessment strategies. The respondents indicated a preference for assessment strategies like classroom participation because they offer students multiple ways to communicate what they are learning and doing.

Figure 11. Grow teachers' perceptions of the usefulness of specific classroom-based assessment strategies (Grow sample, N=96).

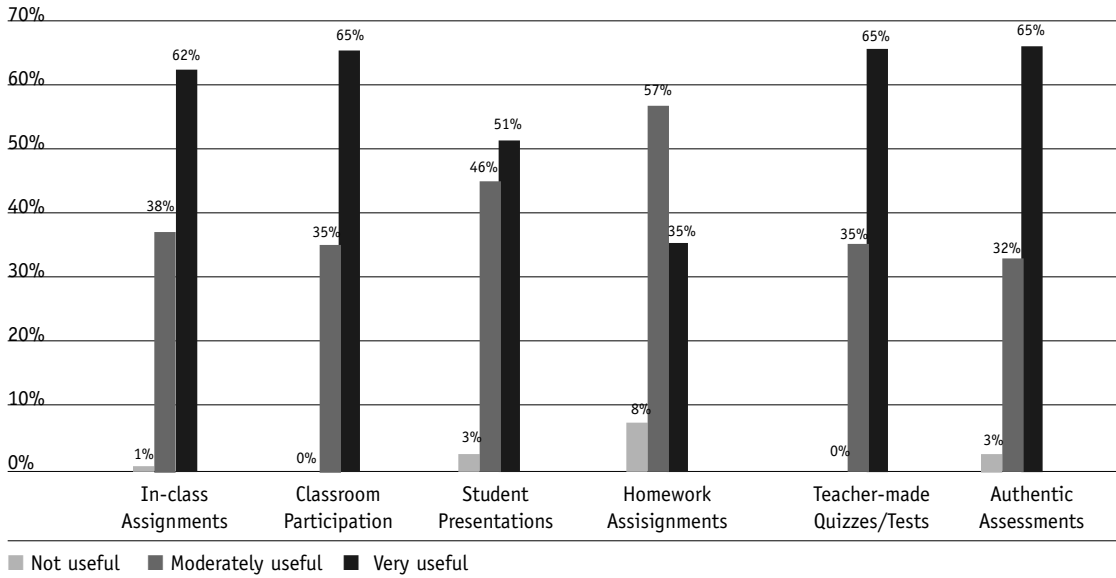
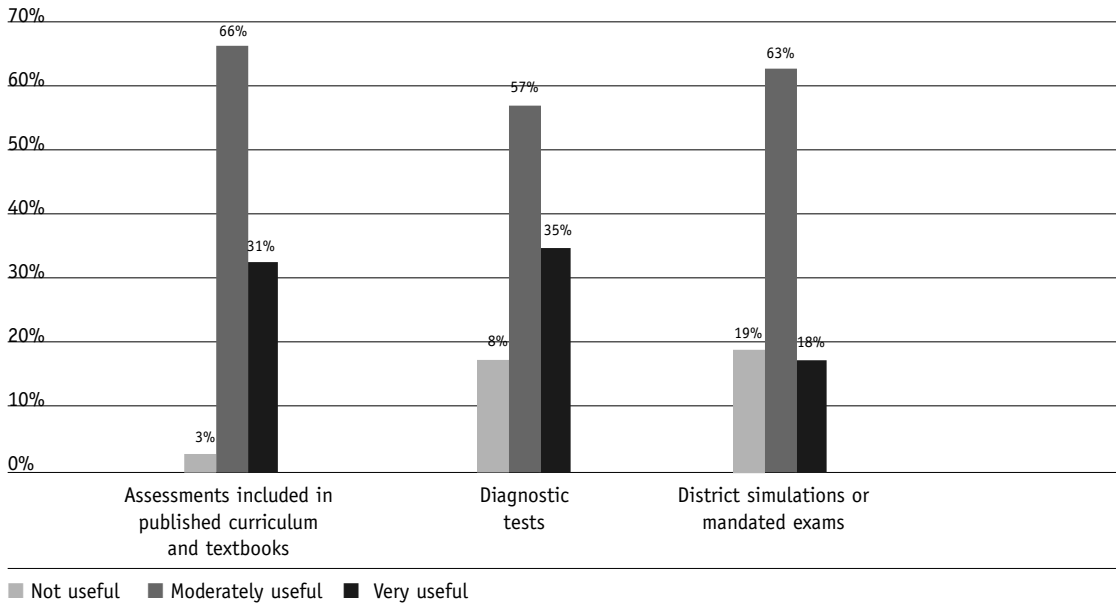


Figure 12. Grow teachers' perceptions of the usefulness of external assessment strategies (Grow sample, N=96).



The teachers were not asked to rank assessment strategies during the interviews; however, several teachers said that they appreciated open-ended and observational assessments, through which they could hear their students grappling with concepts and ideas as well as see them putting skills into practice. The following comment exemplifies what teachers find valuable about these types of assessment strategies:

If they can explain it, you know they understand it. Also, I watch them pair up with somebody. If they can show someone else, then you know that they understood.

Sixth-grade teacher in Harlem [TINT14]

The teachers we interviewed factored many different kinds of data and assessments into making instructional decisions. They expressed the need to monitor different aspects of student development and learning, not all of which could be measured or evaluated with a single assessment. This is why many used other assessment strategies. They spoke clearly on the relative importance, as well as the limited meaning, of standardized test results in assessing their students' progress. They often pointed out concerns that standardized tests measure one point in time and one set of items. As educators, they said they felt it necessary to consider multiple aspects, from personal and social development to intellectual and academic growth. As one teacher made clear, they have many sources of data, but they also need to consider the child who is in their classroom every day.

If you are an experienced teacher you know that you cannot take this data and pigeonhole your children [...] So I would look at Grow and look at Gary. As a teacher I have my curriculum, I have my Grow chart, I have Gary's report card and his student profile sheet. I have all of this and I have Gary. Most importantly I have Gary.

Fifth-grade teacher on the Upper Westside [TSC14]

Administrators' Assessment Perspective

Administration support for classroom assessment

To broaden our perspective on classroom assessment in a high-stakes setting, we also sought to understand how administrators perceived and supported classroom assessment practices. We asked building level administrators during our interviews about their views on classroom assessments. Because they tended to look more broadly across their entire school system rather than at the classroom level, our conversations did not focus as much on classroom assessment practices: however, these discussions did yield useful insights into how administrators approach this subject. A principal on the Upper Westside remarked that there was a greater emphasis on "the culture of testing" this year. In relation to teachers using test data, she commented that many of her teachers were "not sure they wanted to go there, but they realize they have to," even though they resisted a school culture driven by a test "that happens once a year" [SADINT1].

The administrators clearly felt pressure to improve test scores and most of their interviews focused on how they support teachers' testing needs and how the growing accountability culture has influenced their schools and their own decisions.

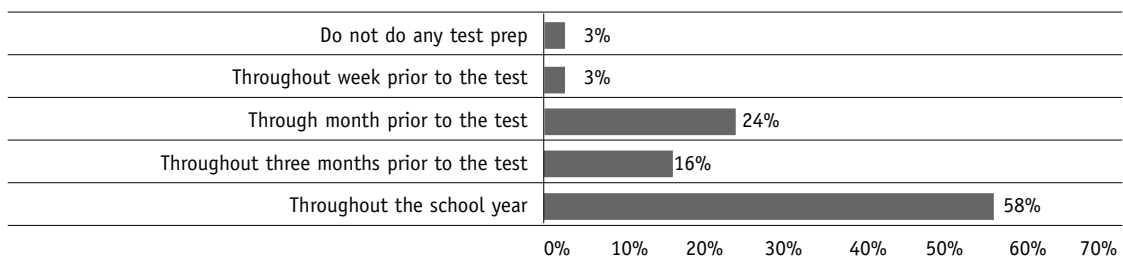
A Deputy Superintendent commented that his district uses the data to pressure principals to help the whole district meet accountability demands, an approach that requires principals to target students near the proficiency level with extra resources and supports. Though the district has shown substantial short-term gains in math scores, the superintendent suggests that deeper improvement will require a “mutual commitment” between principals and district leaders focused on wide-scale reform.

While the state test certainly commanded much of their attention, administrators, like teachers, largely expressed a desire to assess student progress from different angles in an ongoing fashion. Three-quarters of administrators spoke about encouraging their teachers to continually assess their students using multiple approaches. Only one-quarter spoke about using formal assessments, exclusively. A middle school principal in the Bronx commented, “I want teachers to look at assessment as ongoing, not as summative, but to embrace assessment [...] It doesn’t make sense for someone to wait six weeks or even one week to find out that a child has an issue” [SADINT15]. All of the administrators interviewed were concerned about monitoring student progress. The few administrators who didn’t discuss using multiple assessment strategies did require that student progress be assessed throughout the year, albeit using formal assessments. One elementary school principal explained that teachers give monthly review tests, bi-weekly simulations and a diagnostic test twice a year.

Test preparation

Administrators were asked how long before the test did their school require teachers to engage in explicit test preparation. Fifty-eight percent of administrators indicated that they require teachers in their school to do test preparation activities with their students throughout the entire school year. The administrators’ willingness to weave test preparation into the school year contrasts with the teachers’ overall reluctance to do so. Only 41% of teachers supported integrating explicit test preparation throughout the entire year.

Figure 13. Amount of time that Grow administrators encourage their teachers to engage in explicit test preparation (N=135).



Despite encouraging teachers to use multiple assessments, a significant majority (91%) of the administrators surveyed reported that, in order to prepare students for the test, they encourage teachers to “teach the students test-taking skills.” The interview data further elucidates this survey finding. It suggests that test preparation included teaching students how to take the test, how to become a good test taker, and what type of information is on the test – skills that could be taught throughout the school year. As an assistant principal in a Bronx elementary school explained: “How to take the test is another ball game. It’s not just about question and answer because a student needs the knowledge to be a good thinker,” then come the rules of test prep (i.e. introducing a child to test format (filling in bubbles), the process of elimination, reading directions). She says that being a “test taker is different from learning and knowledge. You need stamina for reading and focus, conceptual sense” [SADINT8].

Likewise, a staff developer from a Bronx elementary school explained her view of test preparation as helping students succeed:

I have to do a lot of traditional test prep, which the district doesn't offer ... so we do that in-house. But I think that's one of the reasons that we've done what we have...We need to do test prep and it's built in –the literacy people, the math people, we do test preparation...we built that into the program so that ... for the last two months, part of every session is a test taking strategy of some sort. Because that's the only way to succeed. And my feeling is that until they line up the assessment with the instruction- There'll always be that. Because if you wanted them to do well on tests, you did the practical aspect, is you must do test preparation.

[SADINT10]

The interviews with the staff developers, in particular, opened a window into how the school administration reacts to accountability demands and teachers’ testing needs. In their descriptions of their duties, staff developers generally identified two central tasks that relate to the tests: providing teachers with support for tested skills and working directly with students in pull-out sessions. One staff developer commented that her role is “to raise the scores, boost the math scores.” A math staff developer in the Bronx reported how the test calendar divides her work with the ELA coach: “Before February, everything is reading, and then I go in for the math between February and May. We have little groups and direct instruction” [SADINT10]. Later in her interview, she commented on how she felt accountability demands required certain sacrifices:

“Yeah, some [students] will be left behind, but hopefully they'll catch up and be pulled forward, you know, teachers can't worry about making it individualized. It's not practical with a group of 32 to 33 kids and you have a test to take. You have to cover the material”

[SADINT10].

Her school had recently become one of the top scoring schools in her district.

EDUCATOR PERSPECTIVES ON THE USE OF GROW

Teachers' Reported Uses of Grow Reports

During the qualitative phase of the research, we interviewed educators at every level of the New York school system as well as observed grade-wide meetings and professional development sessions offered by schools, districts, the Grow Network, and other external organizations. We spent a considerable amount of time talking with classroom teachers about their use of data in classroom decision-making and more specifically, their use of the Grow Reports®. Later, we developed a survey—based on much of what we learned through these conversations and observations—to validate our findings with a larger sample. In this section, we report on what the teachers and administrators told us about how they use the Grow Reports® in their schools and classroom.

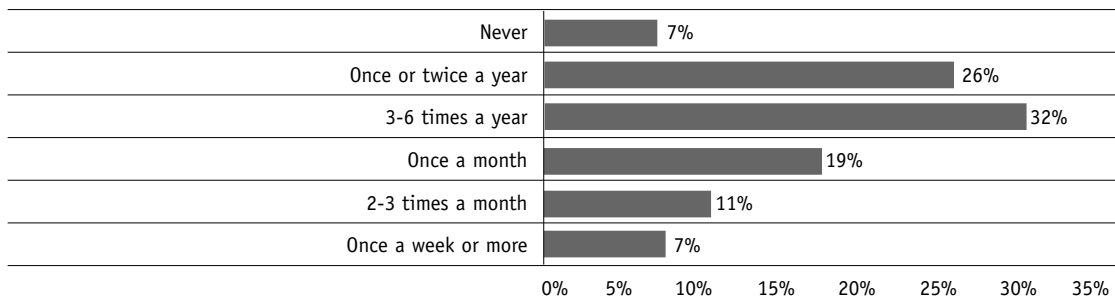
In October, all elementary and math and language arts teachers in grades four through nine received the paper version of the Grow Reports® as well as a password for the extended online version. The Grow Reports® provide each teacher with his or her current students' standardized test results from the previous year. The test results are broken down into to the level of the skill tested for each student. The online version also supports teachers' analysis of the information with a few additional features: access to explanations of standards, access to instructional materials, individual student reports, and "Flexible Groupings" of students, which group students by performance in each standard into three categories of performance. The online reports contain hotlinks to material that defines each skill or standard and explains challenges students face when mastering this skill. The class priorities lists and the explanations of the tested concepts are also hotlinked to instructional materials and resources for teachers and administrators and to external resources approved by the New York City Department of Education.

Throughout the survey and in the interviews, teachers across New York City reported using the Grow Reports® in various ways to meet their overall classes' as well as their diverse students' academic needs. Grow-using teachers discussed making decisions within the context of certain areas of their instructional practice. We grouped those 'areas of instructional practice' into the following five categories: **(1) targeting instruction**, with decisions about class priorities, yearlong pacing calendars, and weekly lesson plans; **(2) meeting the needs of diverse learners**, seen in strategies such as grouping, creating Individualized Education Plans (IEPs), and giving individualized assignments and materials appropriate to the students' levels; **(3) supporting conversations** with parents, students, fellow teachers and administrators about students' learning; **(4) shaping teachers' professional development** by reflecting on their own practice; and , **(5) encouraging self-directed learning** by giving the data to students.

Before exploring teachers' uses of the Grow Reports® more deeply, we want to mention two basic findings about using Grow. First, during the interviews, we learned that teachers found the reports clear and comprehensible. All of the teachers understood the basic aspects of the reports—the students' scores and the New York State performance levels. Teachers reported that the Grow Network, as a company, was very responsive to suggestions from them about improving the Reports.

Second, with respect to frequency, we found that about 40% of teachers used the Grow Reports® monthly, while a little less than one-third of the teachers reported using the Reports three to six times throughout the year.

Figure 14. Teachers’ reported frequency of using the Grow Report in 2003 (N=92).



Meeting the Needs of Diverse Learners

Noting that the Grow Reports® provided them with more information about students than what they had access to previously, teachers said that they looked at the data mostly to know “where their students are.” In the interviews, teachers commonly understood the test results in relative terms of their students’ strengths and weaknesses. The survey results support the interview data as most respondents (91%) reported using the Grow Reports® to determine class-wide strengths and weaknesses. Several teachers noted that they rely on the data more at the beginning of the year when they don’t know their class or students as well. A fifth-grade teacher in the Bronx said, “In the beginning of the year, before I really know the kids well, [Grow] is a good sort of first idea of what I’m dealing with and what their needs are” [TINT5].

Identifying strengths and weakness helps teachers to address long-range planning. When asked how they use the Grow Reports®, several teachers responded that they use the data when deciding where to target their instruction. These included activities such as setting class priorities and/or creating a more extensive yearlong pacing calendar as well as when planning mini-lessons to review and reinforce certain concepts or skills.

Table 3: Using the Grow Reports™ to target instruction I use the Grow Reports™ to... (N=96)
(includes “sometimes,” “often,” and “always” responses)

Response	%
Identify class-wide strengths and weaknesses (n=92)	91
Set priorities about what to teach (n=94)	89
Plan test preparation activities (n=92)	79
Make decisions about grouping (n=94)	77
Do lesson planning (n=91)	76
Learn more about the NYC standards and tested topics (n=90)	73
Decide when to teach mini-lessons (n=92)	71
Create homework assignments (n=90)	66
Make year-long pacing calendars (n=90)	51

Teachers at both the elementary and middle school levels explained that by looking at the Grow Reports®, they can see in which areas their overall class and individual students scored high and in which they scored low, and then can tune their instruction accordingly. Many teachers use the information to focus more intensely on, or include more practice in, the areas where students scored low. For example, one teacher explained that when looking at the reports, she asks herself, “Where is my class lacking?” and then, once identifying those skills, she searches for ways to thread those skills throughout her instructional program. Another teacher stated that she doesn’t rely solely on the Grow data to inform her instruction. “I spiral,” she reflected. “If they know something I will make it a homework assignment to see how much they remember and if I need to review or re-teach. I will spiral topics throughout the whole year” [TINT17].

Conversely, some teachers shift instruction time away from areas where students scored well. For example, one first year sixth-grade math teacher in the Bronx said that she likes the Grow Reports® because once she learns from the data where her students have performed well, she doesn’t have to “teach what they are stronger in” [TSC3]. But teachers did not rely solely on the test results. Several teachers said that certain concepts or skills were more difficult for students to master, so they would incorporate them into larger units or create focused mini-lessons for constant reinforcement.

The survey results supported what we learned in the interviews and give a larger picture of how widely used the Grow Reports® are for certain types of decisions. On the survey, 89% of Grow users reported using the resource to “set priorities about what they teach.” A majority of Grow users also responded that they use the Reports to inform decisions when planning lessons (76%) and when deciding upon mini-lessons (71%). A little more than half of the teachers reported that they used Grow data when creating yearlong pacing plans (51%).

In addition to using the Grow Reports® for guidance in classroom instruction, many teachers also reported using the Reports to meet the needs of individual students, especially those who are struggling. By knowing an individual student’s strengths and weaknesses, teachers claimed that they could better plan instruction specifically tailored to address student needs. By adding modifications to lesson plans, teachers said they could attempt to address the needs of both the low and high performers within a class—in other words, differentiate instruction. Teachers mentioned strategies such as grouping, creating Individualized Education Plans (IEPs), and giving individualized assignments and materials appropriate to students’ levels. The survey responses supported these observations as 89% of teachers reported using Grow to differentiate instruction. Most teachers we spoke with felt that having timely data that showed individual student performances helped support their decisions to differentiate instruction according to individual students’ needs. Teachers reported utilizing several different strategies to differentiate instruction, such as teaching using small leveled groups, assigning differentiated class work or homework corresponding to students’ needs, constructing peer-tutoring situations consisting of a low performing and a high performing student working together, and/or tutoring students one-on-one.

Table 4: Using the Grow Reports™ to support diverse learners I use the Grow Reports™ to... (N=96)
(includes “sometimes,” “often,” and “always” responses)

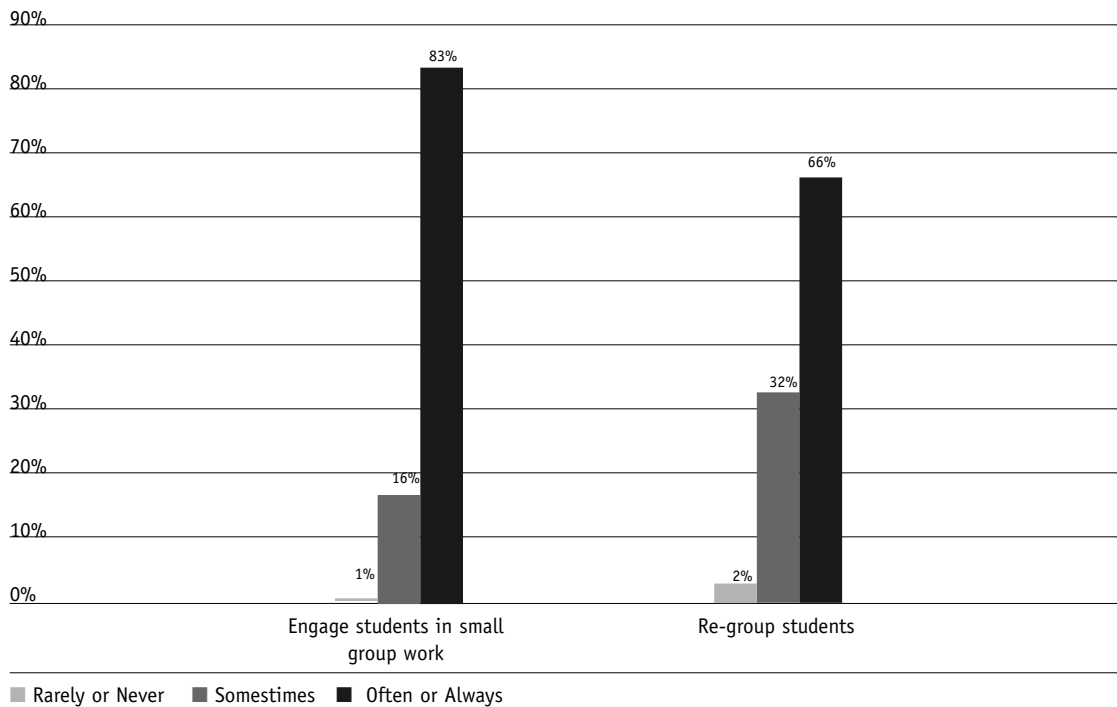
Response	%
Identify low performing students for additional resources (n=92)	90
Differentiate instruction (n=95)	89
Identify high performing students for additional resources (n=92)	83
Organize groups by ability (n=94)	81
Make decisions about grouping (n=94)	77
Identify peer tutors/teacher helpers (n=92)	71

One group of learners who often figured prominently in teachers’ interviews were the students who had scored just above or below the edge of proficiency (the top of the range for Level 2 or at the bottom of Level 3). This population was often called the “bubble kids” because of their statistical location in a bubble of scores near the cut off. Because the probability of moving these students is much higher, some teachers and administrators reported targeting them for extra resources, such as pull-outs, after-school, and special programs, in an effort “to move” them up a level or ensure that they do not slide down a level. The format of the Grow Reports® made identifying these students very easy. A first year seventh-grade language arts teacher in the Harlem said, “The difference is very important. The 615 [score level] could easily be at the next level. To get the kids’ scores to move it’s a discipline issue. If there were less disruptions in the classroom, the kids who are right at the cut offs would be able to learn what they need.” [TSC9] Other teachers handled these students differently. A fifth-grade teacher in Harlem said:

I'd probably put him in the higher group even if he placed below them so he'd be around those kids and give the extra push. So work a little bit above his level because he could probably handle it. A lot of what's done is just for a change of instruction so in other words give him what the higher kids are doing. [TSC25]

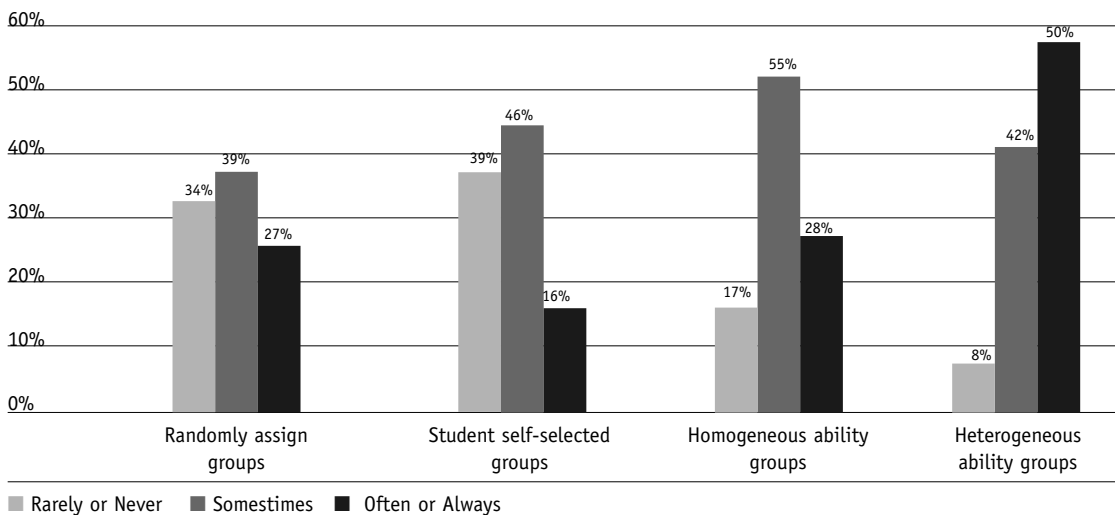
In the interviews teachers described using the Grow Reports® to support grouping by identifying where students performed relative to one another. Teachers could use the data either to group students who were performing similarly together in homogeneous groups or to group students performing at different levels together in heterogeneous groups. Grouping strategies can take on many different forms. For instance, teachers said that while there are some times when they will group students according to their ability levels, as when creating homogeneous groups, at other times, they will create groups based on criteria such as behavior, interest, or motivation levels. Some teachers said that they vary their grouping strategies frequently, sometimes even having students self-select their own group. Grow's "Flexible Grouping" feature allows teachers to create groups of students to suit all of these different grouping strategies.

Figure 15. Frequency with which teachers engage students in small group work or re-group students (N=96).



When asked on the survey if they used Grow for grouping, 81% of teachers reported using the Grow Reports® when making decisions about grouping. We also surveyed the teachers on specific grouping strategies. The most common strategies were ability grouping, heterogeneous ability grouping being the most common with 92% of teachers reporting using this strategy (50% “often” use this strategy; 42% “sometimes” do). Homogeneous ability grouping was also commonly used with 83% of teachers reporting using this strategy. Random grouping and student self-selection were less common with only 66% and 62% of teachers, respectively, reporting using these strategies.

Figure 16. Frequency of teachers’ use of grouping strategies with students (N=96).



A few teachers we interviewed said that having access to data on how students performed on specified skills encouraged them to create small groups to target certain skills. A fourth-grade teacher in the Bronx described the benefit, stating, “The good thing about the Grow Reports® is that everything is broken down, so you can have four different groups working on four different things at the same time” [TINT7]. An eighth-grade language arts teacher in the Bronx explained one of his practices:

If most of the kids in the class need help with figurative language, you do a whole class lesson, and then you take the kids who need the most help and do a small group session with them while the class is doing independent work. It doesn't always work that way because if the class is a ‘crazy’ class, then you can't really do small group instruction. [TINT2]

However, even though teachers reported that they could use the Grow data when making decisions about grouping, many were careful to note that the Grow data was one “small piece” that they considered in conjunction with other forms of data collected from a wide array of assessments, including observations, in-class assignments, daily quizzes, unit pre and post tests, just to name a

few. Moreover, teachers added that the data on the Grow Reports® has “declining value” as the year progresses, in part because students take the exam roughly six months before the reports are distributed. According to a first year, seventh-grade math teacher in Harlem, “The more time that goes by before you get it, the less value it has” [TINT11]. Elementary school teachers, especially, noted that their students change so much throughout the year that they must base their decisions about grouping and re-grouping on ongoing assessment. “If they progress, you take them out and move them into a different group,” said a fifth-grade teacher in the Bronx [TINT7].

In addition to grouping, teachers reported another way to target instruction on students’ strength and weaknesses was by providing students with a multitude of materials and manipulatives to create multiple entry points into the content. A fourth-grade teacher in Harlem remarked, “We have a number of different materials in different lengths, of different subject matter, so it’s cross content plus curriculum that enables them to have exposure and success in various different activities” [TINT13]. Teachers agreed that when selecting or modifying instructional materials or assignments, it is beneficial to know where students are in terms of the standards and skills. Some said they differentiate instruction by giving certain students modified in-class and/or homework assignments, which correspond to their ability levels. Sixty percent of the survey respondents who identified as Grow users said that they use the Reports when creating homework assignments. In interviews, some teachers also said they use entirely different textbooks or supplementary materials to work with different students. For example, one teacher used information found in the Grow Reports® to create various different math homework calendars, which would target students’ specific strengths and weaknesses.

Peer Tutoring

Teachers also differentiated instruction by creating peer-tutoring situations—where a high-performing student works with a low-performing student—in their classrooms. One sixth-grade teacher in Harlem described how she utilized peer tutoring:

Like for writing, we are doing a research project. Some kids are excelling. They have already done the outline and are ready to go on. Other kids don’t know what to do with the outline, so I pair them – the strong and the weak to help each other. [TINT14]

In addition, some teachers spoke about how they work one-on-one with students, especially with those students who need extra support, whenever they can.

Supporting Conversations

The Grow Reports® also helped bridge discussions about student learning. Most of the teachers we spoke with talked about using the Grow Reports® in conversations with teachers, parents, administrators, and students. They found the Grow Reports® were a good starting point for conversations as well as something “concrete” to show parents, administrators, other teachers, or the students themselves when discussing where the class or the student was in terms of his or her learning and where he or she needs to go. One sixth-grade teacher in Harlem said:

I explain the Grow Reports® to the kids and to their parents. I point out where they are weak and encourage them to focus on that skill. Some kids get nervous about it, but kids handle it well. Kids take the initiative and some kids take the textbook home and will do extra problems in areas where they need help. Some parents put the Grow Reports® up on the fridge. [TINT12]

The responses on the survey show a similar pattern of use. Sixty-two percent of teachers reported using Grow as the basis of conferences with parents, 52% with students, 47% with other teachers, and 38% with administrators and/or staff developers.

Table 5: Percentage of Teachers using Grow Reports™ as a basis for conferencing with...(N=96)

Response	%
Parents	62
Students	52
Other Teachers	47
Administrators and/or staff developers (“coaches”)	38

Shaping Professional Development

Several teachers said that when analyzing the Grow Reports® and identifying their classes’ strengths and weaknesses, they took an opportunity to reflect upon their own teaching practice. Three quarters of the Grow users who completed the survey (77%) reported that they use the Grow Reports® to reflect upon the effectiveness of their own instruction. Teachers explained that seeing, for example, that the majority students scored low on a skill, such as cause and effect, would cause them to reflect upon how they taught that specific skill. Some teachers reported that by looking at the Reports they realized that they weren’t even teaching some of the standards and skills on which the students were tested. Specifically, many teachers told us that they became aware that they were not explicitly teaching their students the concept of author’s purpose, one of the fourth-grade English language arts standard concepts. Furthermore, several new teachers, having taught for three years or less, said that they referred to the Grow Reports® to learn more about what to teach. A fourth-grade teacher in Harlem remarked, “[Grow] motivated me to go and research it on my own. I printed out all of the skills for ELA and math. I wanted to know which skills the students needed to know” [TINT15].

Table 6: Using the Grow Reports™ to support professional development decisions I use the Grow Reports™ to... (N=96) (includes “sometimes,” “often,” and “always” responses)

Response	%
Reflect upon my own teaching (n=92)	77
Learn more about the NYC standards and tested topics (n=90)	73

Supporting Self-Directed Learning

Another interesting use that emerged during the fieldwork was the dissemination of the data to students as a way to encourage them to take ownership of their own learning. Since the Grow Network creates a parent report, many teachers spoke about giving Grow Reports® to the students to take home to their parents. A small but sizeable group of teachers talked about sharing the Reports (or the data from the reports) with their students so that students were not only aware of their performances on the test but were also encouraged to take responsibility in terms of their own academic progress.

While all of these teachers hoped that sharing individual data with their students would help them to take ownership over their own academic performance and learning, the ways by which teachers did so varied. One school provided binders for each student to keep his or her Grow Reports®. Some teachers in that school simply distributed the Reports to students without structuring any time to discuss them with students. They instructed students to look at and keep them in the front of one of their binders or notebooks. Other teachers dedicated time to discuss the Reports' structure with their whole class, explaining what scores mean in addition to offering students strategies for improving their scores. A sixth-grade teacher in Harlem described what she did:

I explain their Grow Reports® to the kids and to the parents. I point out where they are weak and encourage them to focus on that skill. Some kids get nervous about it, but kids handle it well. Kids take the initiative and some kids take the textbook home and will do extra problems in areas they need help. [TINT14]

Students could not only see their scores but also analyze them, so they would know which skills they needed to improve. Students could then hone in on those standards or skills with which they are struggling and take concrete steps to practice toward mastering them.

A handful of teachers actually met with each student individually to discuss their scores, identify their strengths and weaknesses, set goals, and strategize about how to meet those goals. One example comes from a middle school in Harlem, where an eighth-grade communication arts teacher, who has been teaching for 30-plus years, holds individual conferences with each one of his eighth-grade students to discuss scores and identify areas or skills needing improvement. According to the teacher, he uses the student's individual Grow Reports® as a starting place for these discussions and as "a motivator." They look at the individual Grow Reports® together to identify where the student scored high and low. The student then writes down on a note card at least three skill items or standards from the Report that she wants to work. The students are instructed to keep the note card. Whenever they have free time during class, the students can work on the skills listed on their note cards in CTB workbooks, which are stored in the back of his classroom and broken down into chapters aligned to what is on the test. Students can then practice answering test-like questions that specifically address the skills with which they may be struggling.

The teacher shared that this is not only a way for him to incorporate test preparation into his curriculum without letting it overwhelm the curriculum but also a means for addressing individual students' needs. Another teacher in the Bronx printed out activities and resources aligned to each of standards from the Grow web site and prepared workbooks, so that students could work on areas, on which they wanted to improve.

Administrators' Reported Uses of the Grow Reports®

New York City district- and building-level administrators reported using the Grow Reports® to gain a greater understanding of the educational and instructional concerns particular to their level of the education system. One superintendent explained how they use Grow Reports® on both the district and school level. All schools and classrooms receive the Grow Reports®. Administrators felt that "data" – including the data that the Grow Network reported on – would drive their decision-making toward more informed educational choices for students, teachers, and themselves as administrators. They used this data in many ways ranging from articulating administrative needs to putting forth an educational vision for their school or the larger school district, often paying attention to the school's diverse student needs.

Looking across the interviews and surveys, we found that how administrators used the Grow Reports® could be grouped into four main categories: (1) identifying areas of need and targeting resources, (2) planning, (3) supporting conversations, and (4) shaping professional development activities for teachers. However, depending upon the administrators' position (e.g. superintendent for curriculum, district math coordinator, school principal, or staff developer), they synthesized the data organized by the Grow Reports® and implemented their decisions into the school or district in slightly different ways.

Identifying Areas of Need and Targeting Resources

Administrators explained that the Grow Reports® helped them to identify class-, grade-, and school-wide strengths and weaknesses that could then be used to make decisions about planning, shaping professional development activities, and determining student performance and demographics. According to the survey, a large majority of administrators (86%) reported that they "sometimes," "often," or "always" used the Grow Reports® to identify high performing students.

Table 7: Administrator Use of Grow to identify students I use the Grow Reports™ to... (N= 137) (includes “sometimes,” “often,” and “always” responses)

Response	%
Identify low performing students for additional resources (n=121)	93
Identify grade-wide strengths and weaknesses (n=122)	93
Identify class-wide strengths and weaknesses (n=122)	92
Identify school-wide strengths and weaknesses (n=122)	92
Identify high performing students for additional resources (n=121)	86

An even larger majority (93%) said they used the reports to identify low performing students for additional resources. The administrators spoke quite frequently about the “bubble kids.” They felt pressure from their district leaders to move students on the cusp. For example, a director of math and assessment in a Queens district described her yearly data training for principals: “We had some really powerful conversations – I said it is important to know who the children on the cusp are and what it’ll take to move them” [DADINT13]. A Deputy Superintendent described a district policy that had been in place for one year as “moving test scores” [DADINT6]. The district identifies students who are near the proficiency level (near the top of Level Two but just under Level Three) and requires principals to target this small group of students, placing them with the best teachers and extra supports.

Other administrators reported that their district was targeting test prep towards high Level Twos and low Level Threes. One school was using Grow to help them target these students for academic intervention services. As a principal in Canarsie commented, “all some students need, by the grace of God, is to pick up one or two more questions” [SADINT19]. Other schools would ask teachers to identify their “bubble kids,” so that the school could provide a “push in” teacher to cover a specific skill or strategy. A staff developer in Canarsie spoke about pulling students near the cusp from across the fourth-grade classes to prepare them for the test:

I pulled out the 10 kids from each class that narrowly missed level 3. I took some 610, 620, I even took some 590 [referring to scaled scores] if I thought that the kid had a solid test [but] just missed it by a question. And I divided them into groups [...] and what I do is take similar questions from the GROW report. [SADINT20]

One principal in Harlem described how her school uses data:

We have AIS [Academic Intervention Services] teachers – reading and math – our Extended Day program, our Title One funding...so these are the kids we are going to start identifying who are in need of remedial work, not that we are going to forget about the others. Nonetheless, we are going to look at these [kids] because they can sway this way or that way, so I say to the teachers, ‘...Who do we feel could be pushed on to Level 3? Who is in jeopardy, at risk of either staying at the status quo or regressing? We don’t want that, so...what we do

is we divvy up these kids and the AIS teacher then focuses her instructional program on those students. [SADINT7]

Administrators also felt that Grow was helpful at the teacher-level in addition to the student-level. As one superintendent explained, “Grow allows you to combine test results and longitudinal analysis to diagnose a school’s strengths. This helps make decisions about professional development and resource allocation” [DADINT17]. Another superintendent believed that the data could “help new principals to identify good teachers” and “meet one of his challenges by supporting the school leaders” [DADINT17].

Table 8: Administrator Use of Grow to allocate resources I use the Grow Reports™ to... (N= 137) (includes “sometimes,” “often,” and “always” responses)

Response	%
Select curricular and/or test prep materials (n=124)	79
Allocate budgetary resources (n=120)	54
Build class rosters (n=123)	49

However, a little more than half of administrators who responded to the survey (54%) reported that they used the Grow Reports® to “sometimes,” “often,” or “always” allocate budgetary resources. This data suggests that though administrators commonly talk about using data to determine how to allocate resources, they are less likely to put these ideas into practice.

Planning

Once administrators identified which students, teachers, and resources they wanted to target, this information helped them to focus their school or district planning activities. Administrators explained that they used the data on the Grow Reports® to plan for setting school and district priorities and for instructional programs. A district-level math and science coordinator in Manhattan explained during an interview:

We get general test data but they don’t tell us much. Grow can tell us where the weaknesses are... I have schools that are at different levels, I have teachers at different levels. Grow helps map out what District facilitators need to target. [DADINT4]

One assistant principal in the Bronx explained, “Grow helps her target resources, such as after-school and extended day programming, to students in need” [SADINT11].

Table 9: Administrator Use of Grow for planning I use the Grow Reports™ to... (N= 137) (includes “sometimes,” “often,” and “always” responses)

Response	%
Plan test preparation activities (n=122)	84
Set school-wide priorities (n=120)	83
Make professional development decisions (n=119)	82
Conference individually or in small groups with teachers (n=121)	80
Select curricular and/or test prep materials (n=124)	79
Reflect upon the faculty’s strengths and weaknesses (n=123)	74
Learn more about the NYC standards and tested topics (n=119)	72
Make year-long pacing calendars (n=122)	66
Conduct classroom observations (n=122)	33

According to the survey, the administrators reported using the Grow Reports® for planning activities in a variety of ways. Eighty-three percent of the administrators stated that they used the Grow Reports® “sometimes,” “often,” or “always” to “set school-wide priorities.” Eighty-four percent said that they “sometimes,” “often,” or “always” used the Grow Reports® to plan test preparation activities. Meanwhile, two-thirds (66%) said that they “sometimes,” “often,” or “always” used them to make yearlong pacing calendars.

While most administrators made use of the Grow Reports, they did so with lesser frequency than teachers. According to the survey, 4% of administrators used the Grow Reports® once a week or more, compared to nearly 7% of teachers; and 12% of administrators used it once a month, compared to 19% of teachers. However, 43% of the administrators used the Grow Reports® three to six times a year, compared to 32% of teachers. Though administrators, like teachers, noted that Grow provides useful information, they do not look to this data exclusively to make decisions because it is based on a single assessment. As one superintendent said, “You have to take into consideration how valid or current the data is when you are using the previous year’s data. This is why you have to compare Grow with all of the other data sources” [DADINT9].

Table 10: Administrator response to survey question: How frequently did you use the Grow Reports™ last year?

Response	%
Never	7%
Once or twice a year	17%
3-6 times a year	43%
Once a month	12%
2-3 times a month	17%
Once a week or more	4%

School-based administrators also believed that Grow should be used along with other forms of data to assist with planning activities, as an assistant principal in Harlem noted:

Grow is not a sole diagnostic. I see it as one piece. We use it in conferences and goal setting with students, and together with portfolios to create a portrait. It's constant because a student can grow in six weeks. It's wonderful. A fourth grader can recover and jump ahead. A combination of assessments shows change over time. Grow is static. [SADINT8]

A principal in Canarsie echoed this statement, saying, “You must look at the whole picture and never use just one source” [SADINT16]. In other words, using the Grow Reports® for planning purposes was just one in addition to many other helpful resources that for administrators used throughout the school year.

Supporting Conversations

In interviews, many of the administrators spoke about how the Grow Reports® helped framing conversations they had with teachers, parents, or other administrators related to student learning, professional development for teachers, or addressing school or district challenges. According to the survey, 87% of administrators reported using the Grow Reports® to conference with teachers; 73% with other administrators or staff developers; 70% with parents; and 55% with students. In addition, 46% of the administrators also reported using the Grow Reports® to “often” or “always” conference individually or in small groups with teachers.

Table 11: I use Grow Reports™ as a basis for conferencing with... (N=137)

Response	%
Teachers	87
Other administrators and/or staff developers (“coaches”)	73
Parents	70
Students	55

During the interviews, school-based administrators explained that they used the Grow Reports® and related materials in grade-wide meetings “to inform mini-lessons,” “to connect reading and writing activities,” and “as a way to start a dialogue with teachers to identify which categories the students couldn’t answer and develop strategies to respond.”

One principal described how she wanted her teachers at an upcoming grade-wide literacy meeting to “review their Grow Reports® and to identify a grade-wide challenge and then to decide on a few teaching strategies.” She explained that she felt it was “important for colleagues to realize that they all have the same type of kids. Teachers tend to always think that ‘their kids’ are different from other teachers’ students” [SADINT16]. She hoped that this conversation would contribute to teachers sharing strategies and approaches, and that together they would “think through what we need to do” for the school. “Mandating does not work,” she said. She hoped that the teachers would make the decisions.

Some administrators also felt that the Grow Reports® were a useful resource for parents. One assistant principal in the Bronx explained that in her school testing and test scores have been very public and the Grow Reports® are openly shared with students and their parents. “Parents are relieved by it,” she said.

They feel it gives them good feedback and they can see resources, which is important for them to feel like there’s something they can do to help their child. If a child is in need, now parents can see it. It breaks it down... what you can do, what you can make better, and the steps to take to get there. [SADINT11]

District administrators also embraced using the Grow data with parents and spoke about using the Grow Reports® to prepare for parent conferences. One superintendent explained that, “if parents could use the [Grow] data this would be a major push” [DADINT16].

Shaping Professional Development Activities

The instructional resources incorporated into the Grow Reports® are approved by the New York City Department of Education and aligned to state standards. Since the Grow Reports® were aligned with standards and objectives that the school and district viewed as important, administrators found that the reports were often a good fit for shaping professional development activities.

Many district and school-based administrators provided professional development workshops for principals and teachers on how to use Grow; however, some administrators took using Grow to another level. As one superintendent explained, “The ultimate goal is for teachers to have flexible grouping in their classroom. Grow provides the easy part of this. Then the hard part of this is up to us: how to use it!” [DADINT10]. Thus, Grow became used both as the focus of a professional development activity (typically in a workshop format) and for shaping other professional development activities, such as helping teachers to create differentiated instructional activities or learning about school or district-wide standards and goals through their close alignment to the Grow Reports®.

Eighty-two percent of the administrators surveyed said they either “sometimes,” “often,” or “always” used the Grow Reports® to make professional development decisions. Administrators interviewed said that the Grow Reports® were helped them to check whether their own assessments matched with what the Grow Reports® outlined. A staff developer in the Bronx uses the report to help guide professional development at multiple levels – for herself as an administrator, for the teacher, and for developing differentiated instruction and assessment activities for the student.

I use it just as kind of a back up.... I think to find out myself. Each kid’s an individual so I actually like to work with the kid and show the teachers how to work with the children and find out what he needs. And then to double confirm what I think we need to work on, we’ll look at the Grow Report... That’ll kind of be like telling me that I’m on the right track. If I see that there’s something different... we’ll go back to the child and we’ll see where the inconsistency is. [SADINT10]

CONCLUSION

Data and Accountability – Negotiating the High-Stakes Climate

As listeners and readers, students will collect data, facts, and ideas; discover relationship, concepts, and generalizations; and use knowledge generated from oral, written, and electronically produced texts.

—Excerpt from New York State’s Standard 1 for English Language Arts

Broadly stated, standards define the qualities of a sound education, identifying fundamental learning goals that we want all students to master while providing a framework for the design of curricula and a guide for teaching practice. Accountability tries to ensure that students meet these standards and holds teachers and administrators responsible. Though accountability in education can take many forms from active parent-teacher associations to school visits and observations, in the era of No Child Left Behind, state-mandated tests have become the principle lens through which educators relate to standards and accountability.

Not all tests are created equally. Some states have tests closely aligned to standards, while others do not. Even when the test is closely aligned to the standards, as is the case in New York State, reducing broad standards into a test inevitably requires not only distilling fundamental skills into sets of discrete questions but also selecting certain standards over others since no single test can address all standards. The challenge for most practitioners is that satisfying learning standards, as defined in curriculum frameworks, and doing well on state-mandated standardized tests often raise unrelated and sometimes contradictory demands. Given the increasingly high stakes attached to accountability, the standards measured on the test are, therefore, more crucial to the success of students, teachers, and schools than ones not accounted for by the test. As a result, educators must balance the need to address the standards with the need to prepare for the test.

The teachers and administrators who participated in this research study are acutely aware that standardized testing is the reality with which they must contend. In New York third- and fifth-graders are held back if they fail to meet proficiency on the City’s standardized tests. This trend – to end student promotion to the next grade level based on a single mandated assessment – has gained momentum in states across the country. Louisiana, Florida, North Carolina, Mississippi, and Texas are all retaining students based on third- or fourth-grade performance (Ed Week, 2004).

Such policy decisions deeply affect educators. Teachers in the national and Grow samples acknowledge that test preparation has become a routine part of what they do. And while they also recognize the limitations of the test as an accurate measure of student performance, teachers also know their job is to help their students succeed according to the established criteria and high-stakes tests are the criteria by which students, teachers, administrators, and schools are judged.

The findings from this study suggest that negotiating such tensions is a routine part of teachers’ work. While teachers generally find the tests to be reliable indicators of students’ performance and are comfortable using testing information to identify students in need of additional support,

they also routinely question whether the tests adequately measure all that students need to know. As a result, teachers frequently seek to monitor student learning and triangulate assessment data in a variety of different ways.

From the interviews, survey and scenarios, we learned that teachers implement both formal and informal assessment strategies, which they perceive as more useful than information yielded by external exams. The assessment strategies they find most valuable are linked closely to classroom activities and are developed and implemented by the teachers themselves. They favor classroom assessments, whether formal or informal, which reach beyond discrete skills and allow students to demonstrate their abilities and show their thought processes in the context of richer instructional tasks. As a whole, the findings suggest that teachers feel that externally mandated assessments are useful in the decisions they make in the classroom only when part of a collection of assessment strategies that enable them to view students' work through a variety of different lenses.

Interestingly, administrators' attitudes about high-stakes testing are not markedly different from teachers. They generally accept the overall validity of the test as a measure of student performance (more so than teachers). They are comfortable using test data to identify students who need remedial help. And while they believe that a single test can offer a fairly accurate, though not necessarily complete, picture of what students know and what they need to master, they also recognize the importance of using multiple assessment strategies—both viewpoints also supported by teachers.

What is different is how administrators experience accountability pressures and talk about the ways in which the high-stakes climate of NCLB has changed their thinking about the role of standardized test data. There's no question that administrators are under intense pressure. Principals can be removed from schools that are not making adequate progress as measured by testing results; superintendents and their deputies can be replaced; and schools and districts can lose resources when students chose to transfer from failing schools to better performing ones. While administrators may sound like teachers when they talk about the importance of using multiple assessments, the reality is that they are much more likely to think about how to set instructional priorities and direct resources based on the results of the high-stakes exams. Testing carries a different kind of weight for administrators than it does for teachers. Administrators, therefore, actively encourage teachers to engage in explicit and routine test-preparation activities and to teach test-taking skills. They also acknowledge that because everyone is being judged by the same information – superintendents, principals, teachers, and students – they are more likely to align what they do to what the data suggests they should be doing and they are more likely to encourage teachers to do the same.

Data-Driven Decision-Making and the Grow Reports

As outlined earlier in the theoretical framework, data-driven decision-making requires, first and foremost, that appropriate data be turned into useful information that can aid in the making of knowledgeable and informed decisions. Digital technology has played a major part in making it possible for educators to interact with appropriate data that can be used to make decisions on a more informed basis. Specifically, the relative ease of use and sophistication of data gathering, storing and delivering systems has made data accessible in a meaningful format to whole sets of constituents whose access to data in the past was either nonexistent or presented in dense and unintelligible reports. This research focused primarily on how this increased access to data might inform and support decisions about children's education. We were interested in how the Grow Reports® would be used by educators to support decision-making at various levels of the system.

As a tool, the Grow Report® tries to underscore the ways in which test data can be used to inform instruction, not just accountability. By design, it provides a format that builds a bridge between standards, testing results, and instructional strategies. It provides educators with guidance for how to cope with students' differential learning patterns and provides teachers with a rationale for differentiating instruction. Our task was to examine how did this concerted effort to introduce teachers to data in the context of instructional decision-making affect teachers?

The survey comparison between Grow-using teachers and the national high-stakes sample sheds some interesting light on this question. Across both samples, very few teachers, less than a third, believe that the state-mandated standardized achievement test is a comprehensive measure of a student's ability, but more Grow-using teachers than the national sample considered it an accurate measure. More Grow teachers believed that the state-mandated test is aligned to what they teach in their classrooms. These same teachers were somewhat less likely than the national sample to believe that test scores influence their teaching on a daily basis. Despite the high-stakes climate in which they teach, Grow-using teachers felt less forced to exclude anything not on the test from their teaching, and they estimated that they spent less time on explicit test preparation activities than was true for the national sample.

The largest differences between the Grow-using teachers and the national sample were in how appropriate they thought it was to use standardized test scores to group students by ability, and in how likely they were to prepare their students by providing them with items from old tests. More than 80% of the teachers we surveyed thought it made sense to use test data to group students by ability. Less than 20% of the national sample thought that ability grouping by test scores was appropriate. New York City teachers were also far more likely to prepare their students by providing them with test practice in the form of old tests, yet they did not feel that such activities significantly compromised or curtailed their use of other instructional strategies. Based on these findings we can speculate that what the Grow Report® does in bridging the gap between test results, standards, and instructional practices is provide teachers with a framework that helps them to understand the relationship among, and distinguish between, these elements of the educational system.

Compared to the national sample, these New York City teachers can be considered “data-friendly.” In other words, the skepticism many expressed about the value of standardized test data in educational decision-making was often more about its incompleteness and the stakes involved, than a rejection of the whole idea of using standardized tests to measure individual achievement. The vast majority seemed essentially satisfied with the alignment between the state standards, their own curriculum, and the test. Teachers and administrators had a wide range of opinions about the ethical, political and educational implications of high-stakes testing, but they were using the data to inform decision-making appropriate to their role in the educational system.

Teachers use the testing data provided in the Grow Reports® to plan activities, lessons and units. They sometimes use it as a starting point for conversations with students, parents, specialists, and administrators. Some teachers plan their own professional development, based on test data, to focus on areas where they think their students need more help. But more than anything else, teachers use test data to allocate their own resources: time, attention, practice, and homework.

In contrast to teachers, administrators use the Grow Reports® to help them make building-level decisions, such as allocating resources, targeting those students who are most in need of additional support or who could benefit from more advanced instruction, and starting conversations that help them communicate and establish leadership priorities. In this way they help to support and focus the priorities of staff developers and instructional leaders. Reports are also used to identify strengths and weakness in classrooms and across school buildings so that professional development can be focused on those areas of instruction most in need of improvement.

While administrators, like teachers, stress the importance of using multiple sources of data when making decisions they, too, know that raising test scores is the ultimate measure of their success. As a result, administrators in New York encourage teachers to focus on the students who are on the cusp of meeting proficiency on the high-stakes test. In particular, the students who are at the top of level 2 in New York’s four-level system (where 1 = failing, 2 = below standard, 3 = at standard and 4 = above standard) and who can be moved up to level 3 are the most crucial for accountability. Getting the failing students to where they are doing better—though still below standard—or getting students who are doing well enough to do really well does not count for as much as moving as many of the kids who are at the top of level 2 to the bottom of level 3. For teachers and administrators, this means identifying those “bubble” students, making sure they spend enough time practicing the skills they need for the test, pitching instruction to the areas they are having difficulty with, or creating ability groups designed to help these students move up a notch.

As a tool that aligns test results with standards and instructional strategies, the Grow Report® appears to be highly successful in creating a navigational framework for educators. For teachers, the Grow Report® presents data in a meaningful format from which teachers can draw the information they need to support differentiating instruction and thinking about students’ weaknesses as well as their strengths. In this sense, the Grow Report® can be said to be playing a critical role in

helping teachers navigate the tensions that exist in a high-stakes climate between the accountability model of schooling, where data from standardized tests drives assessment and practically defines the standards, and a reform model, where diversity is considered in the curriculum and defined by differentiated pedagogical practices.

For administrators, the Reports represent a means to meet accountability requirements. This is not surprising given that they face real penalties. From an accountability standpoint, the “bubble kid” strategy makes sense from their point of view. The danger of this kind of strategy is that it tends to direct resources to one group of students at the expense of others (Confrey and Makar, 2002). Accountability requires that they get the averages up, which means focusing more attention on the test and the tested skills some students lack than on expanding and improving the skills of those who are failing or succeeding.

The limitations of the Grow Reports® are, of course, that they rely on a single standardized test and as a result, the way in which they present evidence of differential learning is limited. However, the teachers and administrators with whom we spoke were extremely cognizant of this limitation, and teachers, in particular, found creative ways to use the information contained in Grow to inform both instructional practices and administrative decision-making. When teachers in this study talked about data-driven decisions, it was almost invariably in the context of accountability – helping students meet standards as evidenced by their performance on standardized tests. Yet, we also heard teachers discussing data-driven decision-making in the context of a reform agenda in which diversity, both cultural and cognitive, is as central to learning as standards are to the accountability perspective.

Technology and Data-Driven Practices — What The Future Holds

What we have learned from this study is that teachers, in particular, combat a narrowing interpretation of their students’ strengths and weaknesses by using multiple sources of data – impressionistic, anecdotal, and experiential – accrued over the long term and based on many experiences with their students to make most decisions. Having been introduced to the possibilities of using systematic data to make instructional decisions, teachers are eager for more and better data. Digital technology has already played an important role in giving teachers access to this data. We believe that digital technologies – particularly the highly portable, wireless, and multi-functional devices that are becoming increasingly prevalent today – will expand on this role by helping teachers engage in the kinds of data-gathering and data-analysis processes that are inclusive of the multifaceted ways in which children show evidence of learning.

Digital technology makes possible the systematic collection of far more varied and richer data that is tied more closely to practice than the standardized tests we currently use to measure achievement, such as the abilities to display data for manipulation, to allow a variety of interpretive activities, to drill down into data to see detail, and to see larger patterns across data sets with visualization tools.

Appropriate uses of flexible, mobile, easy-to-use digital devices could allow teachers and administrators to keep track of performance data, observational data, informal conversations, portfolios of student work, self-assessment and reflection, the stuff of daily instructional decision-making in the classroom. The digitization of data sources makes it possible to reflect systematically on the data, to organize it, share it and analyze it in a variety of ways. The inclusion of organized diagnostic and authentic, performance-based data in the decision-making process would go a long way toward affirming and supporting the practices that educators routinely engage in. As the educators in this study acknowledge, high-stakes tests communicate only a piece of what they need to know about the complex repertoire of skills and talents that children need to succeed.

The Grow Report® is an important first step in this process. It shows that the promise of making the transition to data-driven decision-making in our schools can be a useful and interesting process. Significantly, by creating a lens through which the relationship between standards, assessments, and instruction can be explored, it helps educators to find reason in and navigate the tensions that prevail in the high-stakes contexts in which they work.

REFERENCES

- Ackoff, R. L. (1989). From data to wisdom. *Journal of Applied Systems Analysis*, 16, 3-9.
- Barr, J.M. (2004). *Teacher location choice and distribution of quality*. Newark, NJ: Rutgers University, Department of Economics.
- Black, P. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-44.
- Breiter, A. (2003). *Information – knowledge - sense-making: A theoretical analysis from management/business literature*. Unpublished manuscript.
- Brophy, J. (2001). *Teaching*. Brussels: International Academy of Education.
- Cizek, G. J. (2001). Conjectures on the rise and fall of standards setting: An introduction to context and practice. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 3-18). Mahwah, NJ: Lawrence Erlbaum & Associates.
- Confrey, J., & Makar, K. (2002). Developing secondary teachers' statistical inquiry through immersion in high-stakes accountability data. In D. Mewborn, P. Sztajn, & D. White (Eds.), *Proceedings of the twenty-fourth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education PME-NA24* vol. 3, (pp. 1267-1279). Columbus, OH: Eric Clearinghouse for Science, Mathematics, and Environmental Education.
- Council for Basic Education. (2000). *Closing the gap*. Washington, DC: Author.
- Cromey, A. D. (2001). Data retreats: A conduit for change in schools. *Newsletter of the Comprehensive Center-Region VI*, 6, 21-23.
- Drucker, P. F. (1989). *The new realities: In government and politics/in economics and business/in society and world view*. New York: Harper & Row.
- Education Week on the Web. (2004). Quality counts 2004: *Count me in: Special education in an era of standards*. Bethesda, MD: Author. Retrieved August 16, 2004, from: <http://www.edweek.org/qc/reports/standacct-t1k.cfm>.
- Elliot, E. (1993). National testing and assessment strategies: Equity implications of leading proposals for national examinations. In M. T. Nettles (Ed.), *Equity and excellence in educational testing and assessment*, (pp. 415-424). Boston: Kluwer Academic Publishers.
- Elmore, R., & Abelman, C. (1999). *When accountability knocks, will anyone answer?* Philadelphia: University of Pennsylvania, Center for Policy Research in Education.
- Fullan, M. (2000). The three stories of education reform. *Phi Delta Kappan*, 81(8), 581-584.
- Fullan M. (2001). *The new meaning of educational change* (3rd ed.). New York: Teachers College Press.

Fuqua, D. R., Newman, J. L., & Dickman, M. (1999). Barriers to effective assessment in organizational consultation. *Consulting Psychology Journal: Practice and Research*, 51(1), 14-23.

Grow Network. (2004). *About the Grow Network*. Retrieved April 1, 2004, from: <http://gownetwork.com/>.

Heistad, D., & Spicuzza, R. (2003, April). *Beyond zip code analyses: What good measurement has to offer and how it can enhance the instructional delivery to all students*. Paper presented at the annual meeting of the American Educational Research Association (AERA) conference, Chicago, IL.

Herman, J., & Golan, S. (1990). *Effects of standardized testing on teachers and learning*. Los Angeles: UCLA Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Herman, J., & Gribbons, B. (2001). *Lessons learned in using data to support school inquiry and continuous improvement: Final report to the Stuart Foundation* (CSE Technical Report 535). Los Angeles: UCLA Center for the Study of Evaluation.

Kearns, D.T., & Harvey, J. (2000). *A legacy of learning*. Washington, DC: Brookings Institution Press.

Lemann, N. (1999). *The big test*. New York: Farrar, Straus and Giroux.

Linn, R.L. (1998). *Standards-based accountability – Ten suggestions* (Policy Brief 1). Los Angeles: UCLA, Center for Research on Evaluation, Standards, and Student Testing.

Linn, R.L. (2001a). *The design and evaluation of educational assessment and accountability systems* (CSE technical report 540). Los Angeles: CRESST.

Linn, R.L. (2001b). *Reporting school quality in standards-based accounting systems (policy brief 3)*. Los Angeles: UCLA, CRESST.

Linn, R.L. (2003a). Accountability: Responsibility and reasonable expectations. *Educational Researcher*, 32(7), 3-13.

Linn, R.L. (2003b). *Requirements for measuring adequate yearly progress (policy brief 6)*. Los Angeles: UCLA, CRESST.

Mandinach, E.B., & Cline, H.F. (1994). *Classroom dynamics: Implementing a technology-based learning environment*. Hillsdale, NJ: Lawrence Erlbaum & Associates.

McMillan, J.H. (2003). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice. *Educational Measurement: Issues and Practice*, 22(4), 34-43.

Meyers, E., & O'Connell, F.O. (2000, May 31). The test doesn't tell all. *Education Week*, 19(37), 34, 37.

- Mitchell, D., & Lee, J. (1998, September). *Quality school portfolio: Reporting on school goals and student achievement*. Paper presented at the CRESST Conference, Los Angeles, CA.
- Mitchell, D., Lee, J., & Herman, J. (2000). *Computer software systems and using data to support school reform*. Paper prepared for the Wingspread Meeting, Technology's Role in Urban School Reform: Achieving Equity and Quality, sponsored by the Joyce and Johnston Foundations. New York: EDC/Center for Children and Technology
- National Center on Educational Statistics. (2002). *The condition of education 2002*. Washington, DC: Author.
- National Commission on Excellence in Education. (1983). *A nation at risk*. Washington, DC: author.
- Newman, F., King, B.M., & Rigdon, M. (1997). Accountability and school performance: Implications from restructuring schools. In *Cool thinking on hot topics: A research guide for educators*. Cambridge, MA: Harvard Educational Review, 67(1), 41-74.
- No Child Left Behind Act of 2001 (NCLB). (2001). Washington, DC: author. Retrieved January 30, 2002, from: www.nclb.gov.
- Pedulla, J.J., Abrams, L.M., Madaus, G.F., Russell, M.K., Ramos, M.A., & Miao, J. (2003). *Perceived effects of state-mandated testing programs on teaching and learning: Findings from a national survey of teachers*. Boston: The National Board on Educational Testing and Public Policy, Boston College.
- Pellegrino, J.W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.
- Robelen, E.W. (2003, September 3). State reports on progress vary widely [Electronic Version]. *Education Week*. Retrieved October 22, 2004, from: <http://www.edweek.org/ew/articles/2003/09/03/01ayp.h23.html>
- Sacks, P. (1999). *Standardized minds: The high price of America's testing culture and what we can do to change it*. Cambridge, MA: Perseus Books.
- Salpeter, J. (2004). Data: Mining with a mission. *Technology and Learning*, 24(8), 30-37.
- Schafer, W.D., & Lissitz, R.W. (1987). Measurement training for school personnel: Recommendations and reality. *Journal of Teacher Education*, 38(3), 57-63.
- Secada, W. (2001). From the director. *Newsletter of the Comprehensive Center-Region VI*, 6(1), 1-2.
- Senge, P., Cambron-McCabe, N., Lucas, T., Smith, B., & Kleiner, A. (2000). *Schools that learn*. New York: Doubleday.

- Sharkey, N., & Murnane, R. (2003, April). *Helping K-12 educators learn from student assessment data*. Paper presented at the annual meeting of AERA Conference, Chicago, IL.
- Shavelson, R. J., & Towne, L. (2002). *Scientific research in education*. Washington, DC: National Academy Press.
- Shepard, L. (2000). *The role of classroom assessment in teaching and learning*. Los Angeles: UCLA, CRESST.
- Spielvogel, B., Brunner, C., Pasnik, S., Keane, J.T., Friedman, W., Jeffers, L., John, K., Herмос, J., et al. (2001). *IBM Reinventing Education Grant Partnership Initiative – individual site reports*. New York: EDC's Center for Children and Technology.
- Spielvogel, B., & Pasnik, S. (1999). *From the school room to the state house: Data warehouse solutions for informed decision-making in education*. New York: EDC's Center for Children and Technology.
- Stiggins, R. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(10), 758-765.
- Thorn, C. (2002). *Data use in the classroom: The challenges of implementing data-based decision-making at the school level*. Madison: University of Wisconsin, Wisconsin Center for Education Research.
- Viadero, D. (2000). High-stakes tests lead debate at researchers' gathering. *Education Week*, 19(34), 6.
- Wayman, J. C., Stringfield, S., & Yakimowski, M. (2004). *Software enabling school improvement through the analysis of student data* (Report No. 67). Baltimore, MD: Johns Hopkins University Center for Research on the Education of Students Placed At Risk. Available online at <http://www.csos.jhu.edu/crespar/techReports/Report67.pdf>
- Webb, N. (2002). *Assessment literacy in a standards-based urban education setting*. Paper presented at the AERA, New Orleans.
- Wise, S. L., Lukin, L. E., & Roos, L. L. (1991). Teacher beliefs about training in testing and measurement. *Journal of Teacher Education*, 42(1), 37-42.