DOCUMENT RESUME

ED 251 195 PS 014 729

AUTHOR Buchanan, Aaron
TITLE Avoiding Pitfalls in the Pursuit of Higher Standards for Elementary Schooling.
INSTITUTION Southwest Regional Laboratory for Educational Research and Development, Los Alamitos, Calif.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
REPO" 'O SWRL-TR-87
PUB D 15 Aug 84
CONTRACT NEC-00-3-0064
NOTE 52p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Academic Achievement; Academic Standards; Addition; *Arithmetic; Computation; Cross Sectional Studies; *Curriculum Development; Division; Educational Improvement; Elementary Education; Elementary School Mathematics; *Elementary School Students; *Mathematics Curriculum; Multiplication; School Schedules; Subtraction; *Time Factors (Learning); Vacations

ABSTRACT

Using four large existing databases on student performance in the elementary grades, this study involves a task-by-task, year-by-year cross-sectional analysis of six topics in computation. Information about what students eventually learn and when they learn it is provided by the data analysis; information about when skills are introduced and expanded is determined by examining school textbooks. This information is seen as being useful in setting higher standards for elementary schools, which currently concentrate on teaching skills piecemeal over several grades. Specifically, topics investigated are addition and subtraction with regrouping, multiplication and division facts with numbers 5 through 9, and multiplication and division beyond basic number facts. Results show that the topography of student performance varies across topics and across grades within topics. Implications of findings are that efforts focusing on learning opportunities in the middle elementary grades will be more productive than efforts undertaken at the point of entry into secondary schooling. In particular, it is concluded that students learn most when topics are taught as part of regular instruction, that the margin for school improvement may be more promising in certain segments of a subject (i.e., subtraction and division), and that a dramatic loss of performance occurs when the first teaching of a topic occurs late in the school year. (CB)

SWRL EDUCATIONAL RESEARCH AND DEVELOPMENT

Avoiding Pitfalls in the Pursuit of Higher Standards for Elementary Schooling

TR 87 August 15, 1984

SWRL EDUCATIONAL RESEARCH AND DEVELOPMENT

TECHNICAL REPORT 87

August 15, 1984

AVOIDING PITFALLS IN THE PURSUIT OF HIGHER STANDARDS FOR ELEMENTARY SCHOOLING

Aaron D. Buchanan

ABSTRACT

The pursuit of higher standards for elementary schooling needs to take full account of what schools spend their time on now, what students in the elementary grades eventually learn and when they learn it. Using large existing data bases on student performance in the elementary grades, this research involves a task-by-task, year-by-year cross-sectional analysis of six topics in computation. Results show that the topography of student performance varies across topics and across grades within topics. The implications are that focus and timing of learning opportunities for all or nearly all students in the middle elementary grades will be more productive than efforts focusing on the point of entry into secondary schooling.

AVOIDING PITFALLS IN THE PURSUIT OF HIGHER STANDARDS FOR ELEMENTARY SCHOOLING

Aaron D. Buchanan

Although the pursuit of higher standards is currently concentrating on secondary schooling, the pursuit will inevitably pick up elementary schooling before the quest for educational excellence is over. If the same get-tough legislation for secondary schools is extended to elementary schools, administrators and teachers will get new marching orders to do what they have always been trying to do: raise scores on tests of basic skills, the same unfriendly tests of achievement in math and reading that have always been a nemesis. This means that, within the elementary school ranks, teachers and administrators will face the same uncertainty they had before the nation began it's pursuit of excellence over what tasks students should be required to demonstrate by the end of each grade level.

What school officials know, but what legislators, government officials, and education scholars often lose sight of, is this: setting standards means one thing in thinking about high school graduates and whether or not they've had a sufficient number of the right kinds of courses to enter the university; it means something quite different in looking at a continuous web of six grade levels in the elementary school--one that has few natural seams for demarcating "real" standards at each grade level.

It is possible to set grade-by-grade standards in the elementary grades, of course, but the results are fairly arbitrary. Since instruction on almost any of the basic skills that are taught in elementary school typically stretches out over two or three grade levels, few of the tasks that often make up standards for the elementary school seem like important benchmarks in the growth, grade by grade, of what students are eventually supposed to know and know how to do.

Standard setting is especially difficult in the elementary grades because schooling at this level does not deal in tangible units, such as fixed, year-long courses, that are tied closely to topics that are covered in classroom instruction. Fixed courses are hard currency in the secondary school. There, standard setting and standard raising involve a fairly straightforward process for administrators and teachers of adding real courses, having definite beginning and ending points, onto existing graduation requirements. By contrast, elementary school administrators and teachers must deal with the slipperier proposition of raising their expectations against units of schooling that tend to float across a long expanse of time having no sharp demarcations in what is taught from grade to grade. Here, the goal is not so much to complete something as it is to add value to the aggregate performance of an undifferentiated mass of students entering kindergarten or moving through six succeeding grade levels.

Instead of completing courses, the current elementary school curriculum is concentrated on the development of skills--expansive and complicated ones that are taught piecemeal over several grade levels. Any piece of a developing skill that is taught within any single grade level seldom has a beginning or ending point that is very striking.

Students who successfully learn what is taught will have value added to what they were able to do before, but they won't have reached a point where they can do much of anything next year except to learn more pieces. What students have learned by the end of any single grade level may well be significant in relation to what they knew before, but the nature of what students have completed is not very significant.

At any point in time, the substance of what elementary students should know or know how to do is fragmented. As elementary schooling currently works, nothing is ever really taught once and for all at any elementary grade level, much less learned by an entire grade level's worth of students. Always, there are pieces to be added by teachers at another grade level. Since the end points of instruction at any single grade level seldom represent the completion of something that, by itself, is very useful, most grade-by-grade standards that are established for the development of a basic skill in the elementary school are indefinite. The development of a skill is never really complete, at least not with the same degree of finality that a course is completed, because there is always some task, usually a more complicated one, that could be added to students' repertoires. To make things even more indefinite, the pieces of a developing skill that elementary teachers teach at one grade level are usually retaught extensively at the next grade level and are often reviewed at still another grade level. This recycling has the effect of stretching real benchmarks (i.e., things that students need to be able to do with reliability and confidence) over several grade levels.

The conventions that are often used to set standards are not very compatible with the conventions of schooling in the elementary grades. Standard setting tends to concentrate on what individual students need to know or know how to do before they move to another level of instruction. On the other hand, the process of schooling in the elementary grades depends on timely acquisition of important skills by a generation of students. Most standard setters have their thinking grounded in the performance of academic tasks that individual students must demonstrate before they can enter successfully into secondary coursework. Too often, they have little understanding of how these performances actually "grow" within a generation of students across the elementary grades. As a consequence, most standards consist of tasks that individual students must perform either just before they leave elementary school or very near the point where these students enter secondary school. (Strange as it may seem, the standards for leaving elementary school are often somewhat lower than standards for regular entry into secondary school.) Higher standards are established by requiring individual students to perform more tasks or more complicated tasks near the point of entry into secondary school. Most often, these standards can only be met through massive remediation, also near the point of entry into secondary school.

When standards are set grade by grade for the elementary school, they often don't align grade by grade with the bulk of opportunities for teaching and learning that are provided in school textbooks. When misalignment occurs, it's usually because standards have been set too conservatively. The academic tasks that students must perform to meet standards at a particular grade level are often stressed one or two grade levels earlier in textbooks. The misalignment of standards with learning opportunities is often most serious when standards become "minimal competencies" that individual students must demonstrate before they can be promoted to regular, non-remedial instruction at a higher grade level. There is nothing unique about the academic tasks that constitute minimal competencies, but there is usually something wrong with the points in students' grade-by-grade progress where they are required. Because the academic tasks involved in minimal competencies often act as gates to instruction at a higher grade level, the inability of students to perform them successfully carries the stigma of failure, not only for individual students, but, to some extent, for teachers. Students are often required to perform the academic tasks involved in minimal competencies well past the grade level where most learning opportunities might be expected to occur. As a consequence, minimal competencies have come to represent a "last" chance for individual students to perform routine academic tasks rather than a "best" chance to learn them.

CHARTING GROWTH IN THE PERFORMANCE OF SCHOOLING

What is missing in the current call for higher standards is a sound awareness of students' cumulative accomplishments on academic tasks as they move through the elementary grades. Without a clear picture of what effects the schooling process is producing now, it is difficult to know where the margins for improving this process are located. Many standards are set well past the point where "working harder" on the tasks would be most beneficial. Other standards are set in place at one grade level even though many prerequisites for these tasks are not now being learned by many students. Without these prerequisites in place at earlier grade levels, the progression of schooling cannot work as intended.

The intelligent pursuit of excellence should mean more for the elementary school than establishing standards that amount to little more than elaborate but regressive mechanisms for triggering remediation. What this pursuit should lead to is adjustments made early enough in the schooling sequence so that risks to the regular progress of a generation of students on routine academic tasks can be reduced. Often, these adjustments begin with full exploitation of opportunities that exist in that part of school instruction that is, in practice, sustained by school textbooks. Full exploitation does not imply that individual teachers must attend slavishly to a sequence of textbook pages. What it means is that administrators and teachers should take full advantage of what they can know about critical points in the mainstream of grade-by-grade instruction for a generation of students where major opportunities for teaching and learning are most likely to occur.

Full exploitation of learning opportunities means knowing as much as possible about the gaps that exist in the grade-by-grade sequence where regular learning opportunities are scheduled to occur in school textbooks and the points where a critical mass of students, say 75% to 85%, are able to demonstrate the accomplishments that these learning opportunities are designed to produce. Sometimes, nothing needs to be adjusted because the gap between opportunity and accomplishment has little impact on other things that students must learn how to do for the process of schooling to move along smoothly through several grade levels. At other times, these gaps are important, because new learning opportunities assume that certain routine tasks are already in place. These gaps trigger formal remediation, which moves large numbers of students onto separate tracks where catching up becomes unlikely.

Making adjustments in the growth performance on routine academic tasks requires an understanding about two things:

(1) How an academic skill is instructed task by task and grade by grade--when the most rudimentary tasks are introduced; when the range of simpler tasks is expanded to include more complicated ones; when the capability to do certain routine tasks becomes a "taken-for-granted" part of work on some other task.

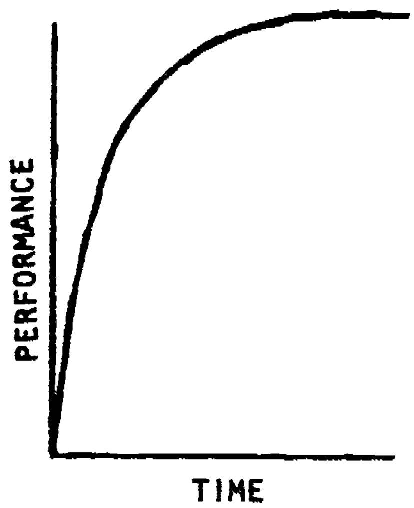
(2) What a large cross section of students knows how to do grade-by-grade; what proportion of the aggregate of all students at particular grade levels can perform a specific academic task; how this performance changes as the momentum of teaching moves from a less inclusive task to a more inclusive and more complicated one.

Looking at the details of task-by-task performance by large groups of students is not something the education community has done very much, especially from a clear perspective of when these tasks ordinarily receive concentrated instruction in regular elementary school programs. When it comes to higher standards, the concern has focused on "new" ideas, techniques, and materials to remedy deficiencies in student performance at the point where these deficiencies become obvious and troublesome.

Part of the trouble has been a lack of interest by most researchers in the anatomy of school textbooks, the primary conduit for most learning opportunities that focus on routine academic tasks. Another thing that has hampered task-by-task analysis of student performance has been overreliance by researchers on conventional standardized achievement tests as their data source. Unfortunately, the information that is most accessible from these tests comes in the form of percentiles and grade-equivalent scores which are mainly intended to show how the performance of individual students compare with the average performance of all students nationwide who are at the same grade level. Percentiles and grade equivalent scores reveal nothing about the aggregate performance of a generation of students on particular academic tasks. For example, it is possible for individual students (or all students in a classroom, or all students at one grade level) to have high grade-equivalent scores on a range of tasks when measured against the population of all students in the nation at that grade level. At the same time, the proportion of this population of students that can perform each of the separate academic tasks may be quite low. In dealing with percentiles or grade-equivalent scores, the meanings of phrases such as "pursuit of excellence" and "raising standards" get lost in a shifting frame of reference--what is good or not good is relative to how well everybody else performs.

Fortunately, it is possible to get a better understanding of student performance on routine academic tasks using other kinds of information from conventional, standardized achievement tests and from other data sources that are currently available. With this information, we can do a reasonably good job of generating some task-by-task profiles of performance by an aggregate of students at different elementary grade levels. In general, the cumulative instructional accomplishments of schooling can be sketched out by looking first at the proportion of students, all at the same grade level, who can perform some routine academic task. Second, we can see how this performance changes as the emphasis in instruction moves, grade by grade, to include the task's more complicated "relatives." In a very practical sense, this kind of analysis gives us some feeling for how well a generation of students tends to keep up with instruction--not, as with norm-referenced information from conventional achievement tests, how well individual students keep up with everybody else.

The growth of student performance on a single task might be expected to look like the simple learning curve shown below. There is a huge



increase in the proportion of students who can do a particular task early in instruction. But as instruction continues, the increases in the proportion of students who can perform the task become smaller and smaller.

In practice, the topography of the performance of a generation of students isn't this clean, especially when teaching is stretched out over several related tasks at different grade levels. A lot depends on the point in time when one looks at performance of a cross section of students during the school year. For example, we should expect some decrease over the summer in the proportion of students who can do a particular task. For almost all students there is a gap of two or three months between the end of one grade level, when a task is introduced, and the beginning of the following grade level, before it is reviewed. Second, we would also expect some dips in performance as instruction shifts from simpler tasks to more inclusive ones. It's also possible we might see some eventual decline in performance because of forgetting, but it's unlikely because routine academic tasks are ones that get a lot of use in the learning of other things.

To the extent that elementary schooling is successful, we would expect to see the curve on a series of closely related tasks to eventually level off at a point where more than 80% or even 90% of all students at a grade level are successful. On other series of tasks, where schooling is less successful, we would expect to see the performance curves flatten out at a much lower level.

Are there practical limits to how well a generation of students can perform routine academic tasks? The obvious answer is yes. If one looks at performance of a group of students on some routine academic task, no more than 100% of the students can come up with the right answer. But seeing 100% who can do anything is rare, especially for a large group of students. Performance levels in the range of 90-95% are likely to be about as good as schooling ever gets.

More important than the ultimate limits of the performance curves are the methods at our disposal to see how these curves actually behave now. For example, researchers now have more than ten years worth of data accumulated from the National Assessment of Educational Progress (NAEP) showing, in great detail, what students at ages 9, 13, and 17 in a national sample are able to do. We can also determine, with reasonable accuracy, the approximate time when different school topics are likely to be taught, just by looking at when these topics are covered in most school textbooks.

To find out more about what students know and when they know it, we took a careful look at changes over time in the level of student performance on several routine tasks that are part of six very fundamental topics in elementary school mathematics. These topics include:

- Addition with regrouping (carrying)
- Subtraction with regrouping (borrowing)
- "Hard" multiplication facts
- "Hard" division facts
- Multiplication beyond basic facts
- Division beyond basic facts

The intent was to find points in time where the basic curve showing the level of student performance begins to flatten out after a substantial amount of formal instruction has taken place. We also wanted to see what performance curves look like for two series of tasks that are part of two different but related topics in mathematics, and we wanted to find out how soon after the end of formal instruction that increases in levels of performance begin to flatten out.

Four existing data bases were used to examine student performance on the six computation topics. Two are large SWRL data bases that contain information on elementary students at each of grades 1-6. A third is the entire set of results for students at ages 9 and 13 who were part of the second mathematics assessment of the National Assessment of Educational Progress (NAEP). The fourth data base consists of statistics on individual items used in establishing norms for students at different grade levels on Form U of the Comprehensive Tests of Basic Skills (CTBS). Use of all four data bases allowed us to look at the growth of student performance on routine computation tasks that are basically taught, but not necessarily learned, by the end of grade 4 and continue to show improvement at later grade levels. Altogether, the four data bases included information that was generated during a period from 1978 through 1982.

Most of the growth in student performance was determined by looking at performances on the same or similar items given to students at different grade levels in school. The population represented by each data base was a little different, so the information from each data base was kept separate in the analysis to minimize the likelihood of concealing differences between apples and oranges.

The basic procedure was to retrieve selected items from different data bases for the same grade level and from different grade levels within the same data base. The objective was to follow across several grade levels the growth of student performance on routine tasks that are part of the six topics on computation with whole numbers. It was impossible to follow only a single task, such as 2-digit addition, over several grade levels, because our data didn't include addition with only

2-digit numbers past about grade 3. It was possible to pick items that represented gradual increases in the range of tasks that students were asked to do. For example, we could shift our attention from 2-digit addition to 3-digit addition, as the analysis moved from grade 3 to grade 4 within our data bases.

Results of the analysis across the six computation topics are reported in three sections: Addition and Subtraction; "Hard" Multiplication and Division Facts; and Multiplication and Division Algorithms. Discussion on the first topic, addition and subtraction (with carrying and borrowing), is fairly extensive. The purpose is to demonstrate the analytic procedure for finding places in the grade-by-grade sequence of instruction where the performance curves flatten out. Results that deal with the remaining topics are then summarized with briefer discussions.

Addition and Subtraction (with Carrying and Borrowing)

Using the PVS, Essential Skills, and NAEP data bases, we were able to construct a profile of growth on routine tasks involved in addition and subtraction with carrying or borrowing. We begin with 2-digit numbers, about the middle of grade 2, and end with 4- or 5-digit numbers in grades 5 and 6 (see Figure 1.1). From the PVS data base, we picked up performances on addition and subtraction at mid-year in grade 2, which is about the earliest point that textbooks begin to provide instruction. By the middle of grade 2, a little less than half of the students in the PVS data base could do a routine task involved in addition with carrying:

$$\begin{array}{r} 29 \\ +\underline{16} \end{array}$$

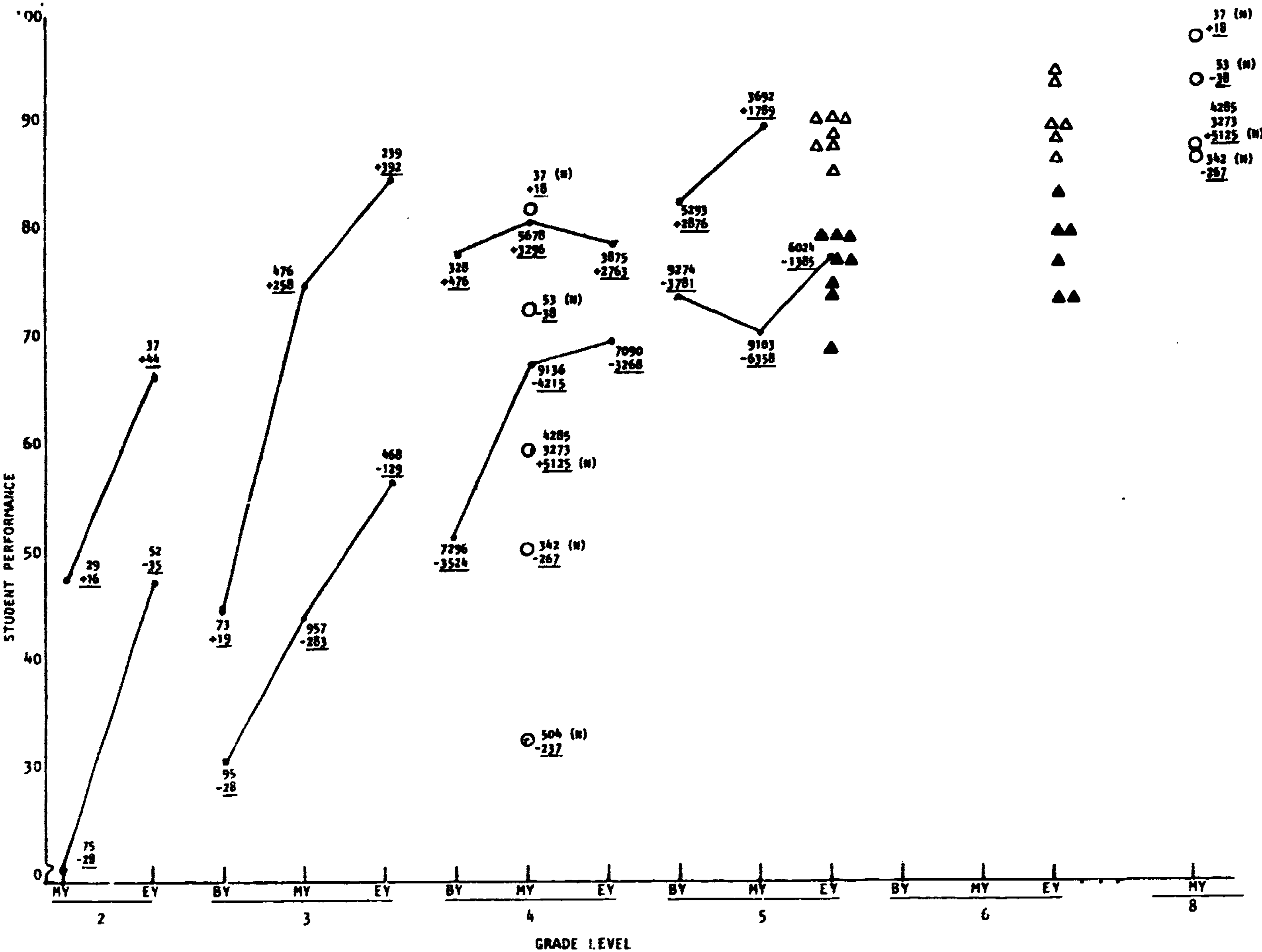


FIGURE 1.1 PERFORMANCE LEVELS ON TYPICAL ITEMS FOR ADDITION AND SUBTRACTION OF WHOLE NUMBERS (with regrouping)

*PVS Items (1978-79)
BY = Beginning Year
MY = Mid Year
EY = End of Year

Performance Levels on Comparable Essential Skills Items
△ Addition
▲ Subtraction

○ Coordinate NAEP (1978) Items (N)

Looking at the top curve in Figure 1.1, it is clear that by the end of grade 2, almost 70% of the students could now do a similar task involved in addition with carrying:

$$\begin{array}{r} 37 \\ +\underline{44} \end{array}$$

At the beginning of grade 3, student performance drops back to about the point where it had been at the middle of grade 2. By the end of grade 3, however, the performance level on a more complicated task involving addition of 3-digit numbers with two carries, climbed to almost 90%.

$$\begin{array}{r} 239 \\ +\underline{392} \end{array}$$

Our analysis shifted from 2-digit to 3-digit numbers between the beginning and end of grade 3, because there were no suitable tasks involving addition of 2-digit numbers in the PVS data base at the end of grade 3. However, the shift from two-digit to 3-digit numbers is consistent with the shift that textbooks make when they stress addition with 3-digit numbers in grade 3. Addition with 2-digit numbers is covered in textbooks for grade 3, but it is only an introduction to more complicated forms of addition involving numbers with three, and sometimes four, digits. There is some drop in performance on addition with 3-digit numbers during the summer between grades 3 and 4, but not nearly as much as there was with 2-digit numbers between grades 2 and 3.

Beyond the end of grade 3, our analysis doesn't show any other dramatic increases in the level of student performance as addition with carrying shifts to larger numbers. Items in the PVS data base shift to 4-digit numbers at the end of grade 4, consistent with a shift in what is stressed in textbooks. In the PVS data base, the level of student performance on 4-digit numbers hovers near the 80% level through the middle and the end of grade 4 and then climbs to almost 90% by the middle of grade 5. It's fairly safe to assume that, if we could continue to look at addition tasks involving only 2-digit and 3-digit numbers through grades 4 and 5, student performance might inch upward beyond the 90% level we saw with 3-digit addition at the end of grade 3, but it's unlikely that they would move very far. Grade 3 is the last time that students get much intensive instruction on addition with carrying. By grade 4, textbooks shift their emphasis in computation to multiplication and division, so they may have only two or three lessons that deal directly with addition, and these usually cover subtraction at the same time.

To extend our analysis of addition skills beyond grade 5, we looked at performances on comparable addition tasks in the Essential Skills data base. We do not show the actual items in Figure 1.1, but we do show levels of student performance, using hollow triangles, on 4-digit addition at the end of grade 5 and on 5-digit addition at the end of grade 6. These performance levels appear to be about the same at the end of grades 5 and 6 as levels in the PVS data base at the middle of grade 5.

Student performance levels in the NAEP data base are generally lower than those in our other data bases, especially at grade 4. NAEP's second assessment of 9-year-olds doesn't include many items on addition with carrying, and none of the ones that are in the data base involve addition of two 4-digit numbers, the kind of routine task we are looking at in the PVS data base at grade 4. One NAEP item does involve addition of <u>three</u> 4-digit numbers. On this task, the performance level for 9-year-old fourth graders is about 55%, which is almost 25 percentage points lower than performance levels in the PVS data base on addition with <u>two</u> (rather than three) 4-digit numbers. The NAEP data base does have an item involving 2-digit addition with carrying. A little more than 80% of the 9-year-olds at grade 4 answered it correctly. This NAEP level of performance on 2-digit addition at grade 4 is a little lower than the levels in PVS for 3-digit addition at the end of grade 3.

Why are performance levels in NAEP generally lower than performance levels in PVS and Essential Skills data bases? It's hard to say. A partial answer may reside in the way students are required to do tasks in NAEP assessments. NAEP uses free-response items, for which students write rather than select the correct answer, for all of its tasks involving basic computation with whole numbers. The other two data bases involve multiple-choice items. It is possible that selecting a response on a multiple-choice item is easier for the kinds of routine computation tasks we are observing. On the other hand, the wrong answers in a multiple-choice item may actually increase the likelihood of common errors in computation. The wrong answers in multiple-choice items often represent a very obvious and easy-but-wrong way to "attack" the problem. When students consider the wrong-answer choices, it's very likely that many of them see a quick fix for getting through with a computation task, even though their answer is wrong. Wrong answer choices for computation tasks usually involve simple intermediate steps that are not appropriate, but they are easy for students to recognize. Because students can easily see how a wrong answer choice could be derived from numbers that are given, the answer becomes plausible.

There is another reason that NAEP performance levels may tend to be lower than performance levels in the PVS and Essential Skills data bases. NAEP items are administered under conditions that are very carefully controlled to such an extent that testing sessions may, in fact, be rigid and uncomfortable for many students. All NAEP items are administered by specially-trained personnel, not by a student's regular classroom teacher. Some directions are read aloud by the person administering the assessment; others involve special, tape-recorded directions. For some tasks students are shown the directions; for others they aren't. Individual students respond to a wide variety of item types, many involving tasks that are much more difficult than ones the students are regularly asked to do. This combination of unfamiliar tasks and testing environment inherently introduce more complications than students ordinarily encounter in testing that is done in the classroom.

CTBS performance levels on addition tasks are comparable to results in the PVS and Essential Skills data bases. Figure 1.2 shows performance levels on five addition tasks that cover several different grade levels.

The CTBS-derived performance level on 3-digit addition with carrying begins at 78% at the end of grade 3. This result is a little lower than the 85% level observed in PVS, but it moves on to almost 90% by the end of grade 4. The CTBS performance level on addition of three numbers having up to four digits begins at about 72% at the end of grade 4. This is considerably higher than the 60% level observed in NAEP (see Figure 1.1) on a comparable task at the middle of grade 4, and the performance level on this task continues to move upward, reaching about 80% by the end of grade 6. CTBS includes tasks that involve addition of three numbers having up to six digits in its test for students who are at the end of grade 6, 7, or 8. The level of performance on this task begins at almost 80% at the end of grade 6, and moves up to 83% at the end of grade 8. The 83% level is about the same as what we see in the NAEP data base (see Figure 1.1) on addition of three numbers at the middle of grade 8. Thus, it seems that schooling can produce levels of performance on addition of whole numbers that eventually reach 80-90% for a population of students.

The growth of student performance in subtraction runs almost parallel to performance on addition tasks, but at a somewhat lower level. The data in Figure 1.1 shows that, for numbers of about the same size, subtraction with borrowing runs anywhere from 10 to 30 percentage points lower than addition with carrying. The biggest difference occurs at the end of grade 3, where about 85% of the students in the PVS data base could do a task involving 3-digit addition that required two "carrying" steps, but only 55% could do a comparable task involving 3-digit subtraction that required only one "borrowing" step. In part, the

239
+392

468
-129

difference in difficulty has to do with the fact that subtraction with borrowing involves a sequence of steps that is different from addition with carrying. In many ways, the subtraction task may be more complicated, even though it requires only one borrowing step. For one thing, the student must decide whether or not borrowing is necessary before beginning to subtract one digit from another. Within the addition task, the carrying step is taken care of after the initial addition of digits has already begun. But there is another factor that is undoubtedly involved. Teachers at grade three will provide instruction on addition of 3-digit numbers with carrying to almost all students by the end of grade 3, but many will not get to subtraction with borrowing, especially with students who typically require a lot of time and patience, before they learn new things. Work on subtraction of 3-digit and 4-digit numbers comes later in grade three textbooks than work on comparable forms of addition, and many teachers don't get to it.

From grade 4 to grade 6, the level of performance on subtraction shows about the same rate of increase as performance on addition, although there is still a difference of about ten percentage points between levels of performance on addition and subtraction.

By the middle of grade 8, NAEP data show performances on addition and subtraction of 2-digit numbers that have come quite close together; both of them reach a level where about 90% of the grade 8 students are successful.

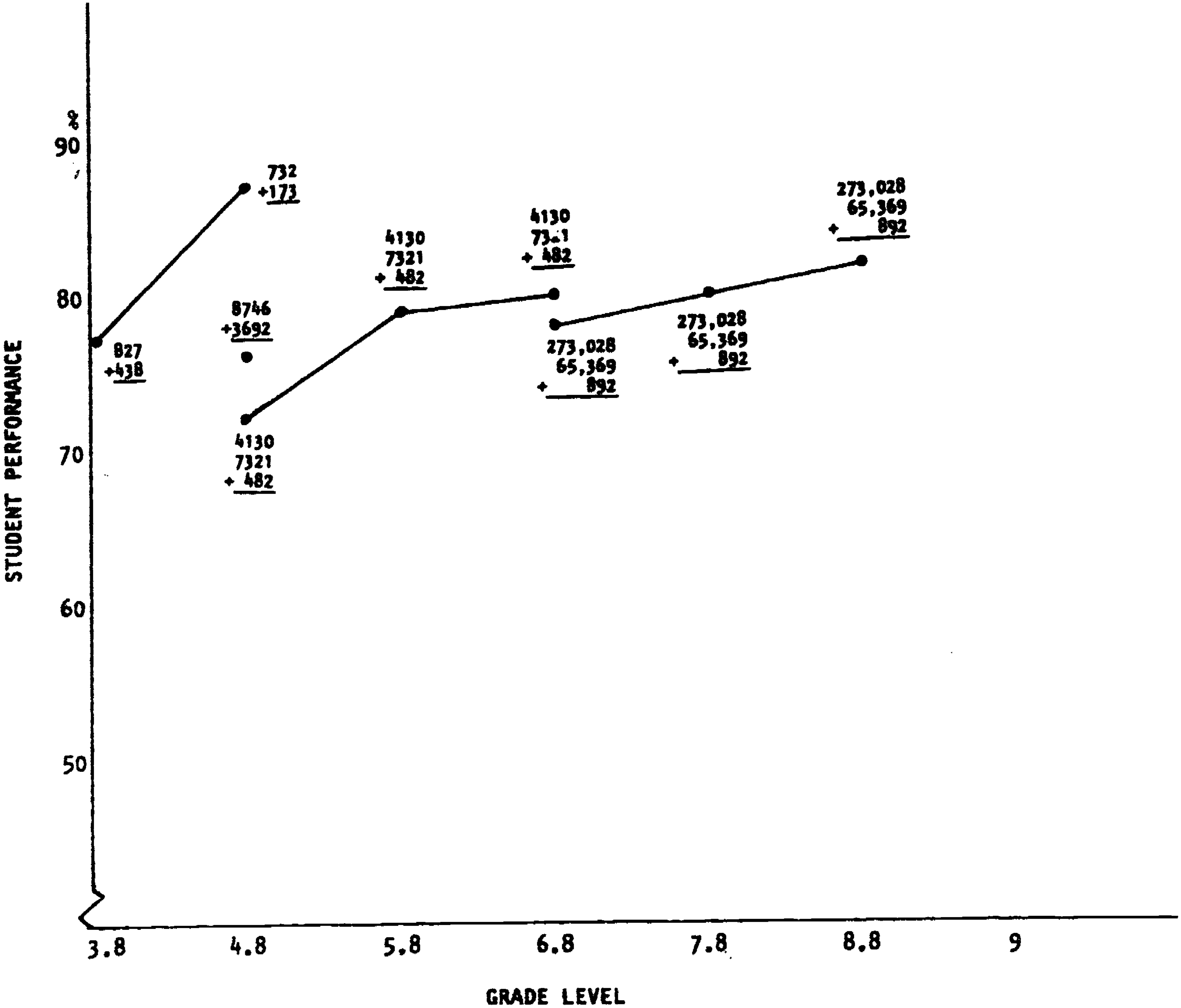


Figure 1.2 Performance Levels on Addition Problems Similar to Items on CTBS Form U

Performance levels derived from CTBS are considerably lower for subtraction with borrowing than they are for addition with carrying through all of the elementary grades. However, the two come very close to each other by the end of grade 8. As shown in Figure 1.3, a little more than 50% of the students at the end of grade 3 can solve a 2-digit subtraction problem that involves borrowing. By the end of grade 4, 70% can solve a 3-digit subtraction problem that has one borrowing step, while 55% to 65% could solve 4-digit subtraction problems that have one or more borrowing steps. The subtraction task showing the lowest level of student performance at grade 4 requires three "borrowing" steps. It is especially tricky, because many students will subtract 8 from 8, instead of 17, in the tens column. By the end of grade 5, the level of student performance on this problem moves from 54% to 67%, and it moves on up to 74% by the end of grade 6.

Overall, our data on addition and subtraction show the following patterns:

1. The growth of student performance on both addition and subtraction takes on about the same zig-zag profile, especially across grades 2, 3, and 4 where the bulk of regular instruction appears in school textbooks.

2. A lot of what students learn in grade 2 is forgotten by the beginning of grade 3, but is quickly relearned and extended.

3. Most of the "improvement" in addition and subtraction occurs in grade 3.

4. The level of performance on addition with carrying flattens out in grade 3, although the range of performance at the 80-90% level is extended in subsequent grades to include addition with larger numbers.

5. The level of performance on subtraction does not begin to flatten out until near the end of grade 4.

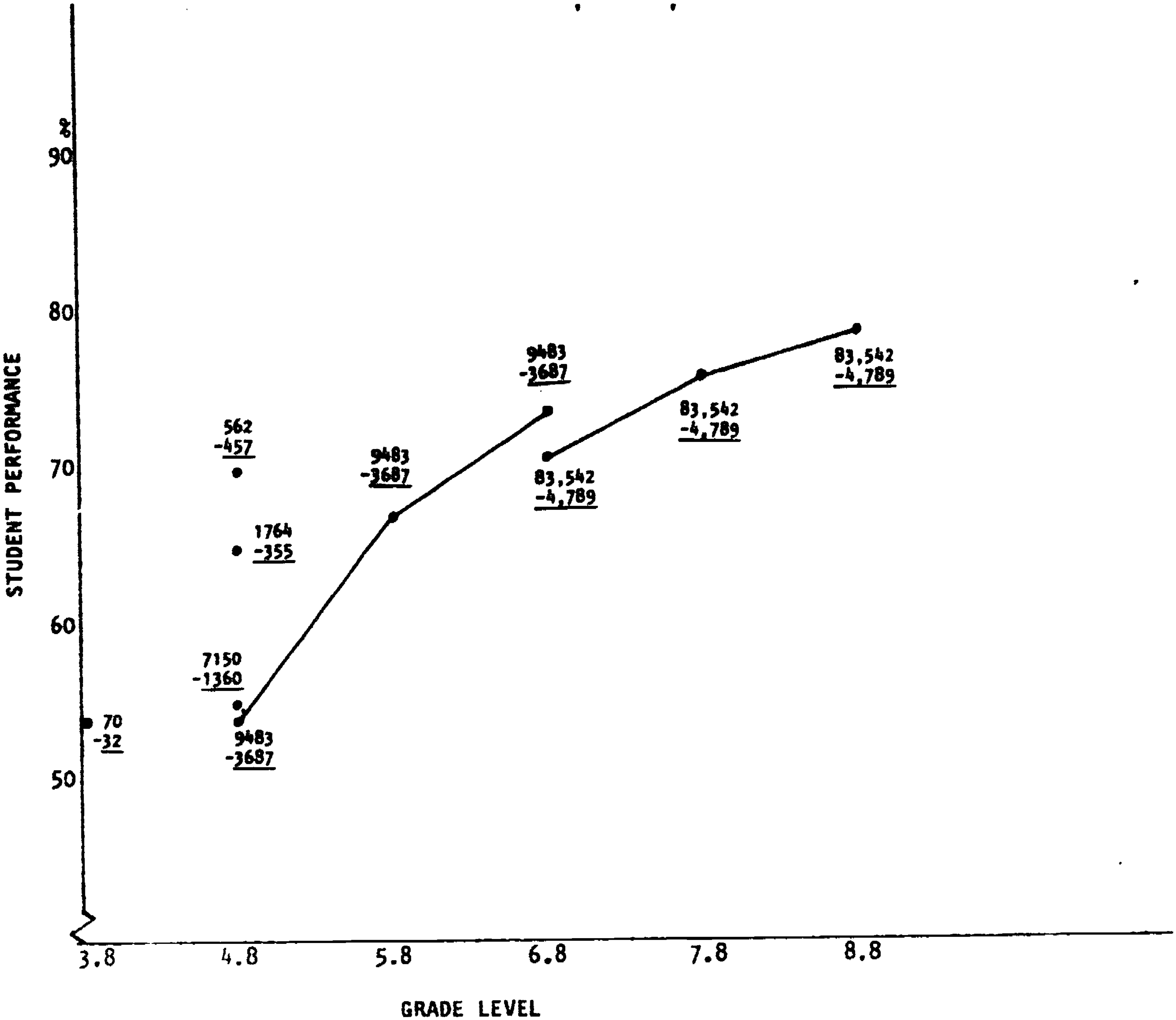


Figure 1.3 Performance Levels on Subtraction Problems Similar to Items on CTBS Form U

"Hard" Multiplication and Division Facts

Using the PVS data base, we were able to track the growth of performance on hard multiplication and division facts from grade 3 through grade 6. No single multiplication or division fact is included in every PVS instrument. But we were able to follow several closely-related tasks through the beginning of grade 6 that require recall of multiplication and division facts. Textbook lessons on hard multiplication and division facts don't begin until grade 3, and the bulk of instruction is completed by the middle of grade 4. However, multiplication and division facts get a lot of indirect practice when students learn and practice algorithms for multiplying and dividing larger numbers and also when they are learning how to generate equivalent forms for common fractions and mixed numbers.

The results in Figure 2 show that the level of performance on "hard" multiplication facts moves from a low at mid-year of grade 3, where a little more than 50% of the students in the PVS sample could find the answer to 6 multiplied by 9, to a high of better than 90% on about the same kind of multiplication problem in grade 5. The very high level of performance on 9 multiplied by 5 at mid-year of grade 5 may have something to do with the fact that multiplication involving "fives" has more regularity than multiplication involving "sixes," "sevens," "eights," or "nines" and may therefore be easier for students to remember.

On facts that involve about the same numbers, student performance on division starts out at about the same level as performance on multiplication. But, after the middle of grade 3, performance on division trails performance on multiplication by five to ten percentage points all the way through the beginning of grade 5.

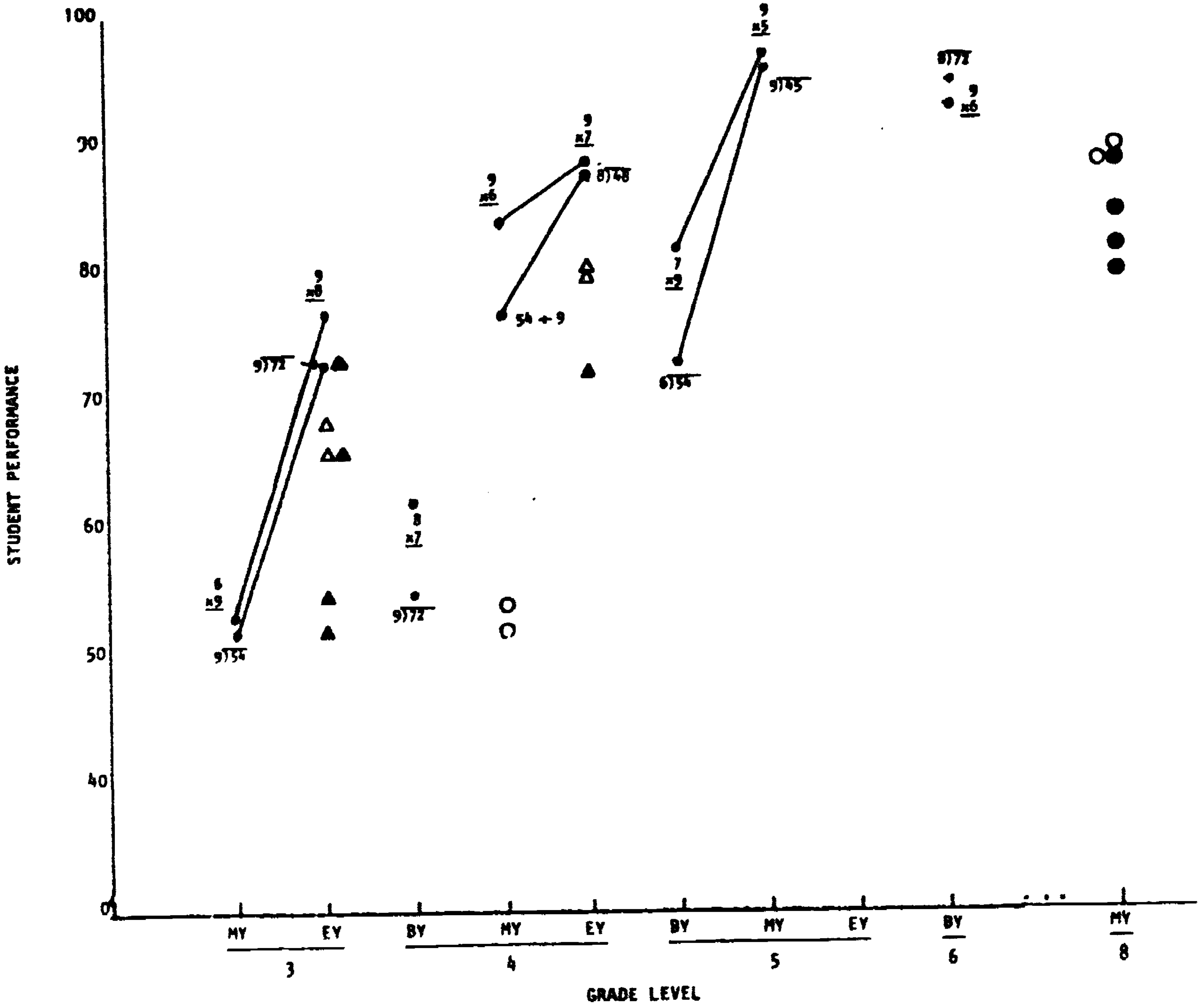


FIGURE 2 PERFORMANCE LEVELS ON TYPICAL ITEMS FOR HARD MULTIPLICATION AND DIVISION FACTS

*PVS Items (1978-79)
BY = Beginning Year
MY = Mid Year
EY = End of Year

Performance Levels on Comparable Essential Skills Items
△ Multiplication Facts
▲ Division Facts

Comparable NAEP (1978) Items
○ Multiplication Facts
● Division Facts

At grade 3, performance on division in the Essential Skills data base is very low. Two "hard" division facts have performance levels around 55% and the third, which involves division by 5, is at 70%, about the same as the "nines" division fact from the PVS data base. The low level of performance on hard division facts in the Essential Skills data base undoubtedly has something to do with the recommended schedule for teaching multiplication and division in the school district where these data were obtained. The district's program calls for multiplication and division facts to be taught in grade 3 and again in grade 4. The textbook series which this district has adopted for regular instruction includes hard division facts at grade 3, but, like most other textbook series, the lessons come very late in the grade 3 textbook. It is quite likely that a large number of teachers don't get to hard division facts before the end of grade 3, leaving this instruction for next year's teachers to pick up in grade 4.

Coverage of "hard" multiplication and division facts in the NAEP and CTBS data bases is spotty. CTBS Form U has no hard multiplication or division facts in any of its tests. NAEP covers hard multiplication facts in its assessment of 9-year-olds, but it doesn't cover division facts until the assessment of 13-year-olds.

Within the NAEP data base, the results on two hard multiplication facts at grade 4 are about 30 percentage points lower than results on similar facts from the PVS data base at about the same point in the school year. Part of this difference may be due to the fact that PVS tasks require students to select the best response from among four alternatives, while NAEP tasks require students to write their response in a space that's provided. Moreover, multiplication and division facts are presented in PVS in much the same way that students would encounter them in a textbook. Students taking NAEP assessments see no written problem. Rather, they respond to a phrase such as "five times nine," which is spoken by the NAEP proctor. This situation is quite unlike regular multiple-choice assessment items, which students see a lot. It's quite likely that the unusual conditions for testing associated with NAEP often lower student performance, especially at grades 3 and 4 when students are first learning how to multiply and divide. By the middle of grade 5, performance levels on multiplication and division facts are both at about 90% and they continue at about the same level into grade 6. Results from the NAEP data base show that the performance level on multiplication facts reaches about 90% at the middle of grade 8. Performance levels on hard division facts overlap with multiplication a little bit, but are slightly lower.

Overall, the results from the PVS, Essential Skills, and NAEP data bases show that the main growth in student performance levels on hard multiplication and division facts, about 40 percentage points, occurs between the middle of grade 3 and the end of grade 4. This is a time span when school textbooks provide most of their lessons that deal with hard multiplication and division facts. At the end of grade 3, performance levels in these three data bases vary widely, and they continue to do so through the middle of grade 4. During this period, it would be easy to make unreliable assumptions about what students can do.

Differences in problem format or in the kind of response that's required make a big difference in what individual students are able to demonstrate. There is a fair-sized drop in performance levels that takes place over the summer between grades 4 and 5, but, by the middle of grade 5, after multiplication and division facts are quite likely to be given a review by classroom teachers, performance levels flatten out at a place where 85-90% of the students are able to consistently recall multiplication and division facts.

Multiplication and Division Algorithms

Our analysis of algorithms for multiplication and division of larger numbers shows the same upward zig-zag pattern in performance that we saw with "hard" multiplication and division facts. Beginning midway through grade 3, PVS data in Figure 3.1 show that about 30% of the students can multiply a 2-digit number by a 1-digit number, and that a little less than 30% can do a comparable form of division. By the end of grade 3, almost 70% of the students can now do basic multiplication-with-carrying and more than 50% of them can do the same basic kind of division.

The second half of grade 3 is when many students get their first instruction on multiplication and division of 2-digit and 3-digit numbers by a 1-digit number, although some teachers may leave this introductory work to be done in grade 4. Often teachers choose instead to provide more practice on multiplication and division facts. All textbooks provide several lessons on multiplication of 2-digit and 3-digit numbers by a 1-digit number before the end of grade 3, and most of them also introduce a comparable kind of division involving 1-digit divisors. All

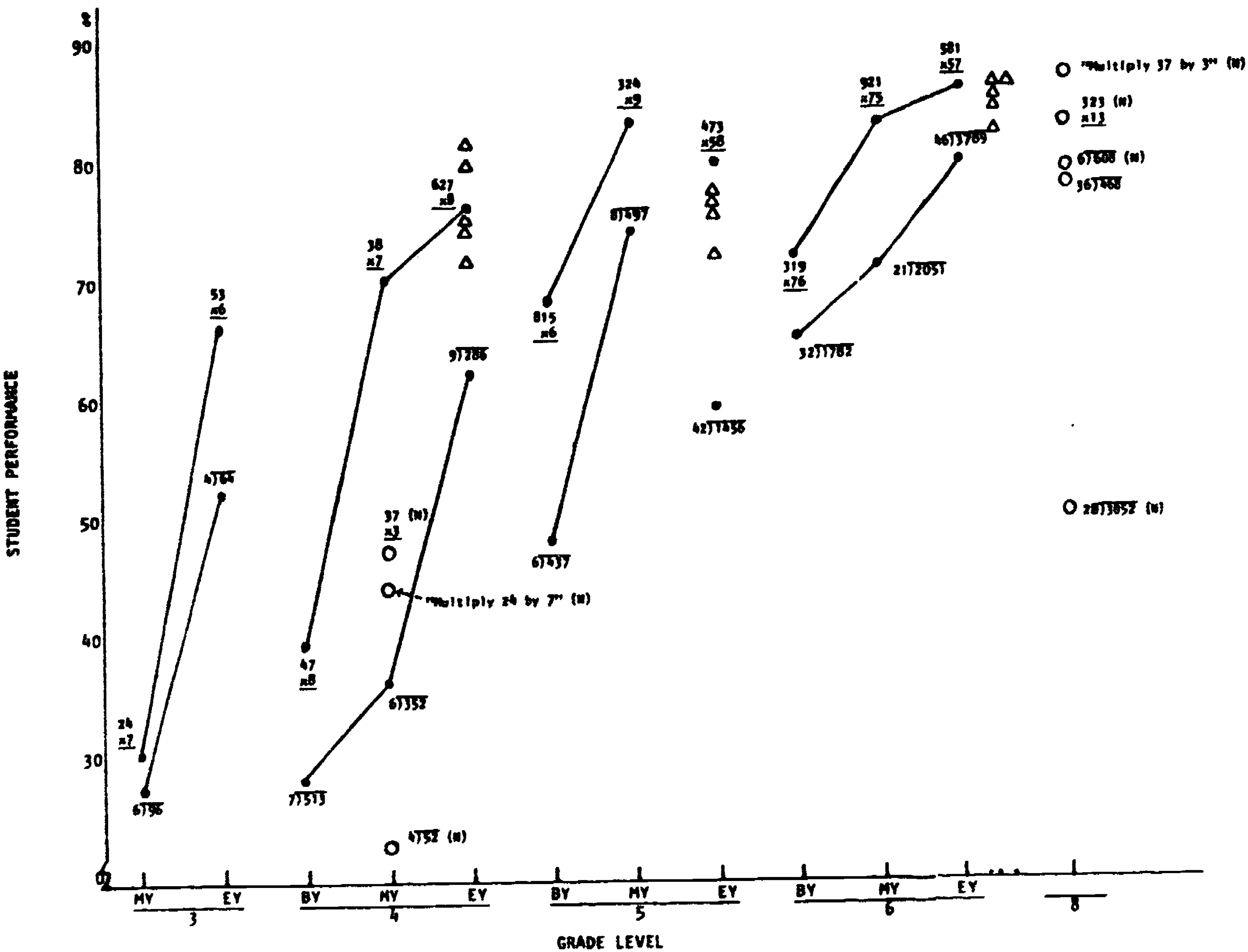


FIGURE 3.1 PERFORMANCE LEVELS ON TYPICAL ITEMS FOR MULTIPLICATION AND DIVISION BY 1- AND 2-DIGIT NUMBERS

*PVS Items (1978-79)
BY = Beginning Year
MY = Mid Year
EY = End of Year

△ Performance levels on Comparable Essential Skills Items: Multiplication

○ Coordinate NAEP (1978) Items (N)

of this work is repeated in grade 4, but at a much earlier point in grade 4 textbooks. Often work in the first half of grade 4 is extended to multiplication and division by tens (e.g., 36 x 20). Before the end of grade 4, most textbooks move to a more general (and more complicated) algorithm that involves multiplication and division by any 2-digit number.

Classroom review of the basic multiplication and division algorithms continues all the way through grade 6, partly because these algorithms are complicated and take a long time to learn. Another reason is that basic algorithms for multiplying and dividing whole numbers are also used in multiplication and division with decimals, which textbooks introduce in grade 5 and 6. Therefore, it is very useful to review multiplication and division with whole numbers just before new work involving multiplication and division with decimals is begun.

We followed the growth of multiplication and division algorithms by looking at performance levels to see how they were increasing in relation to a broadened range of performance that included more complicated multiplication and division tasks. It was impossible to follow student performance on the simple forms of multiplication or division by a 1-digit number all the way through to the end of grade 6. By the middle of grade 5, most textbooks have moved on to 2-digit multipliers and divisors. So have almost all of the items in the PVS data base. In order to have enough data points to show the "form" of increasing levels of student performance all the way through the end of grade 6, we shifted our analysis from 1-digit to 2-digit multipliers and divisors at the end of grade 5.

By the middle of grade 5, there is an increase of about 50 points in the percentage of students who can do problems in multiplication and division beyond basic computation facts. There is a slight, but expected, drop in level of student performance on multiplication tasks between the middle and the end of grade 5 as our items shift from 1-digit to 2-digit multipliers. By the end of grade 6, however, the performance level on 2-digit multipliers recovers to about the same level as performance with 1-digit multipliers one grade level earlier. Overall, the level of student performance on division trails performance on multiplication by anywhere from five to 40 percentage points. A huge difference occurs at mid-year in grade 4. Here, more than 70% of the students can multiply a 2-digit number by a 1-digit number, but less than 40% can do a similar problem where they have to divide a 3-digit number by a 1-digit number (and obtain a 2-digit quotient). In part this difference may have something to do with the tasks, themselves; dividing 352 by 6 (see Figure 3.1) may not be as comparable to multiplying 38 by 7 as, say, dividing 84 by 6. However, a large part of the difference in difficulty between comparably complicated forms of multiplication and division at this point of instruction must have a lot to do with the fact that textbook lessons on the division always come at a later point than lessons on multiplication. As a consequence, many teachers may be deferring instruction on division for many students until the next grade level. Many teachers feel that students who typically learn mathematics at a slower pace are not "ready" to do serious work on anything as complicated as long division at the end of grade 3. Some are even reluctant to ask students to do much with division at grade 4. It is almost certain that the difference we see in student performance levels at the end of grade 5 between multiplication and division by 2-digit numbers is largely a result of the difference in emphasis that teachers tend to give these two topics; multiplication by 2-digit numbers almost always gets covered by the end of grade 5, but the same kind of division often gets postponed until grade 6.

The flattening out in performance on multiplication that shows up in the PVS data base around grade 5 is verified by information on comparable items from the Essential Skills data base. At the end of grade 4, two problems in the Essential Skills data base that involve multiplication by a 1-digit number have higher performance levels than a similar task, 627 x 8, in the PVS data base, while three problems of the same type have slightly lower performance levels. At the end of grade 5, the performance levels on multiplication by 2-digit numbers in the Essential Skills data base are all slightly lower than performance levels on a similar PVS task, 473 x 58. By the end of grade 6, the two data bases agree almost completely.

The levels of performance on items in the NAEP data base are a lot lower in grade 4 than performances on comparable items in PVS or the Essential Skills data bases. By the middle of grade 8, multiplication and division by 1-digit numbers in the NAEP data base has moved up to about the same performance level we saw in the PVS and Essential Skills data bases at the middle of grade 5, although the division task derived from NAEP, 608 divided by 6, is much less complicated than the comparable PVS task, 497 divided by 8.

NAEP tasks that involve multiplication and division by 2-digit numbers are much less complicated than any comparable tasks in the PVS data base. One NAEP task, 323 x 13, requires no carrying steps in finding the partial products, while another task, 468 divided by 36, obviously has a "1" as the first digit in the divisor. A task derived from NAEP, 3052 divided by 28, is more complicated than any of the PVS tasks shown in Figure 3.1, because it has a "0" as the second digit in a 3-digit quotient. This task has an NAEP performance level of about 50% in contrast to the performance level on a PVS task, 3789 divided by 46, which is more than 75% at the end of grade 6.

CTBS-derived levels of student performance on multiplication (see Figure 3.2) show a steady growth on tasks involving multiplication by 1-digit numbers. This growth begins in CTBS at about 70% at the end of grade 4 and moves to almost 90% at the end of grade 7. (There are no items involving multiplication by 1-digit numbers in the CTBS tests in Form U that are intended for students at grade 3.) The level of performance on problems involving multiplication by 2-digit numbers takes a dramatic jump of more than 30 percentage points between the end of grade 5 and the end of grade 6, and it continues upward to about 85% at the end of grade 8. Performance levels in the middle grades are somewhat lower in the CTBS data base than either the PVS or Essential Skills data bases, but they eventually flatten out in the interval of about 80% to 90% in grades 7 and 8, much like performance levels in the other two data bases.

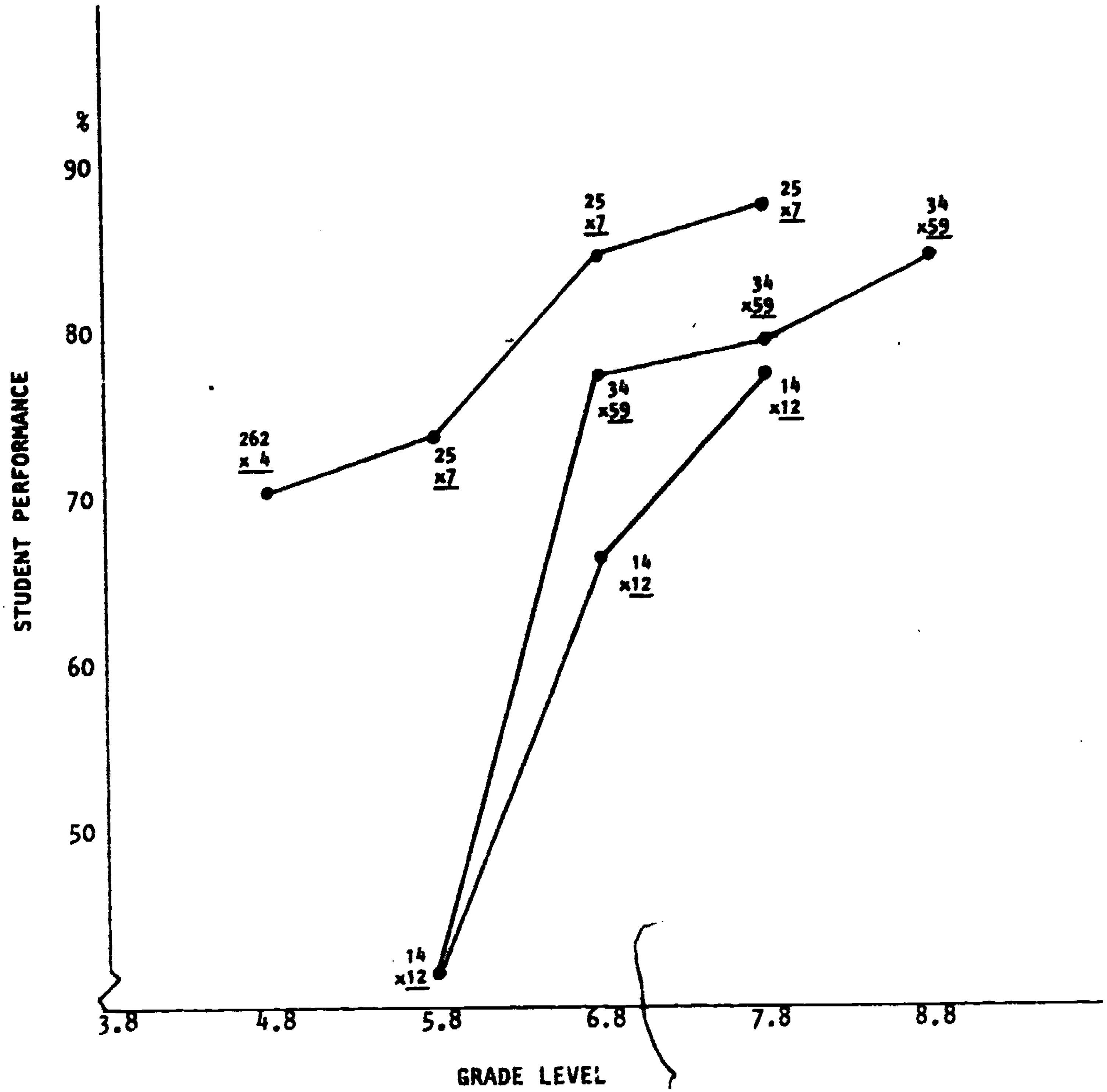


Figure 3.2 Performance Levels on Multiplication Problems Similar to Items on CTBS Form U

CTBS performance levels for division show some special challenges to schooling in the elementary grades. As Figure 3.3 shows, school instruction through grade 8 is only moderately effective for anything but the most straightforward division tasks. None of the division tasks that we were able to track move above the level of 70% to 80%. Performance levels on some special division tasks, having one or two zeros in the quotient (42084 divided by 21 and 6651 divided by 5), show almost no growth past grade 6.

IMPLICATIONS FOR HIGHER STANDARDS

The results of this analysis are important to the consideration of higher standards for elementary schooling. They are important to teachers, principals, subject matter specialists, and superintendents and their assistants who have the broad responsibilities to bring large groups of students along in the pursuit of excellence. Outside the domain of elementary schooling, these results are important for standard setters in government, colleges and universities, and even in secondary schools, who are inclined to support directives to "bring up test scores," often without benefit of knowing much about the details of what is generally being accomplished now in the elementary grades.

People who set higher standards for elementary schools cannot assume that, because students often perform higher on some routine tasks than they do on other areas of the curriculum, the teaching and learning of routine academic tasks is not a problem. The assumption is false. As these results demonstrate, many routine computation tasks come on line

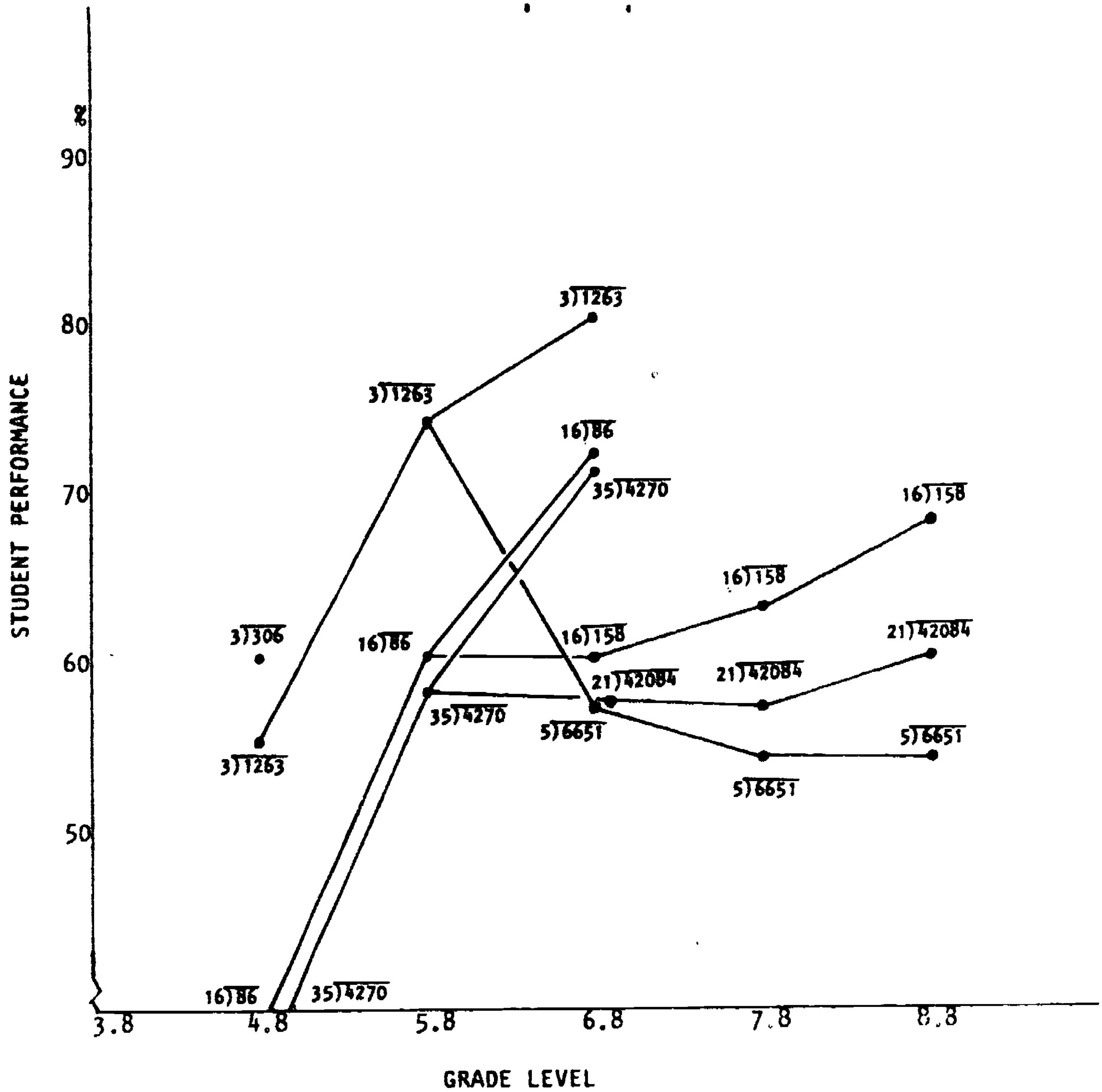


Figure 3.3 Performance Levels on Division Problems Similar to Items on CTBS Form U

for all or nearly all students well past the point in schooling where the core of instruction is provided, and well past the time when these tasks have become a prerequisite to student participation in other, often non-routine, areas of a subject matter.

As far as details are concerned, higher standards for the elementary grades should be concerned with several characteristics in the growth of student performance as a result of elementary schooling.

First, the growth in performance that occurs whenever topics are part of regular instruction for all or nearly all students is much more dramatic than any growth that is likely to occur later, because of intensive efforts to fine tune various aspects of ongoing instruction. Mastery learning and pass/fail programs implemented one or two grade levels after an academic task has been covered in regular textbook lessons may be important for some individual students or even some classrooms that include homogeneous groups of students who are mostly not doing very well in school. However, these programs will not do much to improve the performance of elementary schooling for either the present or future generation of students.

Second, the margins for school improvement are more promising in some segments of the topography of an academic subject than others. Our analysis shows that the margins in elementary school mathematics are most favorable around instruction on the subtraction and division algorithms. The levels of student performance on both algorithms trail their counterparts in addition and multiplication all the way through the development years of basic arithmetic with whole numbers. It would be easy to disregard the difference between subtraction and addition and between division and multiplication, attending only to the obvious fact that subtraction and division are more complicated processes than their computational siblings. Some of the discrepancy that we see between student performance on division and multiplication or subtraction and addition obviously do have roots in intrinsic differences in complexity in these basic arithmetic skills, but not all. We know that subtraction and, especially, division come very late in textbooks the first time that these topics are either introduced or significantly extended to larger numbers. We also know that teachers often tend to skip chapters near the end of their textbooks.

With subtraction and division, there are good-sized margins for school improvement especially if teachers are encouraged to complete all of the instruction on subtraction and division topics at each grade level where these topics occur--no matter how late the instruction comes in textbooks. Often, teachers don't go on to subtraction or division because they had quite a bit of trouble teaching addition or multiplication. Some students are still a little shaky. While sensitivity to the immaturity and anxiety of individual students is a characteristic of good teaching, the complete avoidance of difficult topics because all or nearly all students are not entirely ready represents questionable pedagogy. As long as teachers are careful, students can do significant and productive work on complicated topics even though they may not be able to do well on a complete range of tasks that these topics entail. For students who are likely to experience a lot of anxiety with complicated processes, there are ways, short of waiting until next year, to carefully control the difficulty of the tasks they are asked to attempt.

Third, a dramatic "loss of performance" occurs when the first teaching on a topic occurs late in the school year, and this loss may provide a fairly wide margin for school improvement for entire schools or districts. For each of the topics in this analysis, the instructional treatment on particular tasks comes very late in the textbook and, as we noted earlier may be skipped by many teachers. Even when instruction is provided, the amount of practice that students get on a particular task is often quite skimpy[1] and may come so late in the year that there is not enough time for periodic strengthening, spaced over several weeks, before the year is out.

These implications are not complicated and neither is the action to implement higher standards for schooling. From one perspective, it is fairly clear that more than 90% of elementary school students learn basic multiplication and division facts by the end of grade 5 and that upwards of 80% of them eventually learn how to do routine tasks involved in addition and multiplication and probably subtraction. These are important accomplishments, but they're qualified. The phrase ". . . eventually learn how to do routine tasks . . ." is a big qualifier that often upsets well-intended efforts to set higher standards and, more important, to realize them. From another perspective, schooling looks a little different. By the end of grade 6, almost 20% of all elementary

[1]For information regarding issues involving concentration of practice on pivotal mathematics skills at different grade levels, see Perkins (1983) and Buchanan, Schutz, and Milazzo (1983).

students are unable to demonstrate the most straightforward tasks in addition and multiplication that involve carrying. More than 20% of them can't do basic subtraction. Moreover, fairly high levels of performance at grades 4 and 5 in the PVS and Essential Skills data bases (where problems are not designed to be especially tricky and students have plenty of time) deteriorate badly when the conditions of performance are unusual, as they are with NAEP assessments, or when speediness becomes a factor, as it does with CTBS. When "performance" moves away from basic forms of computation and into tasks that have special complications, such as "zeros in the minuend" in subtraction or "zeros in the quotient" in division, things get worse yet.

Overall, the results show that, through grade 6, students can do a lot of different routine tasks involved in computation, but the reliability of their performance is fragile. A pedagogical sledgehammer, even though wielded under the banner of "excellence," is not the right tool for improving the performance. Bringing all or nearly all students to a status where computation algorithms are completely reliable--subject only to careless or random error--has been a historic struggle for schools. In the long range, it should at least be considered that some other approach now might make better use of the limited amount of time that's available for teachers to teach and students to learn. In the past, there was no reasonable alternative to learning complete, full-blown algorithms for computation; now, with calculators and microcomputers, there is.

For now, however, human memory-driven computation is the dominating factor in mathematics instruction for the elementary grades. The question is how much to expect. Even if schools can push performance levels on computation problems to 85-90%, there is still a question of timing. It is not enough that students "eventually learn" how to compute. Many learning opportunities that elementary schools provide at one grade have strong linkages to skills that will come at the next grade. Little that is of any consequence is learned in isolation. As it is now, opportunities missed are likely to be opportunities lost for a sizable proportion of learners. Timing is critical.

From any perspective, the process of elementary schooling is complicated. Pursued intelligently, the successful achievement of educational excellence in any school subject requires more than the efforts of individual teachers teaching small groups of children. A simplistic view of what teachers and principals need to do to promote excellence ill serves all parties--students, teachers, administrators, and the public. Pushing teachers to try harder is not enough. All in all, better articulation of learning opportunities across grade levels has a lot to do with raising standards and promoting excellence in the elementary grades. Intensity is a critical part of any worthwhile effort--but so is focus and timing.

APPENDIX A

DESCRIPTION OF TOPICS INCLUDED IN THE ANALYSIS

All six topics in the analysis involve computation with whole numbers. Each of the topics receives a substantial number of textbook lessons, compared to other topics that are taught at about the same grade levels, and each one becomes embedded in instruction on some other topic a grade level or so later.

Topic 1: Addition of 2-, 3-, and 4-digit Numbers (with carrying) Addition of 2-digit numbers is introduced in some textbooks by the end of grade 1, but it doesn't require any regrouping of place values (carrying). Addition with carrying begins in the last half of grade 2 and is expanded in grade 3 to include 3-digit and often 4-digit numbers. This topic is reviewed in grade 4 and is usually extended to addition of numbers having more than four digits, but instruction usually amounts to no more than two or three lessons. Textbooks for grades 5 and 6 also include a small amount of review in most textbook series.

Topic 2: Subtraction of 2-, and 3-digit Numbers (with borrowing) Subtraction of 2-digit numbers with borrowing begins at the end of grade 2. It is extended at grade 3 to include 3-digit numbers, but sometimes the subtraction does not involve much work with minuends that have zeros (302 - 215) until grade 4. Nevertheless, the instruction on subtraction with zeros and subtraction with 4-digit numbers is completed by the end of grade 4 and is reviewed briefly, along with addition, in grades 5 and 6.

Topic 3: "Hard" Multiplication Facts This topic includes multiplication of the numbers 6, 7, 8, and 9 by the numbers 5, 6, 7, 8, and 9. Instruction on multiplication facts begins in most textbooks at grade 2 with easy facts such as 3 x 2 and 3 x 4. Instruction on hard facts begins in grade 3, usually in the second half of the textbook, and is repeated in grade 4. Often, there is a small amount of review provided near the beginning of grades 5 and 6, but the bulk of formal instruction is completed in grade 4.

Topic 4: "Hard" Division Facts These facts parallel the multiplication facts identified above (e.g. 54 : 6 and 42 : 7). Instruction on hard division facts begins late in grade 3, usually after all of the work on hard multiplication facts is finished. Instruction is redone in grade 4, usually in the first quarter of the textbook. A small amount of review and practice is provided early in grades 5 and 6 together with practice on multiplication facts.

<u>Topic 5: Multiplication by 1-digit and 2-digit Numbers</u> Instruction on multiplication by 1-digit numbers often begins late in grade 3 and extends well into grade 4. Multiplication by 2-digit numbers usually begins late in grade 4 and extends into grade 5. Although multiplication by 1-digit and 2-digit whole numbers is reviewed early in grade 6, there are other grade 6 topics, such as multiplication by decimals and multiplication by 3-digit numbers, that build directly upon them.

<u>Topic 6: Division by 1-digit and 2-digit Numbers</u> Instruction on division by 1-digit numbers also begins near the end of grade 3 in many textbooks and is covered thoroughly by all elementary school textbooks in grade 4. Division by 2-digit numbers is introduced near the end of grade 4, although many teachers defer instruction until the topic is reintroduced in grade 5.

APPENDIX B

DESCRIPTION OF DATA BASES EMPLOYED IN THE ANALYSIS

- The major data base used for grades 1-6 includes information from SWRL's Proficiency Verification System (PVS) collected at the beginning, middle, and end of the 1978-79 school year. PVS inventories involve over 200 assessment items at each of grades 1-6. Development of the items was based on a careful analysis of how topics are introduced grade-by-grade in mathematics textbooks that schools use most. One feature of PVS inventories is especially useful for analyzing students' performances on similar assessments across two or three adjacent grade levels. Many mathematics problems that are the same, or very nearly so, appear in PVS inventories for contiguous time periods. Just as there is overlap in instruction on the same topic across two or more grade levels, there is some overlap in items across levels of PVS inventories. For example, multiplication and division facts are part of PVS inventories used at the middle of the school year in grades 3, 4, and 5; they are included in inventories for the beginning of the school year at grades 4, 5, and 6; and they are part of the end-of-year inventories for grades 3 and 4. During the 1978-79 school year, over 1,000 students at each grade level took each of the three PVS inventories as part of their school's regular testing program. Almost all of the students were located in three large school districts whose school populations included a wide range of student capabilities. All of the students in each of these districts were involved in the PVS program. Using the PVS data base we could look at student performances on similar items within each topic beginning at mid-year in either grade two or grade three and, depending on the topic, continuing on through the beginning, middle, and end of grade 5 or grade 6.

- Our alternate data base includes the results from three assessments of essential skills given in the spring of 1979, 1980, and 1982. The assessments were specially written to fit the objectives of a large urban school district's instructional program. Each assessment consists of approximately 50 mathematics items at each of grades 1 through 6. More than 25,000 students at each grade level are administered these assessments at the end of the school year. In most instances, the six topics we wanted to follow are assessed about one grade level later in the Essential Skills data base than they are in PVS. This happens because of the logic that underlies the assessment. The intention of the school district is to give as much opportunity as possible for instruction on each objective to take place before students are assessed. As a result, most of the skills are assessed one to two grade levels behind where they are introduced and developed in most textbooks.

- The NAEP data base comes from the second nationwide mathematics assessment, which was conducted in 1978. It represents a large sample of students at ages 9, 13, and 17 from across the nation. The NAEP data base has a unique structure that enhances the utility of the data for some kinds of analysis, but limits it for others. One feature that enhances an analysis of student growth on routine academic tasks is the huge variety of tasks that students are asked to do in each NAEP assessment. Some tasks are very difficult for students at a particular age level, while others are very easy. The nice thing about this feature is that one can almost always find an item in the NAEP data base to represent some topic that is the focus of some research on student performance.

  A major limitation of the NAEP data base for 1978 is that all assessments were administered to students who were all of the same age but not at the same grade level in school. Therefore, data on 9-year-old students mostly includes fourth graders, but it also includes a fairly large number of third graders who were 9 years old in January and February of 1978. In addition, the NAEP data base for 9 year olds includes a small number of students who were in grades 2, 5, and 6 at the time of the assessment. NAEP, in its published reports, has mainly provided results for each of its three age groups; very little information has been published about results for different grade levels within these age groups. Fortunately, the data tapes which NAEP has made available to researchers have codes that identify the grade level of students taking the assessment. Therefore, it <u>is</u> possible to look at student performance for different grade levels as well as different chronological ages. One does have to exercise some caution, however. Performance levels of 9-year-old fourth graders <u>do</u> include most of the students who are supposed to be in fourth grade, but they would not include very bright 8-year-olds, who might be in the fourth grade. Even more likely, NAEP results for 9-year-olds won't include any 10-year-old students who are still in fourth grade because they are seriously behind.

  In our analysis, we looked mainly at data for fourth graders, who made up about two-thirds of the NAEP sample of 9-year-olds, and eighth graders, who made up about two-thirds of the NAEP sample of 13-year-olds. We did not look at data for the NAEP sample of 17-year-olds, because these students have been out of elementary school too long for NAEP results to tell us much about effects of schooling in the elementary grades.

- The CTBS-derived data base includes p values, the proportion of students answering a test question correctly, obtained from a national sample of school students by CTB/McGraw-Hill when various levels of Form U of this test were normed in the spring of 1980. The p values used in our analysis actually represented the "proportion of students answering correctly <u>who attempted to answer a particular item</u>." This kind of p value excludes students who did not try to answer the particular item. We used the "attempted" p value, because the CTBS tests are carefully timed. Some slow-working students who might have answered correctly on items that occurred near the end of the test never got to them. Obviously, they didn't answer these items correctly, but we don't know that all slow working students would have answered them incorrectly, either. About the best we can do is to assume that the "attempted" p values for items near the end of the test may be a little high as indicators of the proportion of all students who would have answered correctly if the test had not been timed.

Altogether, our CTBS-derived data base included information from approximately 80 to 90 items on Levels D through F of Form U. At each level, the items were from two subtests: Mathematics Computation and Mathematics Concepts and Applications.

APPENDIX C

DESCRIPTION OF ANALYTICAL PROCEDURES

Using the PVS data base we moved assessment by assessment from the beginning, to the middle, to the end of the school year, and we repeated this process across several contiguous grade levels to see how the level and the range of student performance increased together. Along the way, we supplemented various PVS-derived data points with student performance levels on similar items from our Essential Skills data base. Going further, we included data points for coordinating items on NAEP results for 9-year-old fourth graders. In some cases, the items we show for NAEP were items actually used in the assessment; in others, the item we show only illustrates an NAEP item type because the real NAEP item is still restricted and cannot be published. Beyond grade 6, we used items from NAEP data for 13-year-olds in grade 8 to observe levels of performance and, to some extent, ranges of performance several grade levels after regular instruction would have (or should have) been completed.

We used results on CTBS items to provide some additional information on performance levels of students in grades 3 and 4 and to fill in some gaps in grades 6 and 7. Not a lot of items in CTBS are useful for this purpose, because some topics, such as hard multiplication and division facts, simply aren't represented very extensively in any of the tests in CTBS Form U. In addition, many items that are included in the computation portions of CTBS represent extremes in item difficulty. For example, division of a 3-digit number by a 1-digit number, one of the early stages in learning about division beyond the basic facts, is represented in CTBS by two kinds of division tasks. One, for example 3)636, is extremely simple because it involves only the use of basic facts and requires no subtraction. The other, for example 3)311, is very tricky because it starts out with a deceptively simple use of basic facts but requires students to work with a zero in the quotient. This kind of task represents one of the most complicated forms of division by a 1-digit number. For the CTBS analysis, the items we show are coordinate items that have numbers of about the same size and involve computations that require about the same number of steps to do.