DOCUMENT RESUME

ED 301 325 PS 017 624

AUTHOR Gottfredson, Gary D.
TITLE You Get What You Measure, You Get What You Don't: Higher Standards, Higher Test Scores, More Retention in Grade. Report No. 29.
INSTITUTION Center for Research on Elementary and Middle Schools, Baltimore, MD.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE Aug 88
GRANT OERI-G-86-90006
NOTE 29p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Standards; Educational Change; *Educational Improvement; Educational Practices; *Educational Testing; Elementary Secondary Education; *Grade Repetition; *Incidence; School Districts; *Standardized Tests; State Programs; Student Promotion

ABSTRACT

This study examined one urbanized Southeastern county school district's experience in attempting to implement an educational reform program that called, among other things, for higher standards. By focusing on standardized norm-based and criterion-referenced educational tests administered to students according to grade level, this school system had: (1) been misled about the magnitude of educational improvements produced by its emphasis on high standards; (2) increased grade retention; and therefore (3) plausibly increased the risk of the school dropout and other varieties of adolescent problem behavior among the community's youths. Retention was especially extensive in grades 9 (33 percent, two-thirds of students promoted), 1 (25 percent), 10 (23 percent), and 7 (22 percent). The percentage of students still in school who were of ninth-grade age who were actually in ninth grade or a higher grade was only 47. Because grade retention is a very potent predictor of dropout, high rates of dropout are to be expected in this district when students age beyond the compulsory school attendance age. The study also shows that precisely what one measures when holding educational institutions to high standards may determine what the institutions produce, and that what one leaves unmeasured may open the door for undesired outcomes of a reform until these outcomes become painfully obvious. (RH)

Center for Research On Elementary & Middle Schools

Report No. 29

August, 1988

YOU GET WHAT YOU MEASURE, YOU GET WHAT YOU DON'T: HIGHER STANDARDS, HIGHER TEST SCORES, MORE RETENTION IN GRADE

Gary D. Gottfredson

Center Staff

Center Liaison

National Advisory Board

You Get What You Measure, You Get What You Don't:

Higher Standards, Higher Test Scores, More Retention in Grade

Grant No. OERI-G-86-90006

Gary D. Gottfredson

Report No. 29

August 1988

Published by the Center for Research on Elementary and Middle Schools, supported as a national research and development center by funds from the Office of Educational Research and Improvement, U.S. Department of Education. The opinions expressed in this publication do not necessarily reflect the position or policy of the OERI, and no official endorsement should be inferred.

Center for Research on Elementary and Middle Schools
The Johns Hopkins University
3505 North Charles Street
Baltimore MD 21218

Printed and assembled by:
VSP Industries
2440 West Belvedere Avenue
Baltimore MD 21215

The Center

The mission of the Center for Research on Elementary and Middle Schools is to produce useful knowledge about how elementary and middle schools can foster growth in students' learning and development, to develop and evaluate practical methods for improving the effectiveness of elementary and middle schools based on existing and new research findings, and to develop and evaluate specific strategies to help schools implement effective research-based school and classroom practices.

The Center conducts its research in three program areas: (1) Elementary Schools; (2) Middle Schools, and (3) School Improvement.

The Elementary School Program

This program works from a strong existing research base to develop, evaluate, and disseminate effective elementary school and classroom practices; synthesizes current knowledge; and analyzes survey and descriptive data to expand the knowledge base in effective elementary education.

The Middle School Program

This program's research links current knowledge about early adolescence as a stage of human development to school organization and classroom policies and practices for effective middle schools. The major task is to establish a research base to identify specific problem areas and promising practices in middle schools that will contribute to effective policy decisions and the development of effective school and classroom practices.

School Improvement Program

This program focuses on improving the organizational performance of schools in adopting and adapting innovations and developing school capacity for change.

This report, prepared by the School Improvement Program, examines the effects of an urban school system's implementation of higher standards on retention rates and dropout in the system.

Abstract

This study finds that the implementation of educational reforms in a large urbanized country school district, especially a focus on standardized norm-based and criterion-referenced educational tests administered to students according to grade level, has shown improved test scores, but this improvement has been accompanied by increased grade retention of students.

Introduction

During the first half of the decade of the 1980s, calls for higher educational standards created something of a sense of urgency about educational reform (e.g., National Commission on Excellence in Education, 1983). The nation was at risk, according to the *Zeitgeist*, because of lax standards, watered down curricula, and an educational system that robbed from the able and serious-minded students to give to the rowdy and uninterested who would really rather not be in school. Several states passed education reform acts and many local education agencies engaged in activities to raise educational standards.

McDill, Natriello, and Pallas (1986) have suggested that part of this emphasis on educational standards was a backlash generated by perceptions that an emphasis on equity during the late 1960s and the decade of the 1970s had eroded academic standards. A focus on basics and minimum competency testing could stanch this erosion. McDill et al. also cited three ways raised standards might produce unintended outcomes: (1) raised standards could increase academic stratification and decrease student choice, (2) greater demands on student time could increase conflicts between school and other sources of demands on students, and (3) higher required standards could lead to more student failure. These potential outcomes, if they occured, would increase the risk of dropout for students who are already at elevated risk of dropping out.

The application of minimum competency tests in grade-to-grade promotion or graduation decisions, in the absence of real improvements in instruction and learning, would be particularly likely to decrease the likelihood of persistence in school to high school graduation for marginal students.

Purpose of the Study

This study examines one school system's experience in attempting to implement an educational reform program that called, among other things, for higher standards. By focusing on standardized norm-based and criterion referenced educational tests administered to students according to grade level, this school system has (a) been misled about the magnitude of educational improvements produced by its emphasis on high standards, (b) increased grade retention, and therefore (c) plausibly increased the risk of the school dropout and other varieties of adolescent problem behavior among the community's youths. It seems likely than many other school systems have done the same.

The study also shows that precisely what one measures when holding educational institutions to high standards may determine what the institutions produce, and that what one leaves unmeasured may open the door for undesired outcomes of a reform until these outcomes become painfully obvious. Put another way, you get what your measure and you may get what you don't.

This research was conducted in an urbanized county school district in a Southeastern state that had passed an education reform act earlier in the decade. The Act called for testing all students in certain grades (1, 2, 3, 5, and 8) with a criterion-referenced competency test and all students in other grades (4, 5, 7, 9, and 11) with a norm-based standardized achievement test. The state recommended standards for promotion from grade to grade based partly on these tests, as did the county school board.

This is only one part of the more general "reforms" introduced as a part of the reform legislation and the climate of the times. Other parts included the allocation of a considerable amount of new money for educational improvement, the requirement that all schools employ a school improvement team to assess school needs and make plans to address them, some support for technical assistance in school improvement and for training, and attention to principal job performance.

The testing programs used to provide evidence about the extent to which schools were applying stringent educational standards and to assess the consequences of the general reform were administered according to grade level. That is, students enrolled in grade 1 take a first-grade competency-based test, students enrolled in grade 8 take an eighth-grade competency-based test, students enrolled in grade 9 take a ninth-grade norm-reference test, and so on. The majority of school systems in American have similar testing programs.

Improvements in Test Scores

Figures 1 through 7 show this district's improvements in student test scores in reading and math. The median national percentile rank for the norm-referenced test for seventh grade reading went from 39 in 1983 to 53 in 1987, for example. Similar gains were made in all areas of assessment according to these norm-referenced tests.

Gains were also made on the criterion-referenced mastery test developed by the State Department of Education: The percentage of sixth graders meeting or exceding the state-set mastery criterion rose from 54% in 1981 to 78% in 1986, for example. The improvements appeared to increase with increasing grade level.

Promotion Rates and Grade Retention

This district's improvement in student test scores is impressive, but the district had concerns about a possibly related problem -- the presence of "overage" students in many classrooms due to retention in grade.

To estimate the rates of retention in grade (and therefore the rates of grade promotion), we determined the actual grade assignment of the students who were enrolled in any county school in 1986

and 1987. Students who were in a higher grade in 1987 than in 1986 were deemed to have been promoted and their counts established the numerator for the promotion rate (the denominator being the number of students enrolled both years). <1> The retention percentages are shown in Figure 8. Retention is especially extensive in grades 9 (33%, two-thirds of students promoted), 1 (25%), 10 (23%), and 7 (22%).

These rentention rates are incidence rates. The lifetime prevalence rates of grade retention for students in this school district is (for all grades above first) is much higher than the rates shown in Figure 8. For example, the percentage of students still in school who are of ninth-grade age who are actually in ninth grade (or a higher grade) is only 47, implying that the prevalance of grade retention by about age 14 is close to 53% in this district. Because grade retention is a very potent predictor of dropout (Bachman, Green, & Wirtanen, 1971), high rates of dropout are to be expected in this district when students age beyond the compulsory school attendance age. <2>

Recall that the test score data imply remarkable progress for this school district over the period of time that tougher standards have been applied. The percentage of eighth graders meeting or exceeding the state standard for reading on the criterion-referenced test rose from about 48% in 1981 to about 78% in 1986, for example. But recall also that the eighth grade test is given only to persons in the eighth grade. But calculations show that fewer than half of the students of eighth-grade age are actually in the eighth grade (or a higher grade). More than half of the students who are around age 13 have been retained one or more times. Only about 42% of students of eighth-grade age meet or exceed the state criterion for the eighth grade at the age appropriate for that grade. Figure 9 shows

---

<1>This is only one way to estimate grade retention and it has defects (e.g., it excludes from consideration persons who left the school system). Surveys of principals conducted by the district education agency provide another method but also incorporate defects (e g., reporting errors and promotion due to summer school completion is not captured). The two methods agree reasonably well despite defects of each.

<2> A completion rate for this district estimated using the number of students enrolled in grade 12 in 1987 compared to the number of students in grade 5 seven years earlier is 54%. The dropout rate (uncorrected for in- and out-migration) is therefore *higher* than 46% (some 12th graders will not graduate).

the percentage of students who--according to their ages--should be in various grades who meet or exceed the criterion score on the state mastery test for that grade level. This Figure portrays a picture of educational progress that is quite different from that suggested by Figures 1 through 7.

Examinee Age

In this school district, the age of students taking specific grade-level examinations have been increasing in recent years. Figure 10 illustrates this phenomenon for the eighth grade by showing on a single plot the increase in percentage of eighth graders meeting or exceeding the state criterion scores for reading and math and the decrease in the percentage of eighth grade examinees who are at the appropriate age for that grade (i.e., who are not overage). For some elementary and middle school grades, the mean age of students increased as much as three-tenths of a year over the mean ages in 1980 and 1981, although for most grades the increase is less dramatic.

Part of the apparent increase in scholastic achievement reflected in the grade-specific test scores in this district is due to an increase in the average age of examinees. Put bluntly, the population tested has shifted substantially as a result of grade retention (selection) based explicitly on closely related scholastic tests, at least in part. The actual changes in age-specific scholastic achievement of students, if any, are unknown in this district. The same may be true in many districts around the country that have implemented reforms accompanied by more rigorous standards for grade-to-grade promotion.

Grade Retention Before and After the Reform

The percentages of students remaining enrolled in school who were at the grade appropriate for their ages in 1980 (before the reform) and in 1987 are shown in Figure 11. The general picture for both years is much the same. This district has a grade retention problem of long-standing which has been exacerbated in recent years.

A Gereral Implication for Organizational Outcomes

In the course of exploring the grade retention problem with this school district, we conducted many interviews with principals, assistant principals, superintendents, counselors, and teachers. When we asked about educational goals, few respondents hesitated to mention improving test scores. When grade retention was mentioned, it was mentioned as a problem, not as a target for problem solving. These interview observations are in line with an interpretation that this district has focused on improving test scores, has had no systematic and highly visible way of monitoring other outcomes (such as grade retention) and it has got what it measured and got what id didn't: test scores in grade have gone up but so has the age of the population tested in each grade.

Organizational psychologists tell a yarn, perhaps apocryphal and perhaps true, about an organizational psychologist who was doing consulting work on productivity for a silver mining company. According the the story, the psychologist cited the maxim, "you get what you measure." He suggested posting daily figures on the tons of ore removed from the mine to provide feedback to the miners on their work performance. The tons of rock went up; the silver content of the rock went down. The moral of the story is that if you get what you measure, then you better make sure that what you measure includes the things you want to get.

REFERENCES

Bachman, J. G., Green, S., & Wirtanen, I. D. (1971). Dropping out--Problem or symptom (Vol. III, Youth in Transition). Ann Arbor, MI: Institute for Social Research.

McDill, E. L., Natriello, G., & Pallas, A. M. (1986). A population at risk: Potential consequences of tougher school standards for student dropouts. American Journal of Education, 94, 135-131.

National Commission on Excellence in Education. (1983). A nation at risk: The imperative for educational reform. Washington, DC: U.S. Government Printing Office.

Figure 1

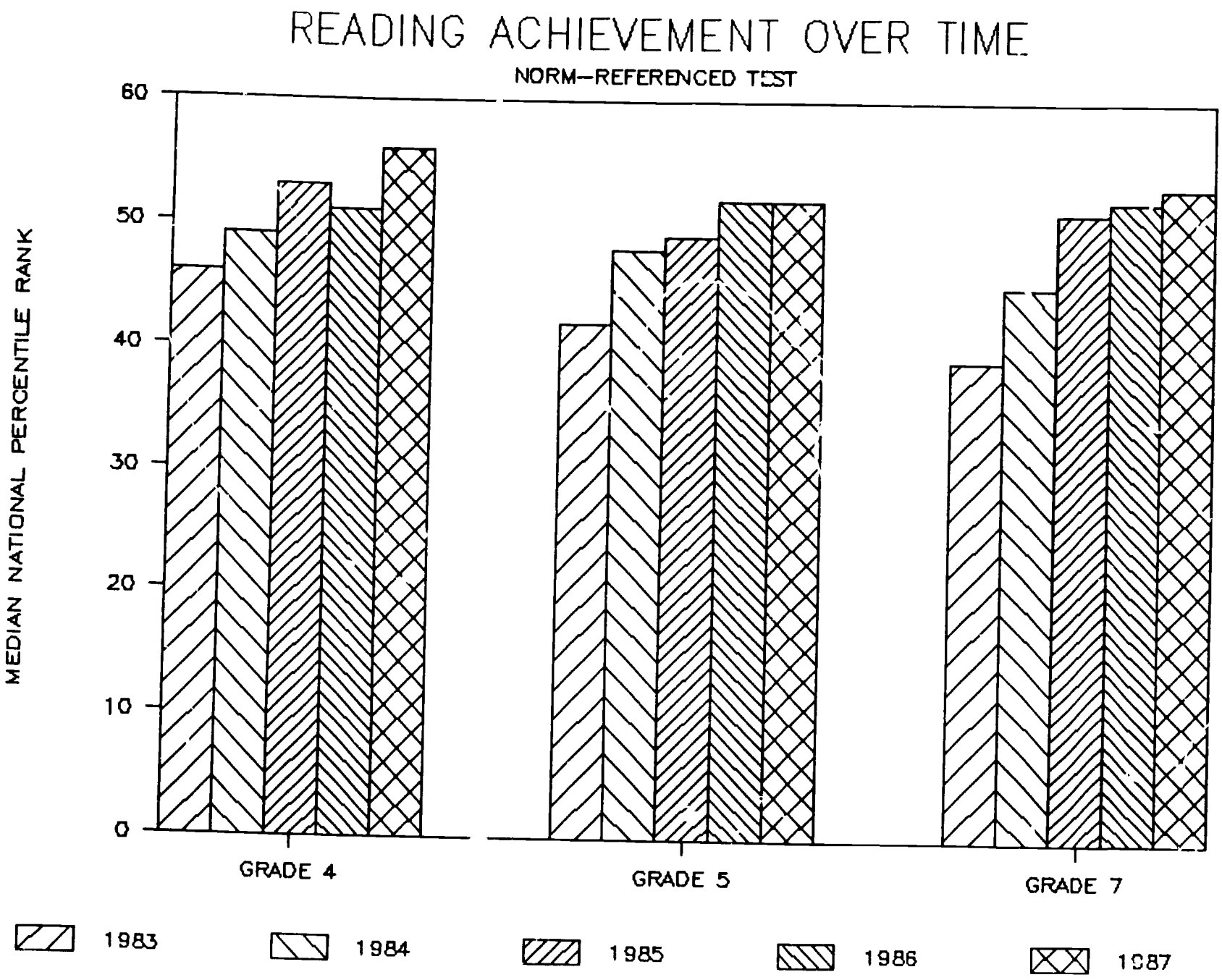

READING ACHIEVEMENT OVER TIME
NORM-REFERENCED TEST
MEDIAN NATIONAL PERCENTILE RANK
60
50
40
30
20
10
0
GRADE 4
GRADE 5
GRADE 7
1983
1984
1985
1986
1987

Figure 2

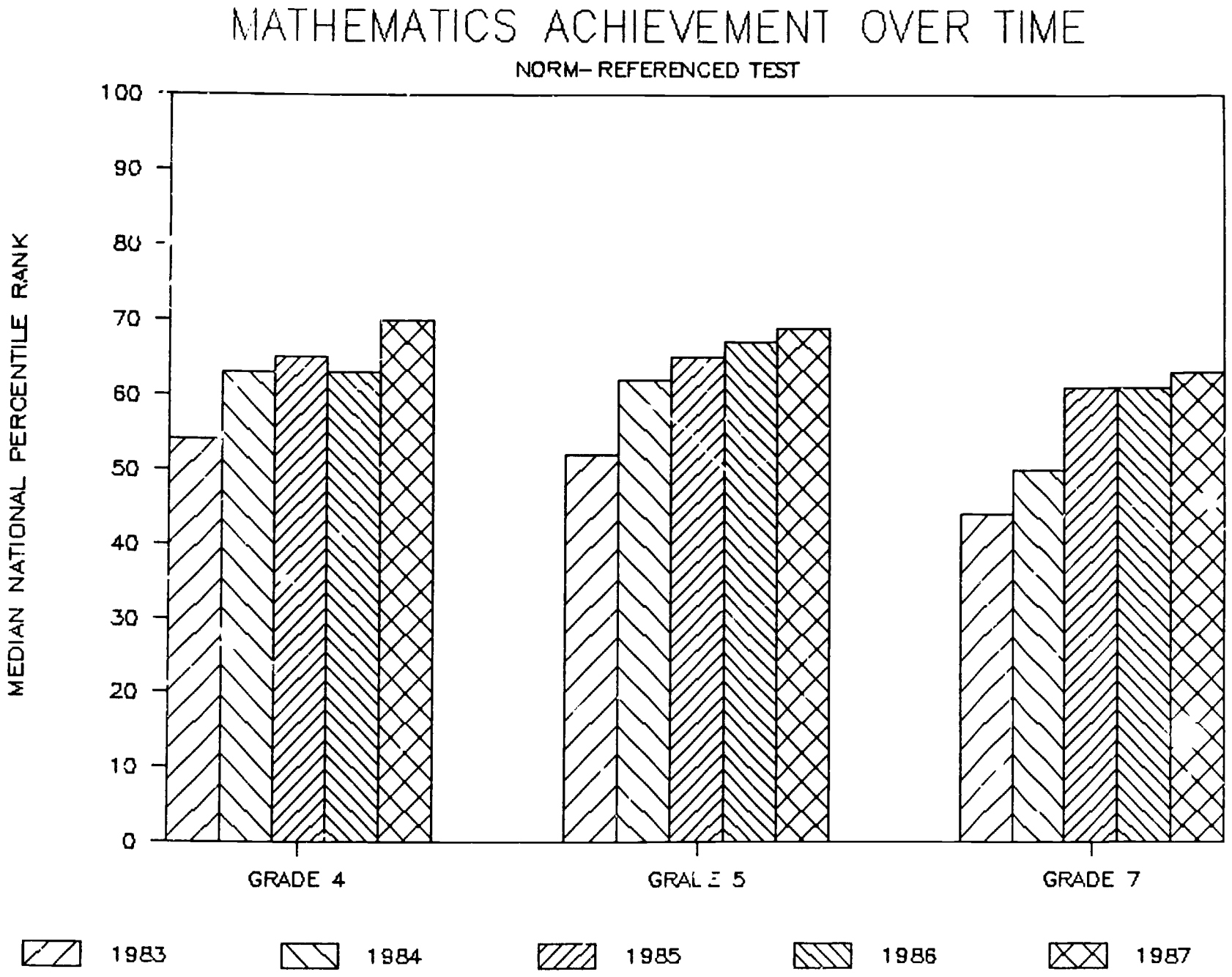
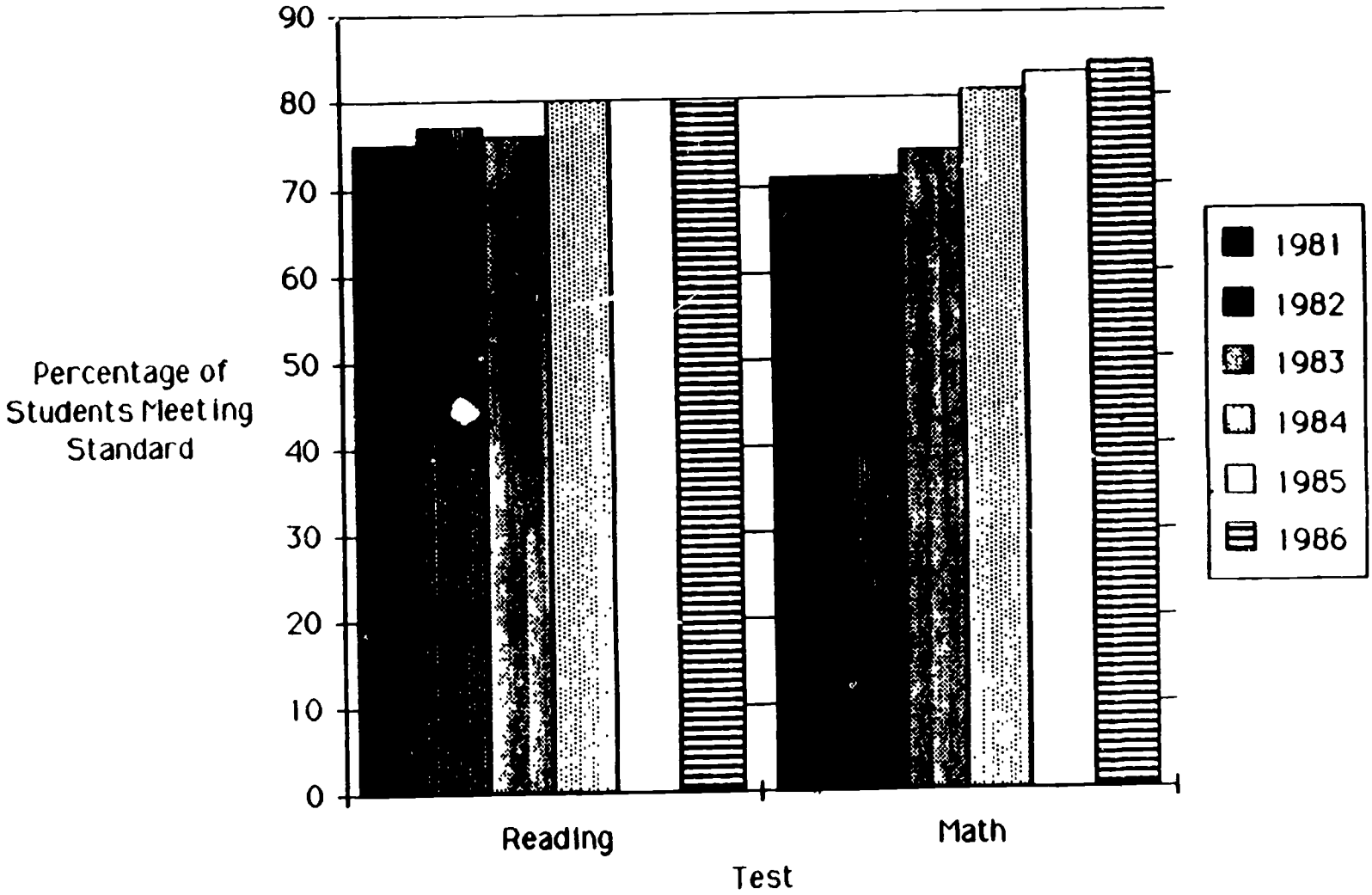
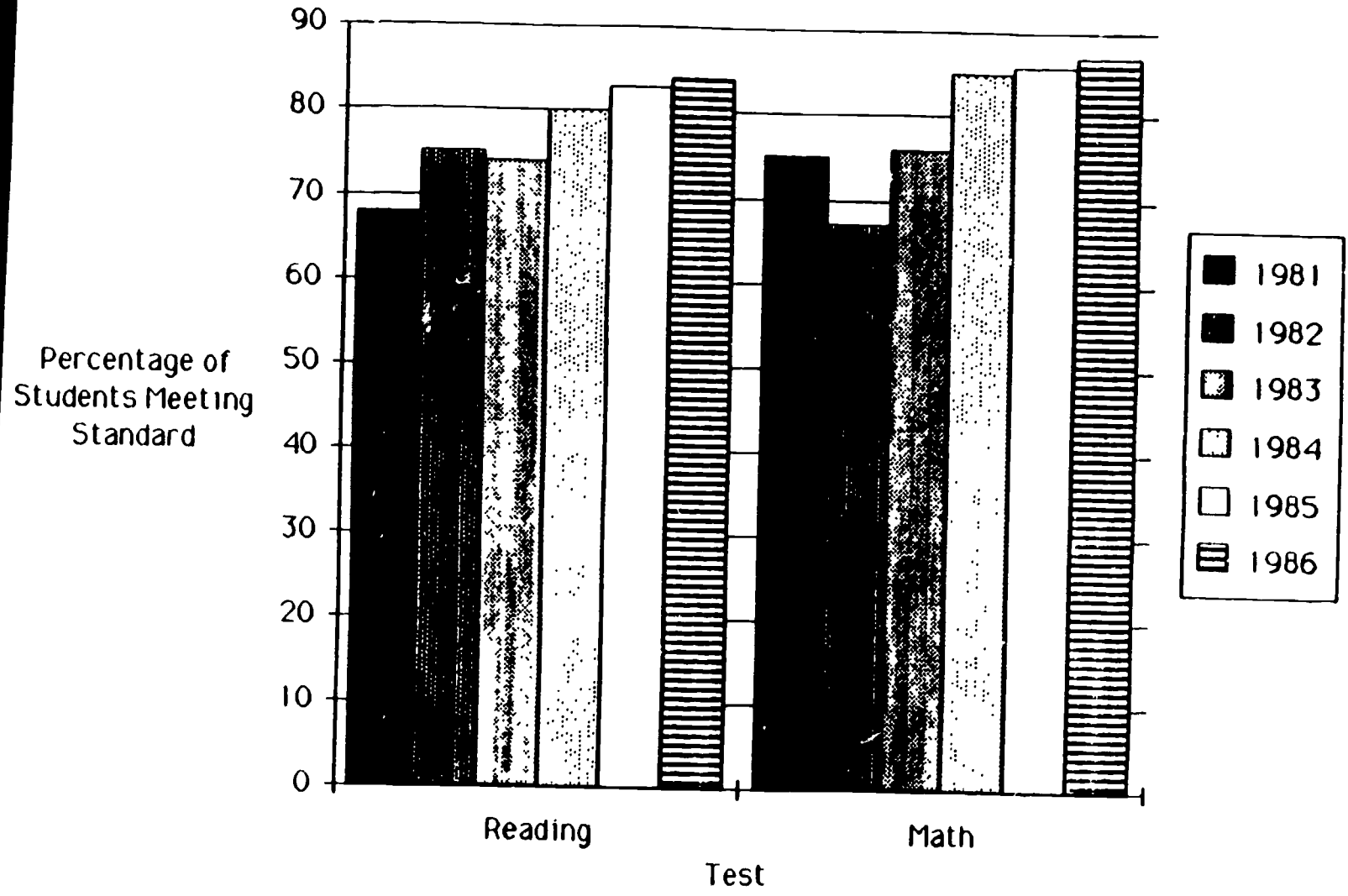

MATHEMATICS ACHIEVEMENT OVER TIME
NORM-REFERENCED TEST
MEDIAN NATIONAL PERCENTILE RANK
100
90
80
70
60
50
40
30
20
10
0
GRADE 4
GRADE 5
GRADE 7
1983
1984
1985
1986
1987

Figure 3



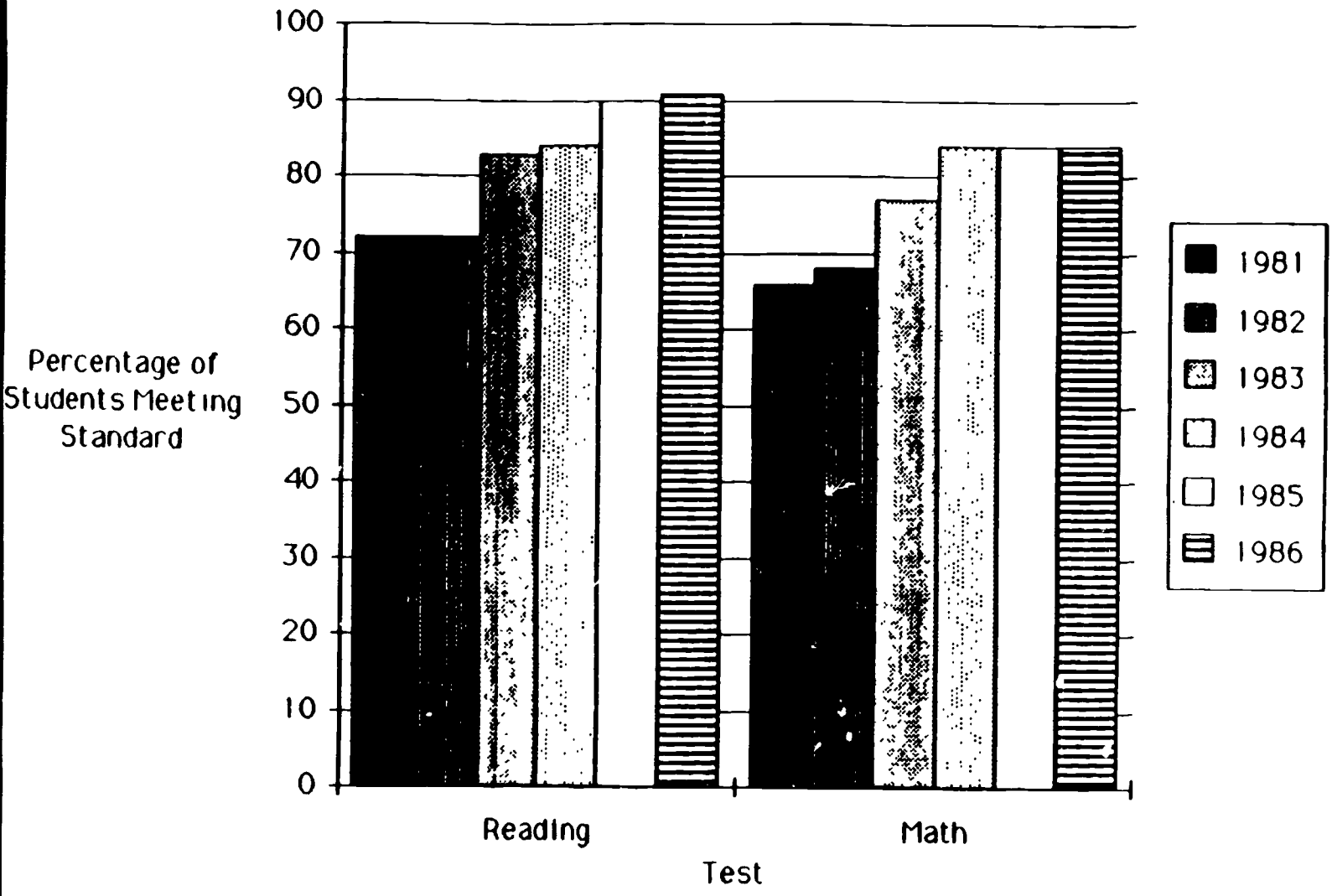
CRITERION-REFERENCED MASTERY TEST, GRADE 1

Figure 4



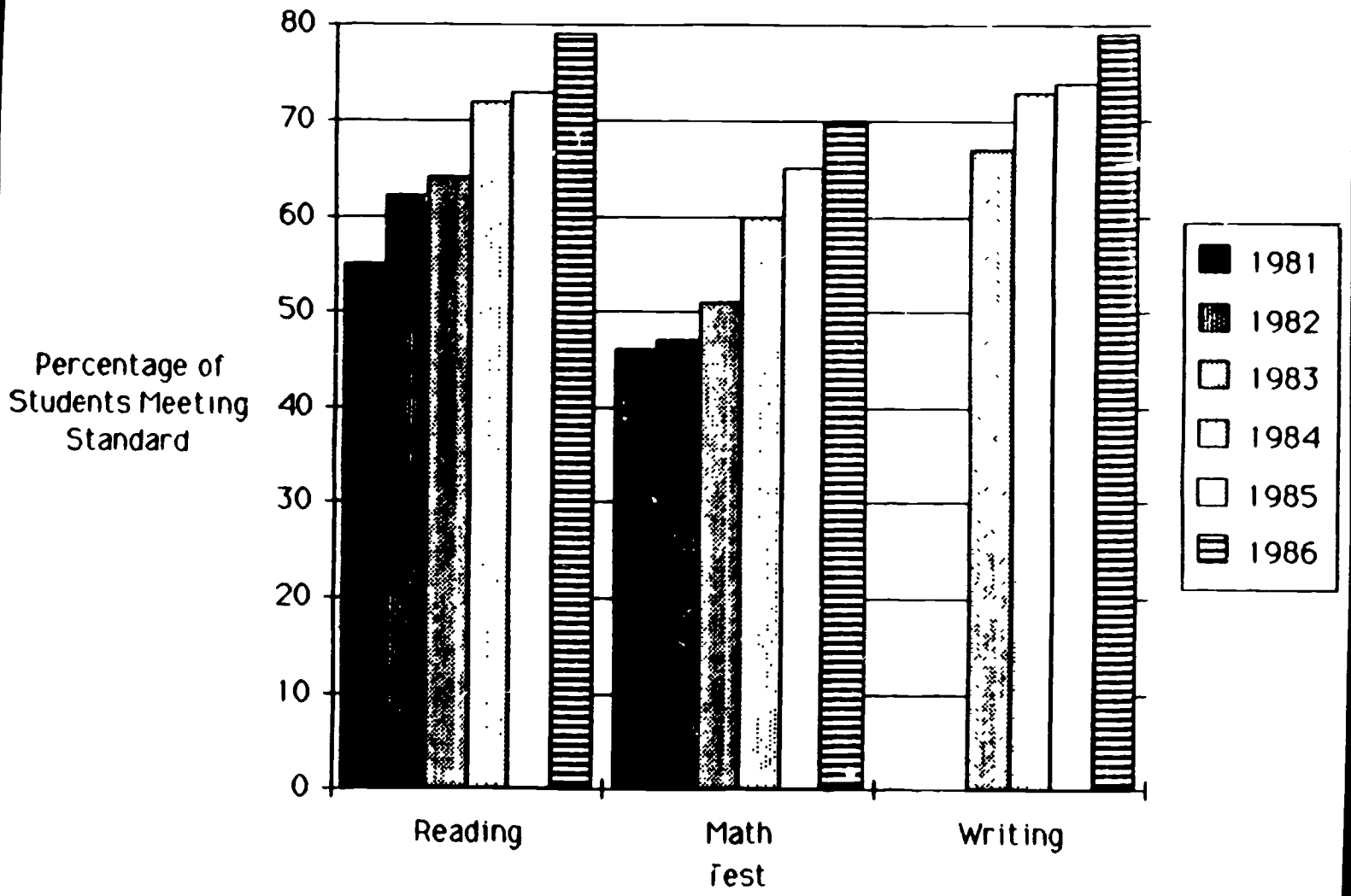
CRITERION-REFERENCED MASTERY TEST, GRADE 2

Figure 5



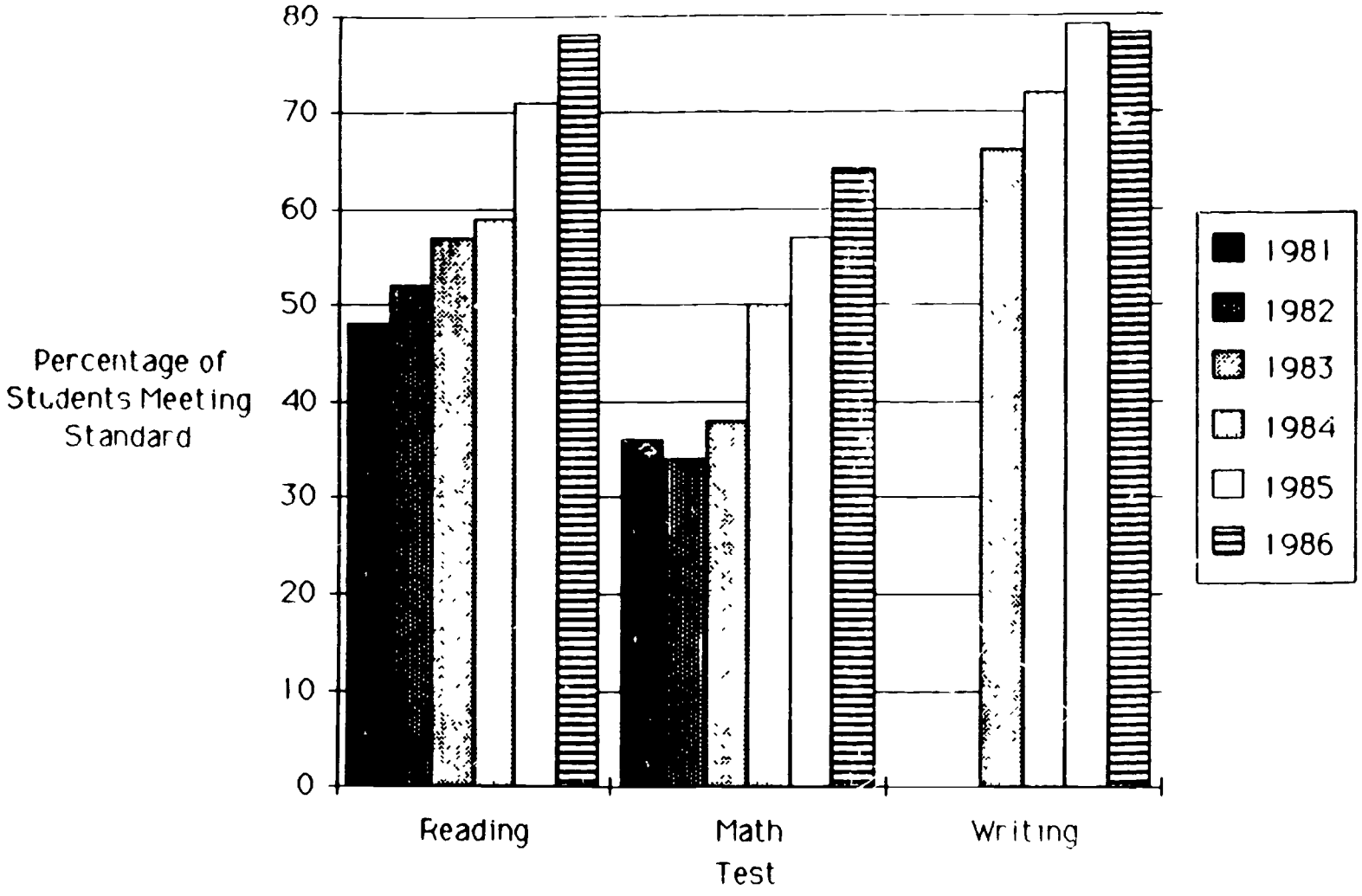
CRITERION-REFERENCED MASTERY TEST, GRADE 3

Figure 6



CRITERION-REFERENCED MASTERY TEST, GRADE 6

Figure 7



CRITERION-REFERENCED MASTERY TEST, GRADE 8

Figure 8

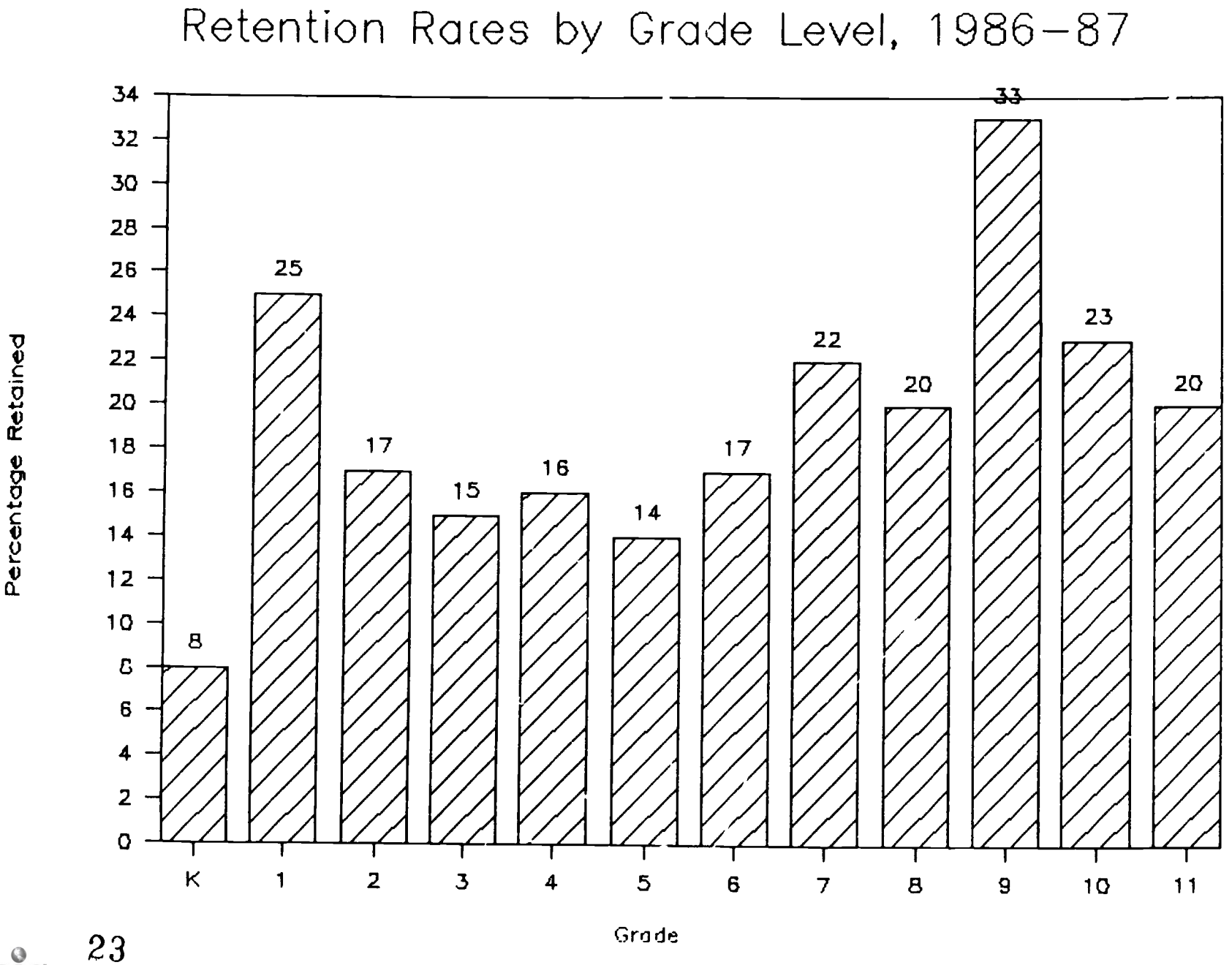

Retention Rates by Grade Level, 1986-87
Percentage Retained
34
32
30
28
26
24
22
20
18
16
14
12
10
8
6
4
2
0
8
25
17
15
16
14
17
22
20
33
23
20
K
1
2
3
4
5
6
7
8
9
10
11
Grade

Figure 9

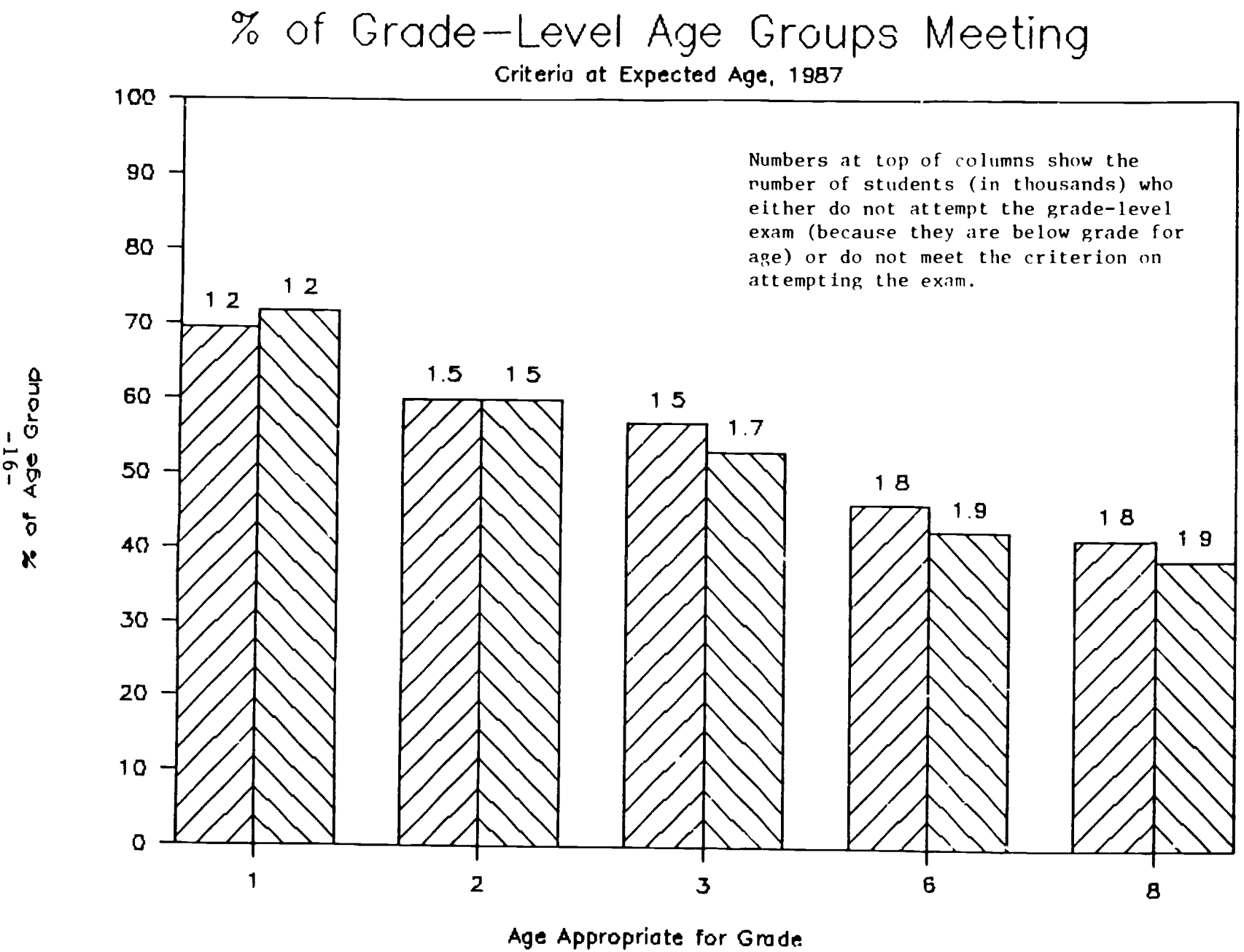

% of Grade-Level Age Groups Meeting
Criteria at Expected Age, 1987
Numbers at top of columns show the number of students (in thousands) who either do not attempt the grade-level exam (because they are below grade for age) or do not meet the criterion on attempting the exam.
% of Age Group
100
90
80
70
60
50
40
30
20
10
0
1.2
1.2
1.5
1.5
1.5
1.7
1.8
1.9
1.8
1.9
1
2
3
6
8
Age Appropriate for Grade
READING
MATHEMATICS

Figure 10

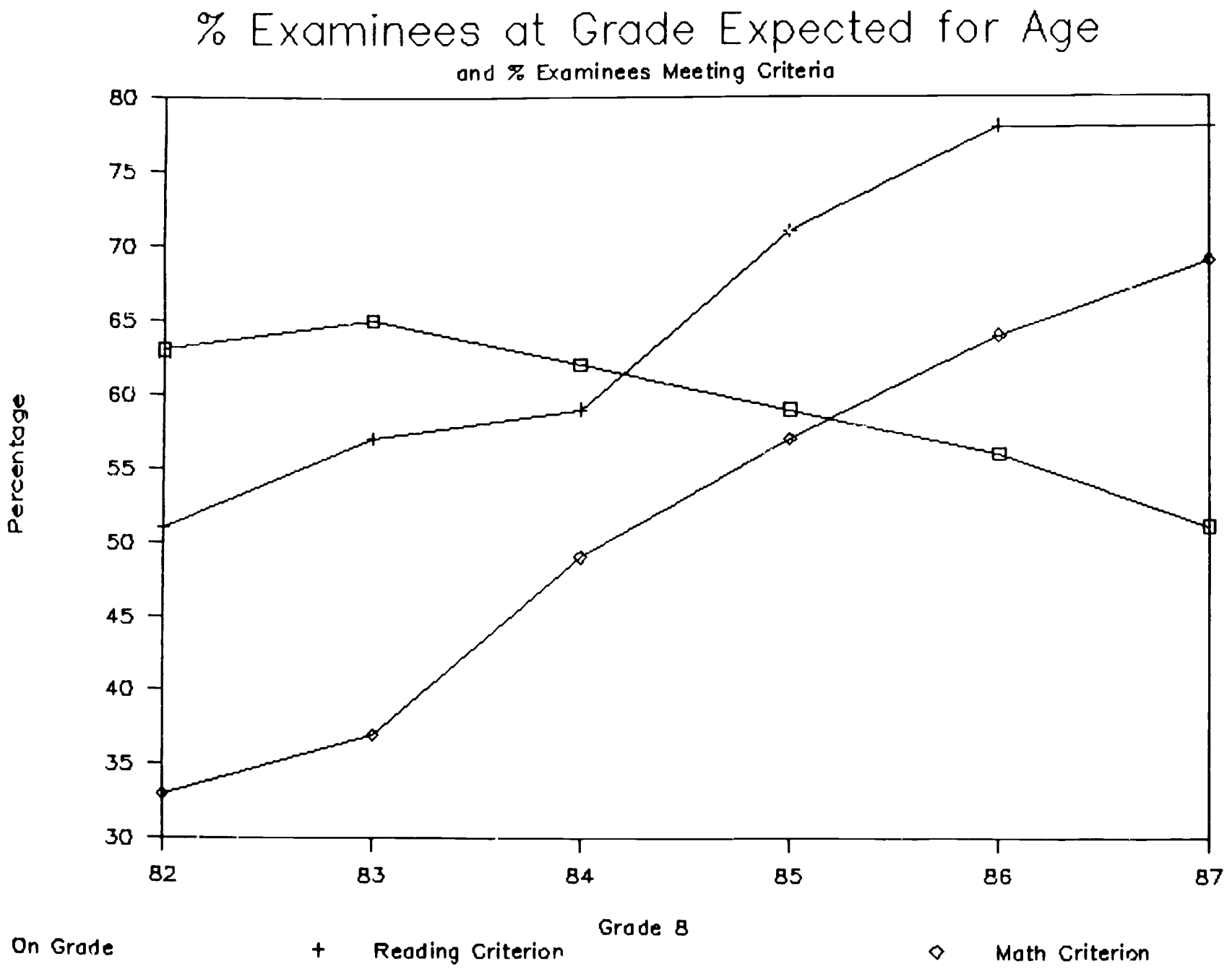
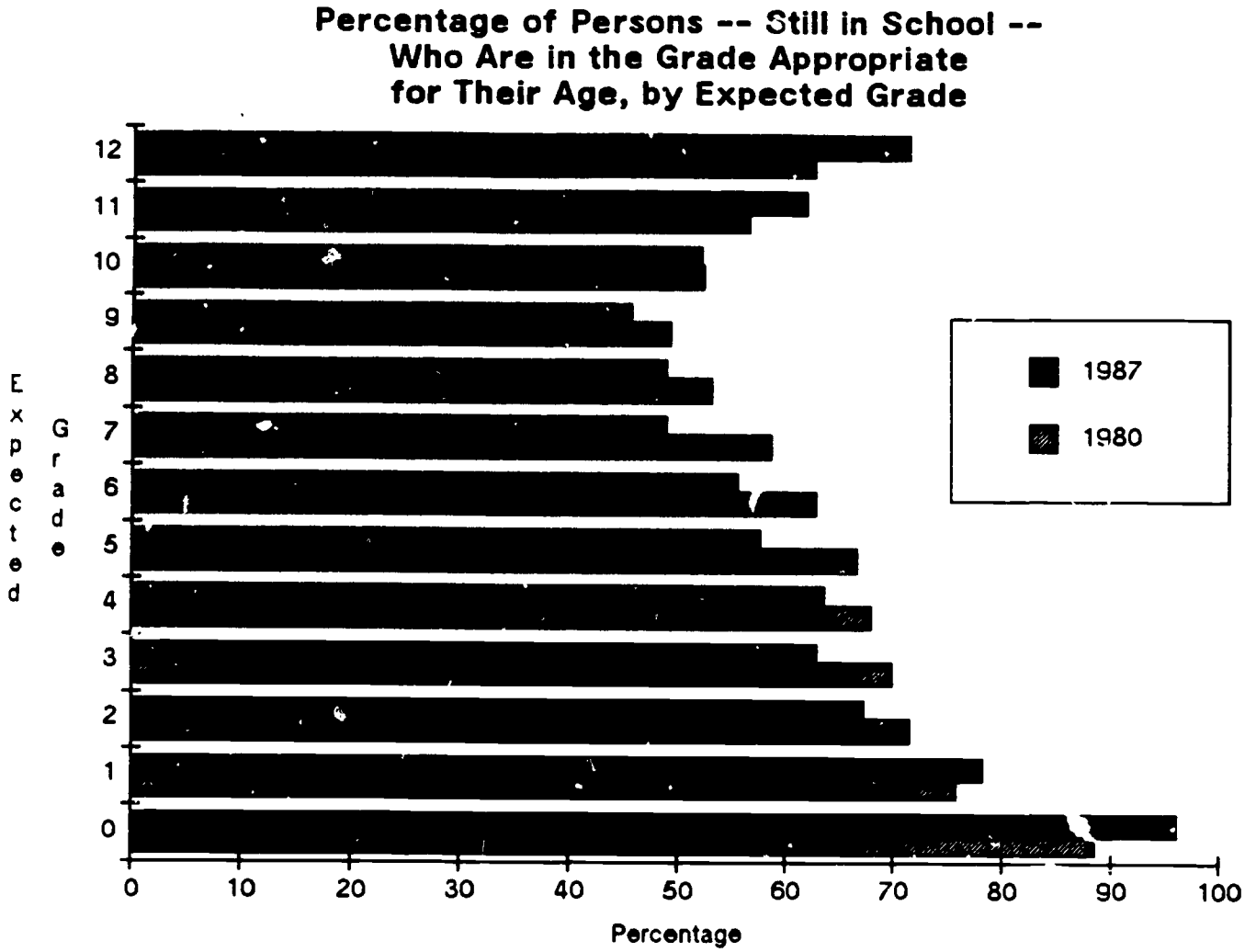

% Examinees at Grade Expected for Age
and % Examinees Meeting Criteria
80
75
70
65
60
55
50
45
40
35
30
Percentage
82
83
84
85
86
87
Grade 8
On Grade
+ Reading Criterion
◇ Math Criterion

Figure 11


Percentage of Persons -- Still in School --
Who Are in the Grade Appropriate
for Their Age, by Expected Grade
Expected Grade
12
11
10
9
8
7
6
5
4
3
2
1
0
1987
1980
0
10
20
30
40
50
60
70
80
90
100
Percentage