5-1994

Using Reading Skills as a Predictor of Success on the Fourth Grade S-PET Test

Francis J. Fiore Jr.
The College at Brockport

USING READING SKILLS AS A PREDICTOR OF SUCCESS ON THE FOURTH GRADE S-PET TEST

THESIS

Submitted to the Graduate Committee of the
Department of Education and Human Development
State University of New York
in Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Education

by
Francis J. Fiore Jr.
State University of New York
College of Brockport
Brockport, New York

May 1994

SUBMITTED BY:

Francis J. Fiore Jr. 5/2/94
Candidate Date

APPROVED BY:

Thesis Advisor Date

Second Faculty Reader 5/12/94
Date

Director of Graduate Studies 5/12/94
Date

Abstract

In the following study, the Program Evaluation Test in Science (S-PET) scores for students in the fourth grade at PS #44 within the Rochester City School District, were compared to their scores on the Degrees of Reading Power (DRP) test. The tests are administered to all fourth grade students every May.

Two groups of students were included: the first group was composed of those fourth grade students who took the S-PET test in May 1992. The second group was made up of the fourth grade students who took the S-PET test in May 1993. These same students also took the DRP test in May 1992 and May 1993, respectively.

This study examined the relationship between the DRP reading test scores and the S-PET science scores to see how strong a relationship there was between them. Further, it compared the level of achievement between the two groups on the S-PET test. The group who took the S-PET test in May 1993 had had extensive class time devoted to teaching specific vocabulary pertinent to the S-PET tests. If reading test scores on the DRP are an accurate predictor of success on the S-PET test, then one would expect the latter group to perform better on the S-PET test. Finally, the

S-PET was evaluated for reading difficulty using the Fry Readability Graph.

Results from a series of t tests indicated that the DRP scores correlated strongly with the S-PET test in both the 1991-92 test group and the 1992-93 test group. Further calculation of the point biserial coefficient of determination (rpb^2) indicated that class time devoted to teaching vocabulary relevant to the S-PET test did in fact have a significant effect upon the test scores of the 1992-93 test group. Finally, according to the Fry Readability Graph the S-PET has an estimated fourth grade level of reading difficulty.

Table of Contents

List of Tables

List of Figures

Chapter 1

Statement of Problem

Purpose

The purpose of this study was to examine the basic reading skills of fourth graders at PS44 in the Rochester City School District, to see whether their skill level in reading, measured on the DRP, could be used as a predictor of success on the S-PET test.

Questions to be Answered

1. How strong is the relationship between DRP scores and S-PET scores?
2. How much of the variation in the S-PET scores can be explained by the variation in the reading scores?
3. How much of the variation in the S-PET scores cannot be explained by the variation in the reading scores?
4. Is there a statistically significant difference between the mean scores of the 1992 group and the 1993 group which received extra assistance with scientific vocabulary throughout the school year in preparation for taking the S-PET?

5. What is the readability level of the S-PET test as estimated by the Fry Readability Graph?

Need for the Study

The Rochester City School District is an urban district that serves a diverse population of students, primarily composed of minority and low income children, as well as others from various ethnic, cultural and economic backgrounds. As a part of the school district's reform effort the district is encouraging teachers to provide a hands on approach to teaching science. It is hoped that this manipulative orientation toward the teaching of science will reinforce the body of knowledge students are exposed to in their science curriculum.

Lincoln Park School (PS44), a magnet school for Performing Arts within the Rochester City School District, wants all students to have the opportunity to be successful in their academic endeavors. However, at Lincoln Park School, and at other sites within the district, average scores on the S-PET test have been falling short of district expectations. A number of students not only do poorly on the multiple choice section of the test but also do poorly on the manipulative portion, even though they have done all the required coursework during the school year.

Why are these students not meeting district expectations? One reason might be that the reading level of the exam poses a problem. Students who lack grade appropriate reading skills might find reading the test questions difficult, even though they possess the necessary science knowledge. This also holds true for the manipulative portion of the test which requires the test taker to read and follow specific sets of instructions.

If it is found that reading scores can predict which students will have difficulty doing well on the S-PET test, a teacher might be able to help those students increase their chances of performing well on the test through the teaching of specific vocabulary and scientific terminology. Further because the reading level of many classes within the district continues to be lower and lower with each passing year, it is crucial that this issue be addressed through pertinent research so appropriate remedial action might be taken both by the classroom teacher and the district.

Definition of Terms

DRP - (Degrees of Reading Power) The Degrees of Reading Power test is a holistic measure of how well the messages within text are understood. As much as is possible in a test situation, the DRP test determines how well a student reads under "real life" conditions in and out of school.

They measure inferential comprehension, a necessary prerequisite to other higher order cognitive abilities. In other words the test is a single objective test measuring how well students understand the surface meaning of what they read.

Each DRP test consists of a number of nonfiction paragraphs and passages on a variety of topics. Each paragraph contains at least three sentences. Each passage contains about 325 words. Each paragraph contains a sentence with a blank space, and each passage has seven sentences that contain a blank space to indicate that a word is missing. For each bank four or five single word responses are provided. Students must select the most appropriate response to complete the sentence.

S-PET - (Program Evaluation Test in Science) The S-PET provides for an annual program evaluation in elementary science to measure the effectiveness of elementary science programs in the schools within New York State. This test measures local program implementation of the content, and skills contained in the New York State Elementary Science Syllabus. The test consists of two components, an objective test and a manipulative skills test.

The objective test contains a total of 45 multiple-choice questions. The manipulative skills test which is a timed test, includes tasks for students to

complete at each of five skills stations. Detailed instructions located at each of the five stations must be read by the student and followed in order to successfully complete each of the five stations.

Limitations of the Study

The data presented here were taken from a small number of students in just one elementary school within the Rochester City School District. The Rochester City School District is composed of a total of thirty-seven elementary schools, therefore, the students from this school may not be representative of all students taking the S-PET test within the district. Furthermore, the students from this study may not be representative of all students within New York State taking the S-PET test. Also, teaching strategies and instructional approaches may have been influencing factors. The students within this study are from a total of five different teachers' classrooms. Therefore, the results of this study may be more useful when considered in relation to PS44 and may not be equally applicable to other schools.

Chapter II

Review of the Literature

Purpose

The purpose of this study was to examine the basic reading skills of fourth graders at PS44, in the Rochester City School District, to see whether their skill level in reading, measured on the DRP, could be used as a predictor of success on the S-PET test.

Review of the Literature

The literature was read to examine the ways in which reading ability might affect student achievement on the fourth grade S-PET test. Many studies have been done that show the connection between reading skills and success in science courses (Corey, 1977; Daugs & Daugs, 1974; Doran & Sheard, 1974; Gallagher, 1979; Gardner, 1980; Thelan, 1976; Walker, 1980; Wright, 1982; Yore, 1987; and others).

Reading materials from science presents difficulties for many students (Thelan, 1976). In any science course there are a large number of scientific terms and concepts that must be mastered. These increase considerably in number and complexity as a student moves through his school

career. Hurd (1983) estimated the number of new vocabulary words introduced during a middle school science course to be about 2,500. Daugs and Daugs (1974) assessed the readability of textbooks used in high school science, and discovered that over half of the students were not able to read their textbooks well enough to utilize the material in them. In general, most high school science reading material is written at a level of reading difficulty two years beyond the ability level of the pupils using them (Corey, 1977; Walker, 1980; Wright, 1982). The inability to understand the written materials used in secondary science courses may frustrate students and hinder their learning (Doran & Sheard, 1974).

There are several reasons for students having difficulty with written science materials. Hurd (1983) discussed the immense vocabulary that must be learned and likened it to the learning of a foreign language. In regard to biology, Daugs and Daugs (1974) found that the biggest obstacle was not the vocabulary involved, but the structure of sentences in the readings. The sentences are too technical and also tend to contain too many words. Gardner (1980) examined the logical structure of the sentences in the texts. Gardner found that students have trouble with logical connectors in science writing, especially those that deal with negation or contrast.

Additionally, Walker (1980) found that the texts most frequently chosen by the teachers were not only difficult to read but were dry and boring.

One solution might be to have texts and related materials written in a simpler, less complex form. Corey (1977) suggested rewriting texts and materials in a simpler form that would approximate the level of reading ability of the students. Wright (1982) found that students comprehended the reading material better when rewritten in a simpler manner. Interestingly though, Wright discovered that students did not necessarily achieve higher scores when tested.

Many have questioned the readability of multiple-choice test questions used to evaluate a student's knowledge of science. Rakow and Gee (1987) have asserted that some tests are not really testing a student's knowledge of science but rather testing a student's reading ability, "Unless you have worded your test items so students are sure to understand what you are asking them, you may be challenging their reading ability rather than their grasp of scientific knowledge" (Rakow & Gee, 1987, p. 28).

The readability of tests, especially multiple choice tests is very important for larger-scale standardized tests such as the National Assessment of Educational Progress (NAEP) test. The NAEP science exam is administered

nationwide. It is a general science test. The NAEP is suppose to be readable by a student in the 8th grade. NAEP offers the same test items to 7th graders and 11th graders, even though one would expect the two groups to vary in reading ability. In 1982 the average score of the 11th grade students was 7.3% higher than the average score of the 7th grade students. A difference greater than 1.6% would be statistically significant. "However, lacking reliable readability measurements we can't know to what extent the results reflected the groups' difference in reading level rather than their science achievement" (Rakow & Gee, 1987, p. 29).

The readability of continuous text are not always easy to assess. Assessment of multiple choice test items is even more difficult. Most readability formulas are designed to analyze several passages of continuous text containing a minimum of one-hundred words each. Readability formulas requiring several passages of one hundred words or more as a sample are: Dale-Chall, Fog, Fry, and Smog (Kahle, 1979).

A multiple choice test item is organized in a pattern that is different from a continuous text. Because multiple choice questions are not long enough to provide continuous passages, readability formulas must be used with caution. The imprecision of readability formulas becomes obvious when

various formulas are applied to the same test item. There were five different formulas applied to one biology item on the NAEP science exam. For this one item, readability estimates varied by as much as 5.8 years (Figure 4, in Appendix B). The estimates ranged from a high of 13.8 years using the Fog Index, to a low of grade 9 using the Fry test.

One researcher recommended the use of the Fry Readability Graph (Figure 3, in Appendix B) because it is simple to utilize and according to her reliable (Kahle, 1979). Kahle found that given a passage of text, the results from various readability tests were generally consistent when applied to science tests. Kennedy (1979) compared the Fry method to several other readability formulas and decided it was useful for assessing secondary school science text books.

It seems obvious that adequate reading ability is of great importance in relation to scholastic achievement in science. Many factors affect a student's problem with reading in science, but it is the influence of reading ability, that impacts upon achievement in science. Daune (1980) found reading achievement to be the most important predictor of student success, followed by math ability.

Other research also points to reading as a primary predictor of success in science courses. Detloff (1982)

gave incoming college freshmen a variety of tests in order to predict which students would not succeed in college science courses. He found reading, followed by math, to be the most important predictor of achievement. Cox (1982) also discovered that math and reading skills were the greatest components to student achievement in science. A study by Yore (1985) of fifth grade students found a strong correlation between reading vocabulary and science achievement.

According to Young, Ruck, and Crocker (1991), students who have difficulty reading science text books need to be taught how to do so. There are special reading skills that are needed in the content area of science. Specific comprehension abilities basic to the reading and understanding of science are noting details, reading analytically, interpretation of charts and diagrams, following directions, visualizing what is read, identification of cause and effect, classification of ideas, and recognition of the steps involved in an experiment (Horn, 1955; Siemons, 1956; Sochor, 1953).

In a study by Howe, Hall, Stanback, and Seidman (1983) of a racially diverse, urban group of ninth grade science students, they found that "at this level, the obstacles to achievement in science are not related to race or to sex but, rather, to poor math and reading skills, and failure to participate in the learning activities of the classroom"

(p. 102). They also found that the best predictors of success in science were reading and math ability. They used the California Achievement Tests to measure math and reading ability.

Summary

It would seem evident, based on the findings of numerous investigators, that reading skills are a prerequisite if a student is to achieve success in science. Science textbooks and related materials are filled with a tremendous amount of vocabulary which a student must learn if she/he is to fully understand the various concepts that are being taught. Further, most textbooks are not written in a user friendly format. The structure of textbook sentences tends to be too technical and to contain too many words. There is also the logical connectors in science writing, especially those that deal with negation or contrast that give students difficulty when reading. It also seems plausible to believe that because of the readability of some test questions in science that one is really being tested on reading ability rather than on one's knowledge of science. Several researchers recommend that science textbooks and related materials be rewritten in a simpler form that would approximate the level of reading ability of the student. In any event, it appears evident that reading ability plays an important role in a student's success in science.

Chapter III

The Research Design

Purpose

The purpose of this study was to examine the basic reading skills of fourth graders at PS44 in the Rochester City School District, to see whether they skill level in reading, measured on the DRP, could be used as a predictor of success on the S-PET test.

Hypotheses

It is expected that there will be a strong relationship between DRP scores and S-PET scores. Further, it is expected that there will be a statistically significant difference between the mean scores of the 1992 group and the 1993 group which received extra assistance with scientific vocabulary throughout the school year in preparation for taking the S-PET.

Methodology

Subjects

The test groups were drawn from students in the fourth grade at PS44 in the Rochester City School District. It did not include those pupils enrolled in special education

programs. The students represented in the study came from the classrooms of five different teachers. The scores used were from the DRP test and the S-PET. Data were collected from the May 1992 and May 1993 tests.

The 1992 group was composed of seventy-five fourth grade students and the 1993 group was composed of sixty-one students. Only test scores were used for those students with a complete set of scores: if the student was absent for any one test, those scores were eliminated. All students included as a part of the sample took level F-7 of the DRP test in May of the same year as the S-PET test. Both S-PET and DRP test data were collected from the Management Information Services department of the Rochester City School District's central administrative offices.

Instruments

The Degrees of Reading Power (DRP) tests are standardized norm-referenced tests of reading comprehension. The tests are a holistic measure of how well the messages within text are understood. They measure inferential comprehension, a necessary prerequisite to other higher order cognitive abilities. In other words the test is a single objective test measuring how well students understand the surface meaning of what they read.

The DRP test consists of a number of nonfiction paragraphs and passages on a variety of topics. Each

paragraph contains about 325 words. Each paragraph contains a sentence with a blank space and each passage has seven sentences that contain a blank space to indicate that a word is missing. For each blank four or five single word responses are provided. Students must select the most appropriate response to complete the sentence.

The raw score on the DRP test is the number of questions answered correctly. Possible raw scores range from 0 to 70. The raw scores are converted into DRP percentiles.

The Program Evaluation Test in Science (S-PET) is composed of two required components, an objective test and a manipulative skills test. The objective test (Form F), contains a total of 45 multiple-choice questions divided into two sections: content (questions 1-29) and skills (questions 30-45). Each section is scored separately. The student's score on each section is the number of questions within that section that the student answered correctly. A raw score for the objective portion of the test can be obtained by adding the number of correct answers and ranges from 0 to 45.

The second component of the S-PET is the manipulative skills test. It consists of tasks set up at a series of five different stations. A station is a student's desk. Examples of student tasks might be to determine: the

height of a glass, the mass of that glass, the temperature of the water in the glass and the volume of water needed to fill the glass to a reference line. A raw score for the manipulative portion of the test can be obtained by adding scores achieved at each station. Raw scores range from a low of 0 to a high of 40. Total raw scores on the S-PET range from 0 to 85.

The readability of the S-PET was estimated using the Fry readability graph which examines the average number of syllables and the average number of sentences per one-hundred words. The text used comes from the objective portion of the test.

Procedures

The study examined the relationship between reading ability and performance on the S-PET to see how strong a relationship exists between them. The test groups were drawn from students in the fourth grade at PS44 in the Rochester City School District. The scores used were from the DRP and the S-PET tests, which are administered each year during the month of May. Test data were collected for both 1992 and 1993 test years.

Additionally, scores from the 1992 and 1993 S-PET test were compared to see if there was a statistically significant difference between them. The basis for this comparison was the fact that the 1993 test group received

extra time devoted to learning vocabulary relevant to the S-PET. A t test was used to compare the mean test scores of the two test groups.

Summary

Basic reading skills of fourth graders at PS44 in the Rochester City School District were examined to see whether or not their skill level in reading, measured on the DRP, could be used as a predictor of success on the S-PET test.

Two groups were examined. One group was composed of those students who were fourth graders during the 1991-1992 school year. The other group were those students who were fourth graders during the 1992-1993 school year. Additionally, the 1991-1992 group received extra assistance with scientific vocabulary throughout the school year prior to their taking the S-PET test.

Chapter IV

Analysis of Data

Purpose

The purpose of this study was to examine the basic reading skills of fourth graders at PS44 in the Rochester City School District, to see whether their skill level in reading, measured on the DRP, could be used as a predictor of success on the S-PET test.

Analysis of the Data

1991-1992 Data

Summary Statistics For the test group composed of the May 1992 S-PET students, the mean on the S-PET test was 45.02 with an average variation around the mean of 9.88. The test scores ranged from a low of 14 to a high of 61. The mean on the DRP test was 46.13 with an average deviation around the mean of 21.74. The test scores ranged from .02 to 189 (see Tables 1 and 2, in Appendix A). The DRP scores correlate strongly with the S-PET scores with a correlation coefficient (r) of .67.

When DRP scores are employed as the independent variable, and S-PET scores as the dependent variable, and

the sample data pairs are plotted, the regression line that best fits the data has the equation Y = 31 + .31 * X where the Y intercept = 31 and the slope of the line is .31. The plot of the residuals for this data does not reveal any obvious trend, so a straight line is assumed to be the best description of the data (see Figure 1, in Appendix B).

The relationship between the DRP test scores and the S-PET scores is strong, as indicated by an r^2 value of .46. This value indicates that 46% of the variation in the S-PET scores can be explained by the variation in the reading scores. It also indicates that 54% of the variation in the S-PET scores cannot be explained by the variation in the reading scores.

1992-1993 Data

Summary Statistics for the test group composed of the May 1993 S-PET students, the mean on the S-PET test was 59.93 with an average variation around the mean of 10.59. The test scores ranged from a low of 28 to a high of 78. The mean on the DRP test was 53.63 with an average variation around the mean of 21.64. The test scores ranged from a low of .07 to a high of .92 (see Tables 4 and 5, in Appendix A).

The DRP scores correlate strongly with the S-PET scores with a correlation coefficient (r) of .59.

When DRP scores are employed as the independent variable, and S-PET scores as the dependent variable and

the sample data pairs are plotted, the regression line that best fits the data has the equation Y = 44 + .29 * X where the Y intercept = 44 and the slope of the line is .29. The plot of the residuals for this data does not reveal any obvious trend, so a straight line is assumed to be the best description of the data (see Figure 2, in Appendix B).

The relationship between the DRP test scores and the S-PET scores is strong, as indicated by an r^2 value of .35. This value indicates that 35% of the variation in the S-PET scores can be explained by the variation in the reading scores. It also indicates that 65% of the variation in the S-PET scores cannot be explained by the variation in the reading scores.

Comparison of 1991-1992 and 1992-1993 Data

It was hypothesized that there would be a statistically significant difference between the mean 1991-1992 S-PET scores and the mean 1992-1993 S-PET scores. It was further hypothesized that there would be no statistically significant difference between the mean 1991-1992 DRP scores and the mean 1992-1993 DRP scores.

A t test for the difference between the two S-PET means was used to compare the 1991-1992 mean score and the 1992-1993 mean score. A calculated t score of -8.472 was obtained. Since the critical value of t for 120 degrees of freedom, unbiased, at the 99% confidence level is $\pm$ 2.576,

and since the t obtained was -8.472, the null hypothesis is rejected and it is concluded there was a statistically significant difference between the 1991-92 mean test score and the 1992-93 mean test score on the S-PET test (see Table 5-1, in Appendix A).

A t test for the difference between the two DRP test means was used to compare the 1991-92 mean score and the 1992-93 mean score. A calculated t score of -2.006 was obtained. Since the critical value of t for 120 degrees of freedom, unbiased, at the 99% confidence level is ± 2.576, and since the t obtained was -2.006, the null hypothesis is retained and it is concluded that there was no statistically significant difference between the 1991-92 mean test score and the 1992-93 mean test score on the DRP (see Table 52, in Appendix A).

As a final step, the point biserial coefficient of determination (rpb^2) was calculated to see just how much of the variation in the S-PET scores between the 1991-92 group and the 1992-93 group was attributable to additional class time devoted to teaching vocabulary relevant to the S-PET test, referred to as the treatment effect. An rpb^2 (treatment effect) of .35 was obtained. What this says is that the treatment effect by itself is explaining 35% of the variation in the S-PET scores between the 1991-92 and 1992-93 test groups.

Fry Readability Test

The text used to estimate the readability of the S-PET was taken from both the multiple choice and manipulative portions of the test. The identical test was used for both the 1991-92 test group and the 1992-93 test group. Both portions of the test have an estimated 4th grade level of reading difficulty according to the Fry Readability Graph (Appendix C). What this says is that for the average fourth grader this test does not pose a high level of reading difficulty. But for those students who are not reading on or near grade level the test would present reading difficulties (see Figure 3, in Appendix B).

Summary

From the data collected in this study it was concluded that DRP scores correlate strongly with S-PET scores and that reading ability is a good predictor of success on the S-PET test. Further, it was determined that there was no statistically significant difference in test scores, between the 1991-92 and 1992-93 DRP test groups. However, there was a statistically significant difference in test scores between the 1991-92 and 1992-93 S-PET groups. Thirty-five percent of that difference was attributable to the treatment effect, the treatment being additional class time devoted to teaching vocabulary relevant to the S-PET

test. It was also concluded, based on the Fry Readability method, that the S-PET has a fourth grade level of reading difficulty.

Chapter V

Conclusions and Implications

Purpose

The purpose of this study was to examine the basic reading skills of fourth graders at PS44 in the Rochester City School District, to see whether their skill level in reading, measured on the DRP could be used as a predictor of success on the S-PET test.

The scores of students on the 1992 S-PET test were compared to the scores of the same students on the DRP test also administered in 1992. The same measurements were repeated in the following year on those students taking both the S-PET and DRP tests at PS44.

Conclusions

In examining the 1991-92 S-PET scores with the DRP test scores it was found that the S-PET scores correlated strongly with the DRP scores, with a correlation coefficient (r) of .67. The relationship between the S-PET and DRP test scores was strong, as indicated by a coefficient of determination (r2) value of 46%. This value indicates that 46% of the variation in the S-PET scores can

be explained by variation in the reading score. Similar results were obtained using the 1992-93 S-PET and DRP test data. The S-PET scores correlated strongly with the DRP test scores with a correlation coefficient (r) of .59. The relationship between the S-PET and DRP test scores was strong, as indicated by a coefficient of determination (r2) value of .35. This value indicates that 35% of the variation in the S-PET score can be explained by variation in the reading score. In conclusion, reading ability is a strong predictor of success on the S-PET test.

In comparing the 1991-92 DRP test mean score with the 1992-93 DRP test mean scores no statistically significant difference was found and the null hypothesis was retained. A $\underline{t}$ test for the difference between the two DRP means was used. A calculated $\underline{t}$ score of -2.006 was obtained. Since the critical value of t for 120 degrees of freedom, unbiased, at the 99% confidence level is $\pm$ 2.576, and since the $\underline{t}$ obtained was -2.006 the null hypothesis was retained. Although there was a difference between the two test groups in reading ability, that difference was statistically insignificant, and could not be used to explain the difference in the performance on the S-PET test.

In comparing the 1991-92 S-PET mean score with the 1992-93 S-PET mean score a statistically significant

difference was found and the null hypothesis was rejected. A $\underline{t}$ test for the difference between the two S-PET means was used. A calculated $\underline{t}$ score of -8.472 was obtained. Since the critical value of $\underline{t}$ for 120 degrees of freedom, unbiased, at the 99% confidence level is $\pm$ 2.576 and since the $\underline{t}$ obtained was -8.472 the null hypothesis was rejected. Therefore, there was a statistically significant difference between the two test groups. The 1992-93 test group had out performed the 1991-92 group and the reason could not be attributed to a significant difference in the reading ability between the two groups. A statistically significant difference in reading ability, based on DRP test scores, between the two test groups had already been dismissed based on the $\underline{t}$ test performed for that purpose.

As a final step, the point biserial coefficient of determination (rpb^2) was calculated to see just how much of the variation in the S-PET scores between the 1991-92 group and the 1992-93 group was attributable to the additional class time devoted to teaching vocabulary relevant to the S-PET test, referred to as the treatment effect. An rpb^2 (treatment effect) of .35 was obtained. What this says is that the treatment effect by itself is explaining 35% of the variation in the S-PET scores between the 1991-92 and 1992-93 test groups. It appears based on these data, that spending time on vocabulary related to

science does in fact improve one's chances of doing well on the S-PET test.

According to the Fry Readability test, the fourth grade S-PET test had an estimated fourth grade level of reading difficulty. What this says is that for the average fourth grader this test does not pose a high level of reading difficulty. But, for those students who are not reading on or near grade level the test would present reading difficulties. At PS44 many of our fourth graders are reading well below grade level, some as low as first grade. Therefore, the S-PET for many, would present a great amount of reading difficulty. In essence, many children may be performing poorly on the S-PET test not because of a lack of science content knowledge but because of their inability to read and interpret exactly what the test questions are asking. It would seem that in many cases we are not testing students' knowledge of science but rather testing their reading ability.

An examination of a student's reading scores will give a good indication of how well a student is likely to perform on the S-PET test. Children with extremely low reading scores would benefit greatly from intensive science vocabulary enrichment throughout the school year. All students would probably benefit. From learning strategies specifically geared toward the reading of scientific materials.

Researchers have suggested ways to help students with reading skills in science. Dettloff recommends remedial work in reading, math, and study skills within the context of a science course. Thelan (1976) outlines techniques for reading practice in the content areas. Yore and Shymansky (1985) focus on the use of science text books in connection with other classroom work to provide reinforcement of science materials. Young, Ruck, and Crocker (1991), feel that in order to be really helpful teachers must explain to students how science books use language. Two ways of accomplishing this are by reading to students, and developing in students an awareness of the text's structure accomplished by asking questions about reading strategies.

Looking at the results of the Fry analysis from the S-PET test it would seem clear that for many students, if for not most, the test is beyond their reading capabilities. Therefore, all students should benefit from learning and practicing reading techniques that will aid them in understanding the S-PET test.

One tool in helping students improve their scientific performance is the simplification of science textbooks. Kulkarni (1987) decided to simplify the textbooks in use in his project, in order to make them more approachable to less-skilled readers. What he did was to change the style of the written text. He removed double negatives, passive

sentence constructs, and complex sentence structures. The results of this were improved student performance. "However, the most significant finding was the removal of disparity in the performance of students from different socioeconomic backgrounds" (Kulkarni, 1987, p. 164).

Yet another means of improving science test performance is through the simplification of the science test itself. Rakow and Gee (1987) and Macinnis (1979) propose the rewriting of tests to minimize the barrier reading poses to students who have a grasp of the scientific material, but lack the necessary reading skills to conquer the test. They suggest using shorter sentences that contain short, familiar words, and active rather than passive verbs.

In a report from the National Research Council, (Fulfilling the Promise: Science Education in the Nation's Schools, 1990) the Council examines current high school science courses and makes suggestions for their improvement. They promote the mastery of concepts through hands on inquiry, not merely through the memorization of words. They also suggest that the number of new vocabulary words must be kept to a minimum. "Students are discouraged by the overwhelming amount of material and the relentless onslaught of technical vocabulary" (National Research Council, p. 12).

Science makes certain unique demands on children's

reading abilities. Among the difficulties in science materials are a highly specialized vocabulary and concepts for which written explanations are often difficult and abstract. Many teachers seem to assume that reading abilities practiced during the reading period develop all the understandings and skills needed by children to read effectively in the content areas. But this is not the case. To guarantee student success on the S-PET test, teachers need to focus on vocabulary skills, reading comprehension, and reading interpretation.

Implications for Research

Further investigations into the role reading ability plays in the academic success of students are suggested. Research into the following areas is needed:

1. A study focussing on how reading ability affects success in other content area subjects such as social studies and mathematics.

2. A study to determine if standardized tests are in fact testing content area knowledge as opposed to reading ability.

3. A study to ascertain whether math ability might be a good predictor of success in science course work.

4. Investigation into whether or not self-esteem hinders performance in science. Poor self-esteem fosters a

lack of risk taking which is so much a part of the trial and error method of science.

For further study, the following changes in the experimental design are recommended:

1. A larger sample size that would involve students from more than just one school building.

2. A more heterogeneous sampling. Rather than choosing students from just an urban school district, involve suburban students in the research.

3. Increase the number of test years that are examined from two to possibly five.

Suggestions for Classroom Practice

Classroom practices should foster and strengthen the following abilities which are characteristic of scientific endeavors.

1. Note details
2. Read analytically
3. Interpret charts, tables, diagrams, etc.
4. Perform specific problems
5. Follow directions
6. Recognize the steps in an experiment
7. Visualize what is read
8. Identify cause and effect relationships
9. Classify ideas

Suggested Classroom Practices

1. Read passages aloud to the class so students develop a feeling of familiarity for the way language is used in a science book.

2. Ask students reading strategy questions which allows a student to explain how they would arrive at an answer or what information they would use from the book to answer the question. The student does not give the actual answer.

3. Practice reading and vocabulary skills within the context of the fourth grade science course.

4. Use textbooks and supplemental material that is simply as well as clearly written.

5. Insist that students consult a variety of materials involving problems rather than relying on a single textbook. This reinforces the skills of locating, selecting, evaluating, and organizing information.

6. Where at all possible allow children to form cooperative learning groups based upon their common interests in science.

7. Select materials which are simpler than the current basic reader the child is reading.

8. Use the words to be learned in the discussions children have for planning and sharing and begin to establish a word list.

9. Present the new words in sentences in which the context will give clues to their meaning.

10. Explain terms to the class according to their level of understanding.

11. Develop pupil made materials such as charts and science picture dictionaries.

12. Provide direct help with word recognition skills: picture clues, context clues, word form clues, structural analysis and phonetic analysis.

13. Help students develop a clear understanding of what is expected of them in the different kinds of study activities. This will allow them not only to acquire skill in dealing with specific situations, but will allow them to adjust their reading and study procedures intelligently for many different purposes.

Summary

It has been shown through this study that there exists a strong correlation between reading ability and performance on the S-PET test. Further, it was shown that additional class time devoted to teaching vocabulary relevant to the S-PET test, with the 1992-93 group significantly increased their performance on that test.

With this knowledge it is imperative for classroom teachers to prepare their students for taking the S-PET

through the practicing of reading and vocabulary skills within the context of the fourth grade science curriculum. In this way students will truly be tested on their knowledge of science as opposed to their reading ability.

References

Corey, N. R. (1977). The use of rewritten science materials in ninth grade biology. Journal of Research in Science Teaching, 14, 97-103.

Cox, D. C. (1982). The effect of type of classroom science, grade level, years without science instruction, and elective science courses on performance level for selected high school process science skill competencies. Unpublished doctoral dissertation, Ohio State University, Columbus, OH.

Daugs, D. R. & Daugs, F. (1974). Readability of high school biology materials. Science Education, 58, 471-482.

Daume, R. J. (1980). The effect of type of junior high school science program on science content achievement and science process Achievement. Paper presented at the Meeting of the Mid-South Educational Association, MS.

Dettloff, J. M. Predicting achievement in community collect science students. Unpublished doctoral dissertation, University of Michigan.

Doran, R. L., & Sheard, D. M. (1974). Analyzing science textbooks. School Science and Mathematics, 74, 31-39.

Fry, E. (1977). Fry's readability graph: clarifications, validity, and extension to level 17. Journal of Reading, 21, 242-252.

Gallagher, J. J. (1979). Basic skills common to science and mathematics. School Science and Mathematics, 79, 555-565.

Gardner, P. L. (1980). The identification of specific difficulties with logical connectives in science among secondary school students. Journal of Research in Science Teaching, 17, 223-229.

Horn, E. (1955). Reading Skills Needed in Content Areas. Reading Teacher, 8(4), 212.

Howe, A. C., Hall, V., Stanback, B., & Seideman, S. (1983). Pupil behaviors and interactions in desegregated urban junior high activity-centered science classroom. Journal of Educational Psychology, 75, 97-103.

Hurd, P. D. (1983). Middle school/junior high science: changing perspectives. In M. J. Padilla (Ed.), Science and the early adolescent. Washington, DC: National Science Teachers Association.

Ivens, S. H. (1989). The Degrees of Reading Power Test. Brewster, New York: TASA, Inc.

Kahle, J. B. (1979). Teaching science in the secondary school. New York: Van Nostrand.

Kennedy, K. (1979). Determining the readability of biology textbooks. American Biology Teacher, 41, 301-303.

Koslin, B. L., Zeno, S., & Koslin, S. (1987). The DRP: An Effectiveness Measure in Reading. Touchstone Applied Science Associates, Inc.

Kulkarni, V. G., (1988). Role of language in science education. In P. Fensham, (Ed.), Development and dilemmas in science education (p. 164). London: Falmer Press.

Lawson, A. E., Nordlund, F. H., & Kahle, J. B. (1975). Levels of intellectual development and reading ability in disadvantaged students and the teaching of science. Science Education, 59, 113-125.

Lawson, A. E., & Renner, J. W. (1975). Relationships of science subject matter and developmental levels of learners. Journal of Research in Science Teaching, 75, 347-358.

Macinnis, P. (1979). Science library of test items, volume 5: readability and science testing. Sydney, Australia: New South Wales Department of Education.

Marke, E. A. (1981). Correlations among cognitive development, intelligence quotient, and achievement of high school biology students. Journal of Education in Science Research, 18, 9-14.

Mitchell, A., & Lawson, A. E. (1988). Predicting genetics achievement in nonmajors' college biology. Journal of Research in Science Teaching, 25, 23-27.

National Research Council. (1990). Fulfilling the promise: Biology education in the nation's schools. Washington, DC: National Academic Press.

New York State Education Department. (1993). New York State Program Evaluation Test In Science. Albany, NY: The University of the State of New York.

Rakow, S. J., & Gee, T. C. (1987). Test science, not reading. The Science Teacher, 54(2), 28-31.

Rochester City School District. Management Information Systems Department (1993). Student test data.

Siemons, A. (1956). Guiding Reading in Science. Reading Teacher, 5, 97-100.

Sochor, E. E. (1953). Special Reading Skills Are Needed In Science. Reading Teacher, 6(4), 4-10.

Thelan, J. N. (1984). Improving reading in science. Newark, DE: International Reading Association.

Walker, N. (1980). Readability of college general biology textbooks: revisited. Science Education, 64, 29-34.

Wright, J. D. (1982). The effect of reduced readability text materials on comprehension and biology achievement. Science Education, 66, 3-13.

Yore, L. D. (1987, April). A preliminary exploration of grade five students' science achievement and ability to read science textbooks as a function of gender, reading vocabulary, and reading comprehension. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Washington, DC.

Yore, L. D., & Denning, D. (1989, April). Implementing change in secondary science reading and textbook usage: a desired image, a current profile, and a plan for change. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Washington, DC.

Yore, L. D., & Shymansky, J. A. (1985, April). Reading, understanding, remembering, and using information in written science materials. Paper presented at the Annual Meeting of the Association for the Education of Teachers in Science, Cincinnati, OH.

Young, P., Ruck, C., & Crocker, B. (1991). Reading, its not quite Jack in the Beanstalk. Science Teacher, 58(2), 46-49.

TABLE 1

STUDENT TEST DATA (1991-1992)

Student	Total PET Raw Score	DRP Reading Percentile
1	29	15
2	52	71
3	49	47
4	44	58
5	42	44
6	53	31
7	45	37
8	28	31
9	14	14
10	53	54
11	45	37
12	56	71
13	43	37
14	46	37
15	34	37
16	30	7
17	54	44
18	58	54
19	45	44
20	39	14
21	35	44
22	47	31
23	47	61
24	53	74
25	60	89
26	48	40
27	29	12
28	58	76
29	41	28
30	61	40
31	55	74
32	55	76
33	49	47
34	54	51
35	57	88
36	46	58
37	57	89
38	49	74
39	40	44
40	44	37
41	38	54
42	46	74

TABLE 1 (Cont'd)

STUDENT TEST DATA (1991-1992)

Student	Total PET Raw Score	DRP Reading Percentile
43	52	71
44	29	40
45	57	61
46	39	28
47	52	51
48	38	37
49	25	9
50	30	7
51	49	58
52	42	18
53	43	37
54	42	25
55	38	15
56	55	76
57	56	86
58	39	47
59	40	40
60	27	47
61	53	47
62	40	31
63	34	74
64	43	54
65	52	31
66	54	47
67	40	28
68	49	54
69	51	58
70	42	25
71	54	65
72	51	81
73	26	2
74	50	28
75	57	37

TABLE 2

SUMMARY STATISTICS (1991-1992)

DRP Summary Statistics

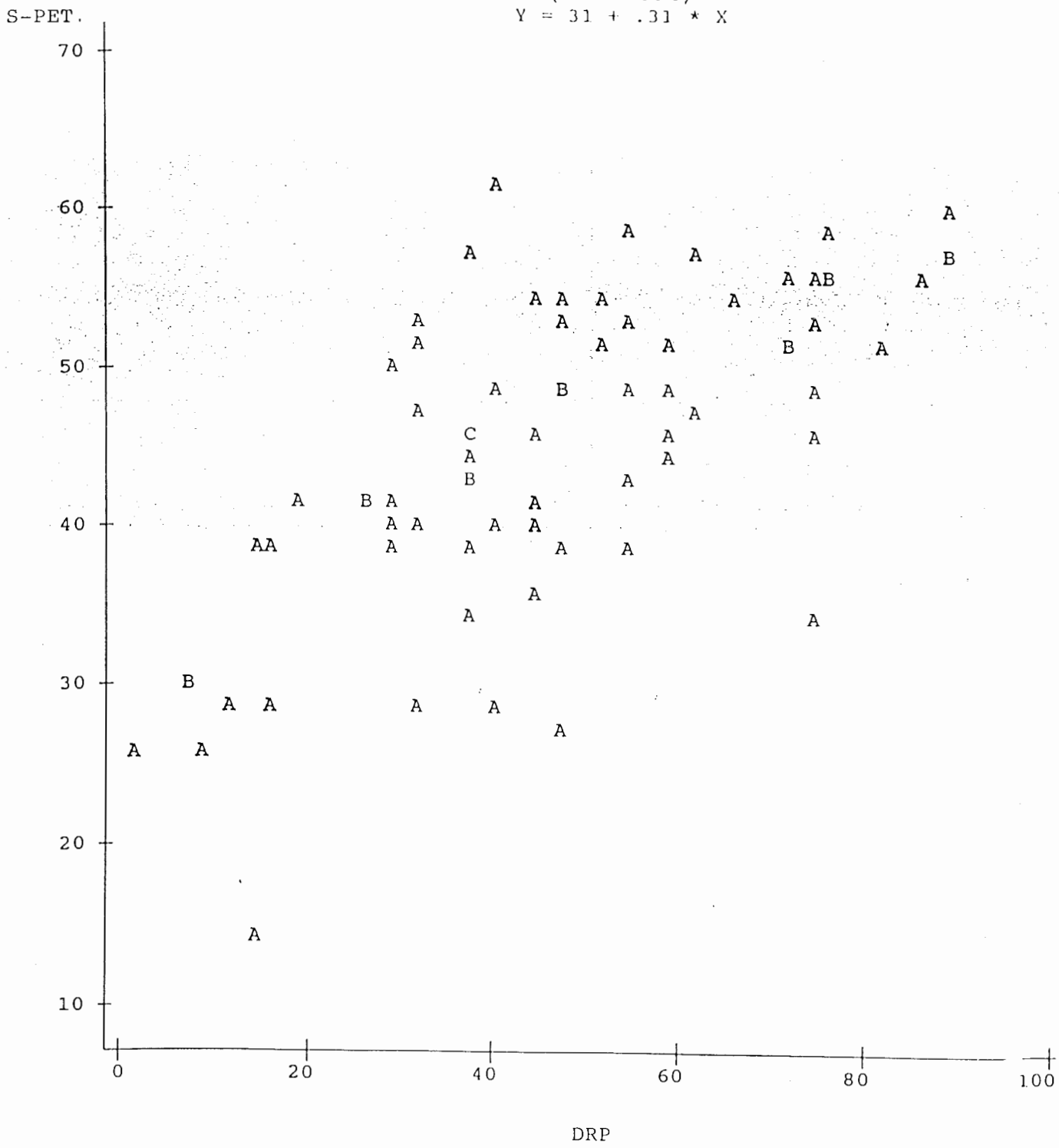
N = 75
Mean = 46.13
Standard Deviation = 21.75
Standard Error = 2.51
Minimum = 2
Maximum = 89
Skewness = 0.10787

S-PET Summary Statistics

N = 75
Mean = 45.03
Standard Deviation = 9.88
Standard Error = 1.14
Minimum = 14
Maximum = 61
Skewness = 0.7100

FIGURE 1

RELATIONSHIP BETWEEN DRP TEST SCORES AND S-PET SCORES
(1991-1992)
Y = 31 + .31 * X



Legend: A = 1 student, B = 2 students, etc.

TABLE 3

STUDENT TEST DATA (1992-1993)

Student	Total PET Raw Score	DRP Reading Percentile
1	44	44
2	42	23
3	37	71
4	71	34
5	58	18
6	40	28
7	48	47
8	61	61
9	69	68
10	70	68
11	57	28
12	71	76
13	67	74
14	61	74
15	51	18
16	55	88
17	56	28
18	51	37
19	60	40
20	70	54
21	47	25
22	57	40
23	61	61
24	59	34
25	78	81
26	53	37
27	70	76
28	49	7
29	74	47
30	60	47
31	50	20
32	72	61
33	70	92
34	65	79
35	75	84
36	59	58
37	73	76
38	60	76
39	63	37
40	67	61
41	52	40
42	59	37

TABLE 1 (Cont'd)

STUDENT TEST DATA (1991-1992)

Student	Total PET Raw Score	DRP Reading Percentile
43	57	61
44	74	74
45	74	71
46	66	51
47	75	91
48	56	25
49	28	25
50	56	65
51	63	37
52	48	31
53	61	71
54	55	61
55	71	68
56	60	54
57	48	65
58	50	51
59	69	71
60	65	54
61	68	91

TABLE 4

SUMMARY STATISTICS (1991-1992)

DRP Summary Statistics

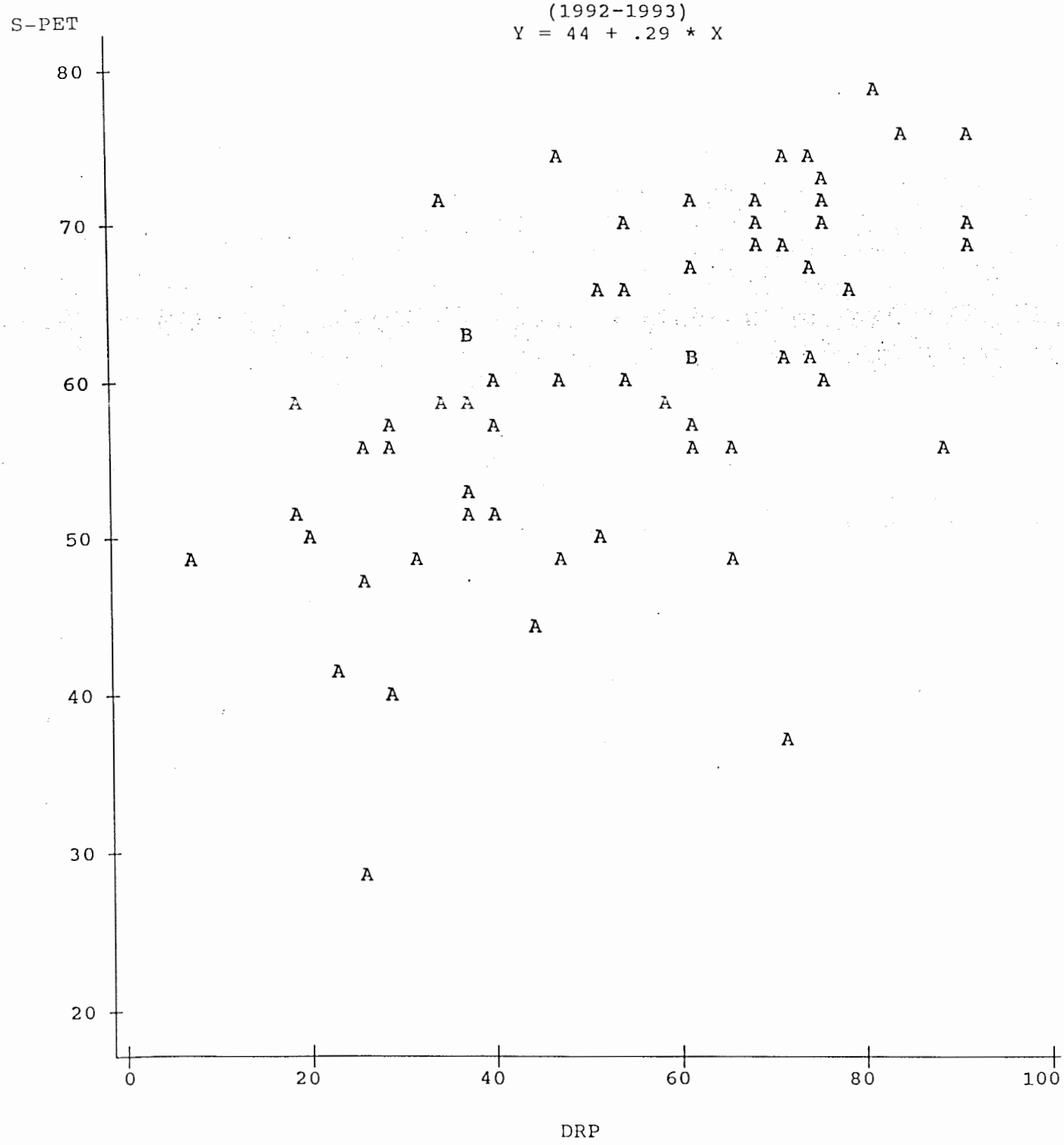
N = 61
Mean = 53.64
Standard Deviation = 21.65
Standard Error = 2.77
Minimum = 2
Maximum = 92
Skewness = 0.11503

S-PET Summary Statistics

N = 61
Mean = 59.93
Standard Deviation = 10.59
Standard Error = 1.36
Minimum = 28
Maximum = 78
Skewness = 0.56819

FIGURE 2

RELATIONSHIP BETWEEN DRP TEST SCORES AND S-PET SCORES
(1992-1993)
Y = 44 + .29 * X



Legend: A = 1 student, B = 2 students, etc.

TABLE 5
t TEST PROCEDURES

1. t Test of the Difference Between 1991-92 S-PET Scores and 1992-93 S-PET Scores.

Year	DF	X	Standard Deviation	t
91-92	120	45.03	9.88	-8.472
92-93	120	59.93	10.59	-8.472

* Critical t = ± 2.576 (at the 99% confidence level)

2. t Test of the Difference Between 1991-92 DRP Test Scores and 1992-93 DRP Test Scores.

Year	DF	X	Standard Deviation	t
91-92	120	46.13	21.75	-2.006
92-93	120	53.64	21.65	-2.006

* Critical t = ± 2.576 (at the 99% confidence level)

FIGURE 3

FRY READABILITY GRAPH

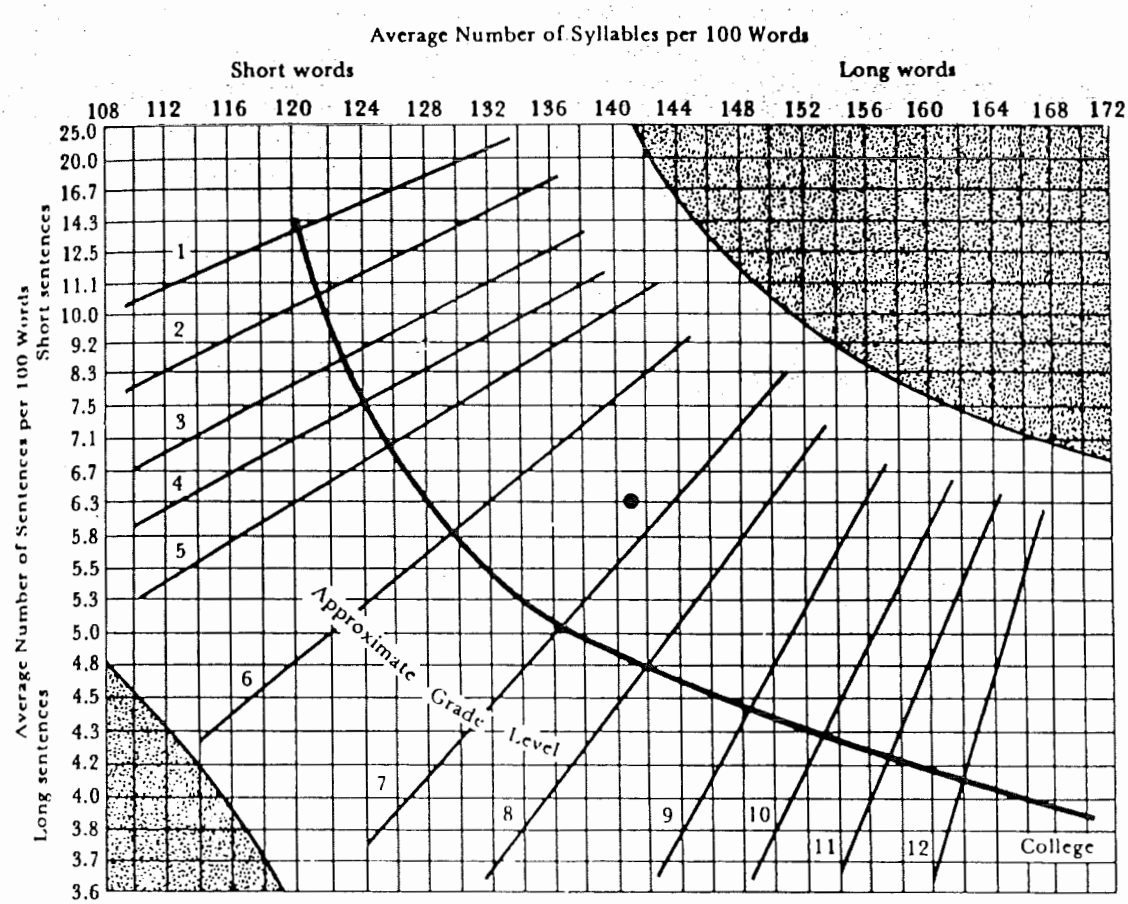


FIGURE 4

READABILITY OF NAEP RELEASED BIOLOGY ITEM

Figure 4. Readability of **NAEP** released biology item.

Biology

Here is a diagram showing the genes for fur color in guinea pigs.

	Male	Female
Parents:	EE	ee
	Brown fur	White fur

	Female	Female	Male	Male
Offspring:	Ee	Ee	Ee	Ee
	Brown fur	Brown fur	Brown fur	Brown fur

According to the diagram, what kind of trait is brown fur in guinea pigs?

- ☐ Recessive
- ■ Dominant
- ☐ Blended inheritance
- ☐ Sex-linked
- ☐ I don't know

Readability formula estimates

SMOG 12.5 RAYGOR 8 FRY 9 FOG 13.8 FLESCH 10