DOCUMENT RESUME

ED 469 763	TM 034 550
AUTHOR	Camel, Christi; Chung, Tom
TITLE	Circumventing the Pressures of Standardized Norm-Referenced Tests.
PUB DATE	2002-05-00
NOTE	73p.; Action research project, Saint Xavier University and IRI/Skylight Field-Based Masters Program.
PUB TYPE	Dissertations/Theses - Masters Theses (042) -- Reports - Research (143)
EDRS PRICE	EDRS Price MF01/PC03 Plus Postage.
DESCRIPTORS	*Achievement Gains; Chemistry; *High School Students; High Schools; Intervention; Scores; *Standardized Tests; *Test Coaching; Test Results; Thinking Skills

ABSTRACT

The state in which the target school is located recently mandated a standardized examination to assess academic progress in high schools. This sparked the interest of the target school's administration and teachers in finding ways to assist students in improving standardized examination performance. A comparison group of 4 classes of approximately 24 chemistry students provided controlled, comparable groups. Two of the four groups received the experimental treatment consisting of four intervention strategies: the Question-Answer Relationship Strategy, the Plus-Minus System, the Multiple Choice Method, and the True/False Method. These methods provided students with the necessary critical reading and thinking skills to use their knowledge and experience effectively in addressing the questions on the standardized examination. Evaluation of the effectiveness of the intervention involved comparison of a standardized examination taken by students from experimental and comparison groups. The change in score indicated the effects of the intervention method, and the statistical significance was evaluated through a two-sample t-test. Results show the average score difference between control and experimental groups to be distinct and significant.The project results show that specific efforts to help students improve their test taking skills, without emphasizing the contents of the examination, can improve the performance of students on standardized, norm-referenced tests. Four appendixes contain sample test items and charts illustrating the intervention strategies. (Contains 18 references.) (Author/SLD)

CIRCUMVENTING THE PRESSURES OF STANDARDIZED NORM-REFERENCED TESTS

Christi Camel

Tom Chung

An Action Research Project Submitted to the Graduate Faculty of the

School of Education in Partial Fulfillment of the

Requirements for the Degree of Master of Arts in Teaching and Leadership

Saint Xavier University and IRI/Skylight

Field-Based Masters Program

Chicago, Illinois

May, 2002

Abstract

The state in which the target high school resides has recently mandated a standardized exam to assess the academic progress of their high schools. This has sparked the interest of the target school's administration, as well as the teachers, to seek methods that could assist students to improve standardized exam performance. A comparison group of four classes consisting of approximately twenty-four tracked chemistry students provided controlled, comparable groups where two of the four groups received the experimental treatments – Question Answer Relationship Strategy (QARS), Plus-Minus System, Multiple Choice Method, and True/False Method. These methods provided students with the necessary critical reading and thinking skills to effectively utilize their knowledge and experience in addressing the repertoire of questions on standardized exams. Evaluation of the effectiveness of the intervention involved a comparison between a sample standardized exam taken by the students from both Experimental and Control Groups before and after the experimental treatment. The change in score indicated the effects of the intervention method, and the statistical significance was evaluated through a two-sample t-test. The results of this test showed the t value of 3.0585. With n – 1 degree of freedom that is equal to 34, the p value equaled 0.0025 with a confidence level of 99.5%; the average score difference between the Control and Experimental Group was distinct and significant. Thus, the project results clearly demonstrate that specific efforts to help students to improve their test-taking skills, without emphasizing the contents of the exams, can improve the performance of students on standardized norm-referenced tests.

SIGNATURE PAGE

This project was approved by

Charles J. Nier

Advisor

Bob Kaplan

Advisor

Beverly Gulley

Dean, School of Education

DEDICATION

We dedicate this paper to Paul, Tessie and Olivia, for all the time that we spent at school and you spent at home. Thanks for the unwavering support and love!

TABLE OF CONTENTS

CHAPTER 1

PROBLEM STATEMENT AND CONTEXT

General Statement of the Problem

Standardized testing has become the instrument of choice to assess the success of high school students across the nation and the world. The state in which the target high school resides has recently mandated a standardized exam to assess the progress of their pupils/schools/ districts. The results of this standardized exam will be recorded in the transcripts of every student, and thus, have an impact on their future education plans. Given the implications from this state assessment plan, the target school's administration, as well as the teachers, began to seek curriculum-based methods to assist the students.

This study investigated the effectiveness of curriculum-based intervention methods to improve students' test-taking skills. The skills enhance the abilities of students to effectively utilize their knowledge and experience by improving their process knowledge – the means to apply their educational experience. This involves necessary critical reading and thinking skills to address the repertoire of questions commonly found on standardized exams.

Immediate Problem Context

This high school is one of two in a high school district that has experienced significant amount of growth over the past decade. In the next decade, the student population is expected to rise 10% and should reach nearly 5,000 students by the year 2005. Recently, U.S. News and World Report (January, 1999), listed this high school as an "Outstanding American High School" and Newsweek Magazine named it one of "100 Schools that Work". In order to meet the wide variety of educational needs of all its' students, the target school offers 194 courses in 14 subject areas, 13 advanced placement courses, an Individualized Instruction Center, a Center for Individualized Curriculum, and an Off Campus Learning Center. The school day is divided into 15 minute "mods", with nonlab-based classes meeting for three mods each day, and lab-based classes meeting for four mods. This allows students more flexibility in class choice and scheduling.

	1996	1997	1998	1999	2000
Total enrollment	4,111	4,172	4,173		4,424
• White	61.0%	60.3%	59.4%	59.3%	60.4%
• Black	1.7%	1.9%	1.8%	2.3%	2.6%
• Hispanic	5.4%	5.8%	5.7%	5.8%	6.2%
• Asian/Pacific Islander	31.7%	31.9%	33.0%	32.5%	30.8%
Low income students	7.2%	7.4%	7.8%	10.4%	6.2%
Limited-English-Proficiency	5.6%	4.9%	5.8%	5.9%	6.1%
Drop-out rates	---	---	0.7%	1%	0.4%

Figure 1: Student Profile

According to the 2001 Illinois School Report Card, the graduation rate of the target school is 94.2%, with 93.8% of its graduates continuing their education, 68.4% of them at a four-year institution. In 2000, the composite ACT score of 22.1 exceeded the national average of 21.5 with approximately 80.2% of the class of 2000 taking the test. In English, students averaged 21.3, in Reading they averaged a 21.7, in math they averaged a 23.2 and in Science Reasoning they averaged a 21.8.

On the Illinois Standards Achievement Test (ISAT) grade 10 students performed as follows:

% of Students		
Subject	met or exceeded standards	Below standards or received academic warnings
Reading	74%	26%
Mathematics	72%	28%
Writing	83%	17%

Figure 2: Student Achievement, ISAT

The average class size is 19.5 students. Approximately 88% of parents have had personal contact with the school staff during the school year. There are 299 classroom teachers in the district:

- 95.1% are White
- 1.0% are Black
- 2.0% are Hispanic
- 1.9% are Asian/Pacific Islander
- 47.6% are male, 52.4% are female
- Teachers have an average of 11.9 years of teaching experience
- 73.6% have a Master's degree or above
- The pupil to teacher ratio is 17.8:1
- The average teacher salary is $67,154 with the average administrator salary at $95,764
- The operating expenditure per pupil is $13,672

The Surrounding Community

The following is a brief summary of the community setting in which the target school resides. Town A and B residents comprise the largest portion of students in the target school with a few students coming from Town C. Town D residents all attend an alternate high school, and are included to show a more extensive community background.

	Town A	Town B	Town C	Town D*
Population	59,432	22,405	28,284	11,365
Average Age	41.4	42.8	54.4	44.9
Male (%)	46.8%	48.2%	46.8%	47%
Female (%)	53.2%	51.8%	53.2%	53%

Figure 3: Community Profile

	Town A	Town B	Town C	Town D*
White	77.1%	81.0%	88.6%	79.4%
Black	2.2%	0.3%	0.2%	0.4%
Asian	15.6%	15.0%	7.0%	15.6%
Hispanic	4.1%	2.8%	3.1%	3.3%
American Indian/ Eskimo/Aleut	0.1%	0.1%	0.1%	---
Other	---	0.8%	1.1%	1.3%
% native born citizens	72.1%	77.3%	78.2%	72.4%
% foreign born citizens	27.9%	22.7%	21.8%	27.6%

Figure 4: Community Racial Composition

	Town A	Town B	Town C	Town D*
Median income • Household	$42,276	$47,808	$38,718	$57,309
• Family	$49,489	$52,815	$45,822	$62,419
• Nonfamily household	$22,775	$22,193	$20,554	$28,576
• Per capita income	$20,595	$20,206	$17,422	$27,154
Unemployment Rates	1990→ 2.4% 1995→ 3.4%(1) 1996→ 3.4%(1) 1997→ 2.7%(1) 1998→ 2.8%(1)	1.8%	2.0%	2.6%
% of individuals living below poverty line	3.9%	2.9%	3.7%	2.2%
% of families living below poverty line	2.7%	2.3%	2.5%	1.5%

(1) Source: Illinois Department of Employment Security, March 1999.

Figure 5: Community Economic Status

	Town A	Town B	Town C	Town D*
Preprimary, public	1.0%	1.1%	0.7%	1.1%
Preprimary, private	1.1%	0.8%	0.5%	1.2%
Elementary/high school, public	12.4%	11.7%	8.4%	12.4%
Elementary/high school, private	2.2%	2.5%	2.5%	2.4%
College, public	4.9%	5.5%	4.5%	4.8%
College, private	2.8%	2.5%	2.6%	2.8%

Figure 6: School Enrollment (all persons ages 3 and up)

	Town A	Town B	Town C	Town D*
Median years of school completed	14.4	---	12.8	15.1
Less than 9th grade	6.4%	7.4%	12.5%	5.2%
9th-12th grade	8.2%	8.5%	12.5%	6.3%
High school graduate	22.8%	25.8%	31.6%	20.5%
Some college	20.6%	22.6%	18.3%	22.2%
Associate's	5.4%	6.1%	4.8%	4.0%
Bachelor's	21.7%	19.6%	14.4%	24.4%
Graduate degree	15.1%	10.0%	5.8%	17.4%

Figure 7: Educational Attainment (all persons ages 25 and up)

As can be seen in Figure 4 (p. 4, above), the community from which the target school draws is a very ethnically and racially diverse one in which a variety of languages other than English are spoken. These include Asian/Pacific Island, Greek, Korean, Polish, Spanish, and Russian among many others. From Figures 5 through 7 (above), the community can be described as a fairly affluent one with high educational attainment.

National Context of the Problem

Since the inception of the educational institution, the assessment processes, giving tests to students and using the results to evaluate student performance, have been a vital component to education (MacLean and Lockwood, 1996). These assessments provide the means to ascertain not only the performance of the students for students themselves, but also for parents, and teachers for the purpose of improving student instruction, and ultimately, their learning in classrooms (Baresic and Gilman, 2001; McLean and Lockwood, 1996). Teachers commonly taught core knowledge and skills that were readily assessed by using exams that were specifically geared to evaluate the concepts taught. The results of these exams provided an evaluation of students within a particular subject and teacher. The results were teacher and content specific, and as a result, had little to no value when compared with the student's score from another teacher, class, or school (McLean and Lockwood, 1996).

In recent decades, educational reform efforts largely have been driven by assessment and assessment results in a context that is not limited to teacher directed in-class performance assessment, but rather in the context of the national comparison. To the American people, higher education has appeared as a stepping-stone to success for themselves and their children; it represented an educational avenue where children of lower socioeconomic status can climb the social ladder of success (Popham, 2000). Up to the middle of the Twentieth Century, U.S. public schools by in large appeared successful in the eyes of the public. However, by 1970s, with the ever more common news of failing educational systems in our country, the American people began an era stressing educational accountability (Popham, 2000). People no longer held the belief that the U.S. public schools provided the necessary educational experience for our children. With an increasing incredulity towards our schools, instead of merely increasing the funds for schools and school programs, the public demanded that schools demonstrate their effectiveness in educating their students (McLean and Lockwood, 1996). The measuring tool by which the educator's quality was to be judged was a standardized, norm-referenced test. It is a test that measures students' knowledge and skills where individual student performance is compared with those possessed by a national sample of students of the same age or grade level (Baresic and Gilman, 2001).

The two major categories of tests include norm-referenced tests and criterion-referenced tests (McLean and Lockwood, 1996). The criterion-referenced test measures the content of specific curriculum. Results indicate the level of mastery of the topics that were taught in a given curriculum; it is reported in percentage of items answered correctly or "as achievement levels that are compared with a predetermined 'cut score'" (McLean and Lockwood, 1996, p.4). The results of the criterion-reference test provide a valuable baseline of information where the teachers and schools can improve the design and plan of both instructional programs and techniques. The norm-referenced test is "an assessment tool that permits someone to make a valid inference about the knowledge and/or skills that a given student possesses in a particular content area" (Popham, 1999, p.9). The results are reported in terms of how students perform with respect to a comparison group. This comparison group may be a national, state, or local sampling of students. When the norm-referenced exam is administered and scored in a predetermined, standard manner, it constitutes a standardized exam. Thus, a standardized national norm-reference test is administered in a predetermined format to a national sample of students where the scores of any individual student represent a percentile rank among the students in the nation who performed at or below the level of that particular student. A norm-referenced test score does not show how much or how well a student has mastered knowledge and/or skills in a content area; rather, it provides "inferences relative about a student's status with respect to the mastery of knowledge and/or skills in particular subject area" (Popham, 1999, p.9). There are two major kinds of standardized, norm-reference tests: the Aptitude Test and the Achievement Test. The standardized Aptitude Test measures the student's potential in succeeding in some subsequent educational setting. The Scholastic Achievement Test (SAT) and American College Testing (ACT) are common examples of aptitude tests. The Standardized

Achievement tests show the student's status with respect to mastery of core knowledge and/or skills in a particular subject area such as mathematics, science, and Social Studies. Nationally, five common tests are in use: California Achievement Test, Comprehensive Test of Basic Skills, Iowa Test of Basic Skills, Metropolitan Achievement Tests, and Stanford Achievement Tests.

The use of the criterion-reference test has primarily focused on instructional improvements often sought by educators; it is a formative, rather than summative, in nature. However, the primary use of the norm-referenced test has exceeded the bounds of its purpose or design (Popham, 1999). Although the purpose and use of the standardized norm-referenced test was to depict relative strengths and weaknesses in a subject area, it has become the basis in which the public determines and judges the teacher and/or school effectiveness, the allocation of funds, and student promotion, demotion, and graduation (Baresic and Gilman, 2001; Whitehead and Santee, 1987). Advocates believe that standardized norm-referenced tests are "the only way to hold schools and teachers accountable, as well as the only way to determine student achievement" (Baresic and Gilman, 2001, p.1). In addition, they see it as "a quick, relatively inexpensive, and an objective way to determine student achievement and growth" (Baresic and Gilman, 2001, p.1). As a result, "students, teachers, administrators, schools, and school systems are often evaluated and compared solely on the basis of test scores" (Baresic and Gilman, 2001, p.2).

Given the increased value placed upon standardized tests, educators have experienced and are experiencing "relentless pressure to show their effectiveness" (p.9), and as a result teachers are using ever-increasing variety of assessment tools (Popham, 1999). Joseph Angaran, a 3rd grade teacher from Egan, Minnesota, notes how 3rd graders in his school are evaluated using a national norm-referenced test for school achievement, a test to determine school ability, two state-mandated tests in reading and math, a writing assessment, and a test for oral retelling of a narrative or a descriptive text. This inordinate amount of testing of students not only "impedes a meaningful education for the students" (p.72), but also inundates and overwhelms the school officials with test data that often leads to misinformation to the teachers who seek feedback for instructional improvements (Angaran, 1999). In addition, Angaran notes,

> Parents receive less information with even less interpretation of the results and, more important, no direct link to what they can do at home to help their child. The public often receives the most dramatic data taken out of context, misinterpreted, or sensationalized. (p. 72)

In addition, due to the heavy emphasis placed on norm-referenced tests, many teachers "find themselves trying to teach to the test or something that is close as possible to the test," and as a result, teachers "concentrate their curriculum on skills and knowledge that are relevant to the test...[they may even] suspend the regular curriculum for weeks in order to provide intense cram sessions for their students" (Baresic and Gilman, 2001, p.2). Thus, the emphasis of classroom instruction is often "placed not on improving student education, but rather on getting their (norm-referenced) scores up to a minimum standard of acceptable performance" (Baresic and Gilman, 2001, p.2).

The standardized norm-reference test is ideal in "supplying evidence needed to make norm-referenced interpretations of students' knowledge and/or skills in relationship to those of students nationally" (Popham, 1999, p.10). In addition, it may even provide valuable information on the longitudinal development of a child in a given content area. However, according to Popham (1999), given the nature of the norm-referenced test and its design, the current method of using the standardized norm-referenced test to measure the educational success or effectiveness of schools is analogous to "measuring temperature with a tablespoon" and as a result, "those evaluations are apt to be in error" (p.8). The goal of the developers of the norm-referenced test is to create questions that can only be answered correctly by roughly half of the assessed students. These questions discriminate among students in order to yield norm-referenced interpretation or normal distribution of scores. Thus, because the norm-referenced test measures relative achievements, it does not assess the amount of knowledge gained through schooling for a given student. Furthermore, given the magnitude of topics that need to be assessed plus the limited time in which these tests can be administered, the norm-referenced tests "contain too few items to allow meaningful within-subject comparisons of students' strengths and weaknesses" (Popham, 1999, p.9). Thus, given current test design, norm-referenced tests at most provide rough approximations of a student's relative status to the norm population, and it does not measure the actual amount learned within classrooms and/or quality of school or education.

The developers of the norm-referenced tests are forced to design "one-size-fits-all" tests that should properly assess curricular emphasis of schools and school districts. However, given the nation's enormous diversity of curriculum, the attempt to "fit " all curricula with one exam is impossible. Therefore, standardized norm-referenced tests will contain items that are not aligned with a particular curricular setting (Popham, 1999). Furthermore, because classroom instruction is often influenced by the contents of textbooks, Freeman and his colleagues (1983), studied four common math textbooks and compared the contents of the textbooks with the standardized norm-referenced test. Their study indicated that between 50 to 80 percent of the test items on standardized achievement tests were not suitably addressed in the textbooks. Therefore, according to Popham (1999), "the assumed match between what's tested and what's taught is not warranted" (p.11).

According to Howard Gardner, there are eight different intelligences. Individuals possess one or more of the intelligences with some being more developed than others. The degree and type of intelligences within individuals vary; they are unique to a given individual. The standardized achievement tests usually emphasize and assess a limited number of intelligences, often only verbal linguistics and logical mathematics. As a result, students who have greater capacity in other intelligences, such as interpersonal and kinesthetic, will appear to possess less aptitude than the norm population. Furthermore, children possess different life experiences due to the different environment or culture in which they have lived. "If children come from advantaged families and stimulus-rich environments, then they are more apt to succeed on items than will other children whose environments don't mesh as well with what the tests measure" (Popham, 1999, p.13). For instance, students from higher socioeconomic status have tendencies to do better on the standardized achievement tests than students from lower socioeconomic status families. Thus, given the properties of the norm-referenced tests, while it may provide a valuable comparative information the on the students performance in a specific content area, norm-referenced tests have limitations when quantifying or measuring the amount of content learned by a student from a teacher or a school. As a result, due to the limitation of the standardized norm-referenced tests, while some schools and/or states continue to mandate higher standards and more tests, many other states (such as Arizona) have discontinued or limited the use of their assessment programs (Baresic and Gilman, 2001).

In the midst of continuing pressure to assess the schools, teachers are seeking data that help to improve instructional quality in order to better educate the students. Both parents and the general public seek data that can be used to measure the quality of education and thus hold schools accountable for their role as educators. The two types of exams, criterion-referenced tests and norm-referenced tests, both have their unique purposes that are premised on the same goal, which is to provide positive learning experiences for all students. However, their results possess significance in limited scope, and consequently results should be interpreted and applied only in the range in which the assessment tool was originally meant to be used. In the realm of assessments in schools today, a common assessment tool that serves the purposes of educators and the public should be implemented; this assessment tool should fairly assess the target group, alleviate pressure on educators, as well as appease the concerns of the parents and other community members.

CHAPTER 2

PROBLEM DOCUMENTATION

Problem Evidence

In order to document the decline in performance on certain standardized tests, both American College Testing (ACT) Program and Scholastic Achievement Test (SAT) scores were obtained through school report cards. These report cards compiled school, district, state, and national data. As shown in Figures 8 and 9 (below), the number of students taking the ACT and SAT is substantial, making it imperative that more students be exposed to an education that not only teaches them content, but teaches them how to retrieve that information and apply it to other situations, including the standardized test.

(% of class)	1992	1993	1995	1996	1997	1998	1999	2000	2001
School	88.2%	82.6%	83.6%	78.7%	81.8%	72.3%	71.7%	75.5%	79.8%
District	89.0%	86.0%	86.0%	83.3%	86.6%	79.4%	76.3%	80.2%	83.3%
State	61.9%	62.2%	63.0%	63.2%	61.2%	58.3%	55.9%	62.4%	62.3%

Figure 8: Students taking the ACT

(% of class)	1992	1993	1995	1997	1998	1999
School	27%	29%	32%	40%	31%	36%
District	15%	15%	13%	14%	13%	12%
State	42%	43%	41%	42%	43%	43%

Figure 9: Students taking the SAT

In Figure 10, a decline in both school and district scores on the ACT show the need to improve test scores, even though the target school scores higher then the national average.

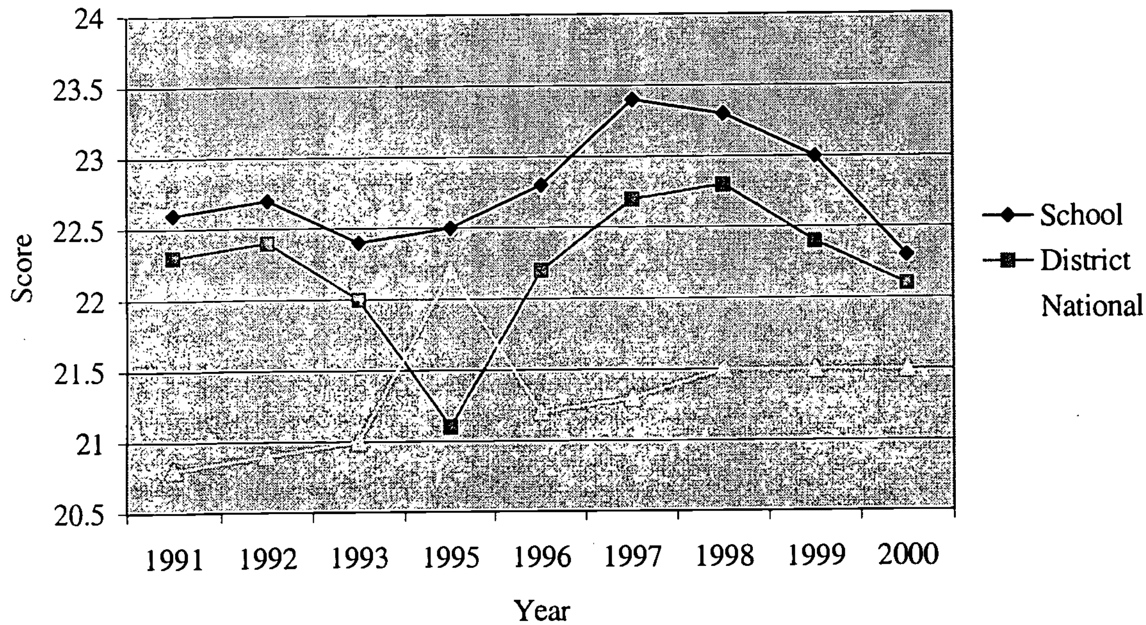


Figure 10: ACT Scores

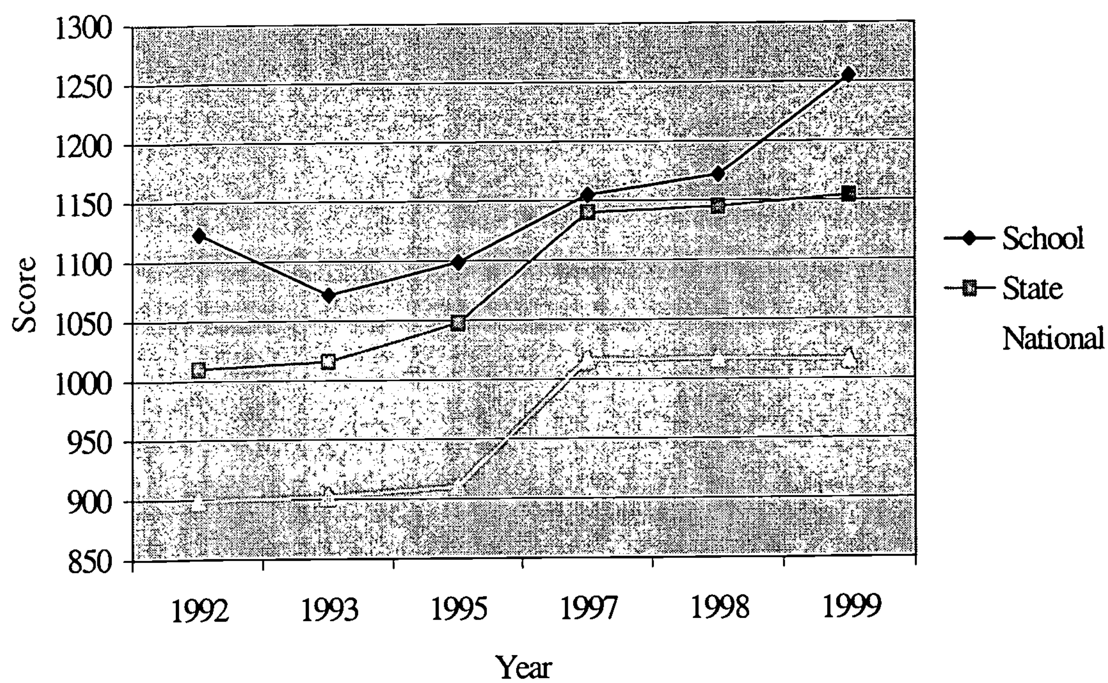


Figure 11: SAT Composite Score

BEST COPY AVAILABLE

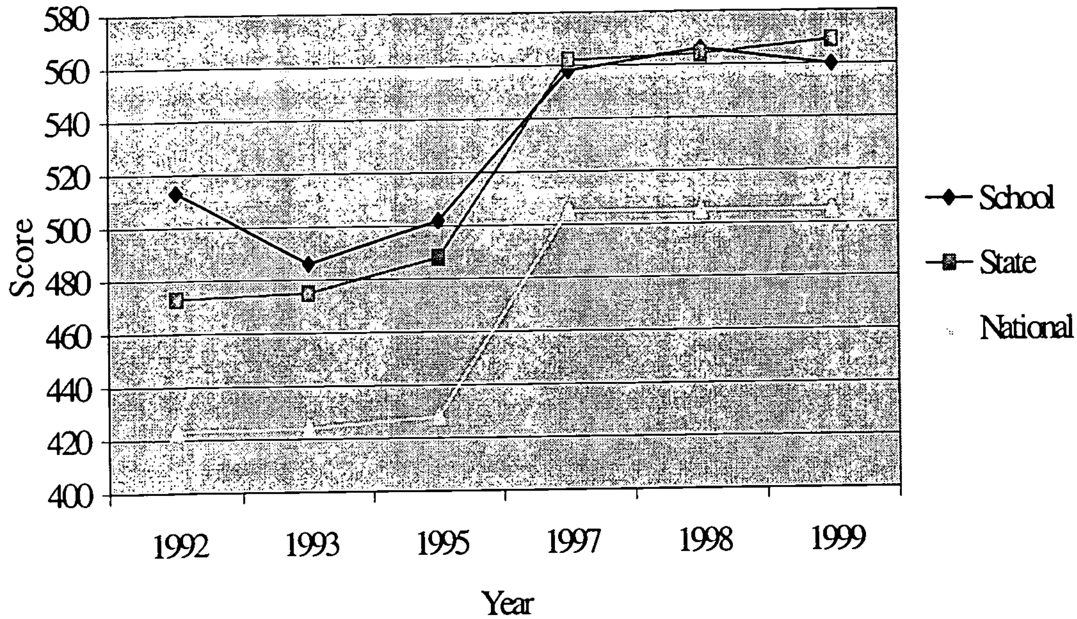


Figure 12: SAT Verbal Scores

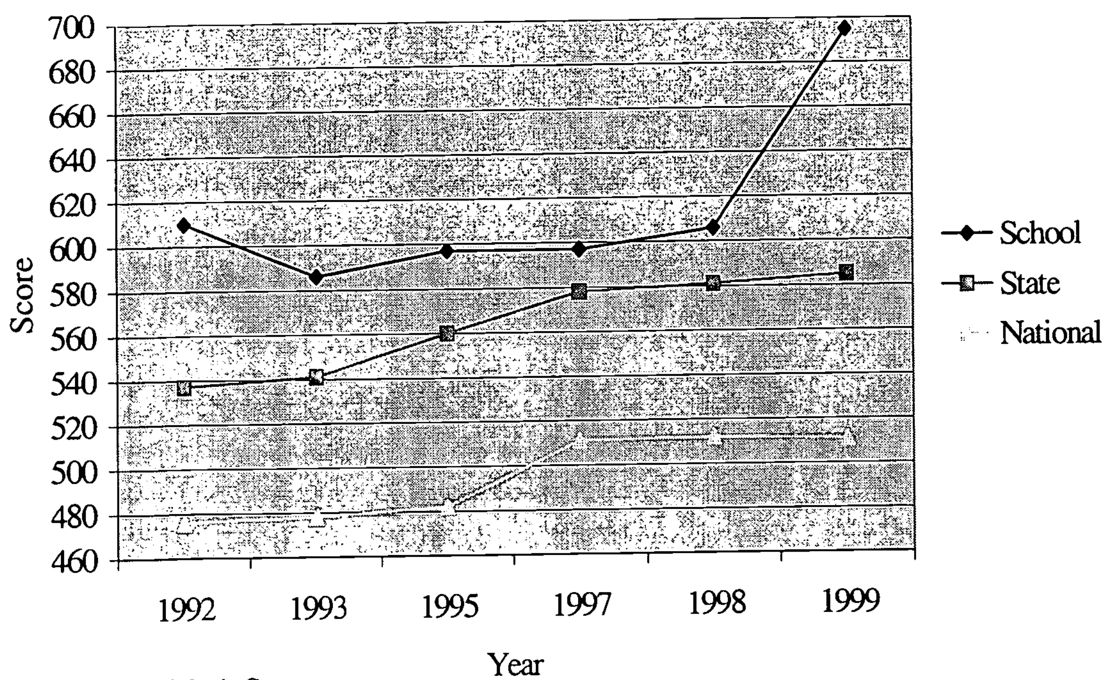


Figure 13: SAT Math Scores

As in shown in Figures 11 and 12 (p. 15 and p. 16), there has been a slight increase in overall SAT scores in the school, but there is still much room for improvement. Note that scores on all four dimensions are well above the national norm, and on three of four scores are above state levels. While this can be seen as an indicator of district/school success, the press for increased pupil achievement by district educators and community is very likely a reflection of the socioeconomic/educational levels of the community itself (see Chapter 1). Quite possibly all parents and school personnel would like to see their children do well; in this community, academic achievement is undoubtedly a high priority in every respect. This concept is detailed below.

Problem Cause

Most teachers will agree that their job is to transfer knowledge to their students. This information makes up a huge knowledge base that students acquire from the time they begin school through college and the rest of their lives. Unfortunately, the norm-referenced test does not gauge how much students know, but rather how well they perform in comparison to other students. So while teachers are teaching content, the school is using a test designed to measure application. This great divide is one of the major issues in education today. At the target school, as is true across the nation, good teachers shy away from "teaching to the test". According to Bond (1995), the answer is not to teach to the test, but rather to teach students content in addition to teaching students test-taking skills. By doing this, students still learn the material for the course, but also learn how to utilize that material to a situation where they are being tested on application (as opposed to content). The assumption in the past has been that if one teaches a subject area well enough, students will learn and apply that knowledge in all types of situations. As can be seen from the decline in standardized tests scores, that is not necessarily the case.

According to Mehrens (1989), it is not merely enough to teach content, but test-taking skills as well. The same is true with teaching to the test. It is not feasible to show students problems and teach the answers to those problems, since that will not improve test scores either. Instead, those students with a strong knowledge base to support the skills necessary to be successful on norm-referenced tests will be the successful ones.

The increasing importance of these tests, as shown by the elevation of property values in some communities, puts pressure on the school board to raise test scores. That pressure is in turn put on administrators who feel that unless test scores improve, jobs will be lost. Finally, that pressure gets transferred to teachers, who historically had to decide between teaching content and teaching the test. As a result, this is raising the anxiety and stress especially among students, parents, and schools, which has shown to negatively influence not only the students but also the schools and the community (Popham, 2000). It is possible not only to teach students content but also help improve their test-taking scores by integrating strategies on how to be successful test-takers. As seen in the notable reputation of the target school, the students are learning the content of their classes. Unfortunately, this does not seem to transfer 100% to an outstanding performance on norm-referenced tests. Standardized tests have been used in the United States as a predictor for future success since their inception. The question is whether or not this is a valid application of the results of a standardized test (Sacks, 1997). Norm-referenced tests have become a tool used by some high schools to even dictate who graduates, and is used by most colleges to decide who enters. Many argue that the extreme amount of classroom time that is taken away just to take standardized tests in a students career is the real problem (Baresic and Gilman, 2001). Schools cannot continue to teach a curriculum that is solely based on content if norm-referenced tests are going to be the ruler against which we measure success.

CHAPTER 3

THE SOLUTION STRATEGY

Literature Review

Standardized norm-referenced exams have become the instrument of choice to not only assess the students performance but also the performance of schools and/or school districts (Popham, 1999). Although the standardized norm-referenced exams are not inherently designed to fulfill this goal, pressures are continuously placed upon schools by society to present evidence of educational success: The public wants accountability from the education institution in the light of common negative media depictions of school inequities (Popham, 2000). Furthermore, the public views the standardized norm-referenced tests as "a quick, relatively inexpensive, and an objective way to determine student achievement and growth" (Baresic and Gilman, 2001, p.1). Given this pressure to perform, many schools have started to align their curriculum around the contents of the standardized norm-referenced exams. According to Mehrens (1989), a more appropriate way to prepare students is to teach test-taking skills that would facilitate students' application of knowledge and experiences in education rather than teaching the exams to the students. Students with a firm knowledge base may be taught the basic test-taking skills with little instructional time in order to elevate scores on standardized norm-referenced tests (Mehren, 1989). Some of the intervention methods that have shown to improve test performances are systematic test-taking approaches, recognition of key words, and, as a last resort, using guessing strategies (Frierson, 1984).

The target school, ranked as one of the top schools in the state, offers a very well balanced education, which includes a strong knowledge base. The target school's curriculum is content driven which aligns with the community's values and emphasis on the need to prepare the students for post-secondary educational standards and expectations. However, in last five years, the average ACT scores of the have shown gradual decline in scores while the national average continuously, yet steadily, increased in the same period of time (see Chapter 2). Thus, Mehren's study, which showed a positive impact of teaching students test-taking skills as long as it is in conjunction with a strong knowledge base, provides an ideal approach to intervention for the target school; this study will provide the basis in which the problem at the target school will be addressed. Furthermore, Berendt and Koski (1999) and Frierson (1984) have also shown the positive effects of teaching test-taking skills through intervention methods to address low standardized exam scores.

Project Objective and Processes

A series of test-taking skills will be taught to groups of students in chemistry classes at the target school as the intervention method. The target school offers four levels of chemistry: Chemistry 10-20 (remedial), Chemistry 12-22, Chemistry 11-21 (honors), and AP Chemistry. Thus, for a given level, student placement will establish a random sampling of students for each class that can be used as comparative groups within a given level -- Experimental versus Control Groups. In group comparison, all variables can be considered to be constant except the experimental factor. For this research, four Chemistry 12-22 classes that are taught by two separate teachers will be used to conduct the study. Each teacher has two chemistry classes. For each teacher, one class will be the Experimental Group and the other class will be the Control Group; the selection process will be random. Thus, upon completing the experimental treatments, two experimental trials will be completed that can potentially be used to support the reliability of the data. The intervention method for this study implemented the following procedures between the period of early October and mid-February.

Project Action Plan

- Week 1: All four of the groups/classes will be given a preliminary standardized and the scores will be collected.
  - A science reasoning section from the ACT exam consisting of 40 questions that are to be completed in 35 minutes will be used as a preliminary exam. A copy of this exam can be found in Appendix A.
  - A standard environment will be set and maintained throughout the examination.
- Weeks 2-5: Both the Control and the Experimental Groups will be taught the same content in the chemistry curriculum, with the Experimental Group receiving instruction in the intervention method. Each intervention will be taught in a one-week period. Once a week, for approximately the first fifteen minutes of the class, students will learn and practice the intervention.
  - Four intervention strategies will be taught. The tables and graphic organizers describing the intervention strategies are included as Appendix D.
    1) The Question-Answer Relationship Strategy (QARS)
       - Helps the students to readily identify the types of questions and readily locate synthesize answers.
    2) The Plus/Minus System
       - Time management strategy where students answer questions according to the difficulty and time consumption for a given question.

3) Multiple Choice Methods

  - Process sequence (questions vs. answers or vice versa)
  - Process of elimination
    - Key word recognition
    - High/Low numbers
    - Long vs. short answers
    - Ridiculous answers.

4) True/False Method

  - True unless obviously false
  - Absolute qualifiers

- In the subsequent week, students will have fifteen minutes to practice using samples from the Science Reasoning section of the ACT test.
  - A sample Science Reasoning section or passage containing on the average of four to five questions will be used to help students practice the test-taking skills learned in the previous week's lesson. An example is included as Appendix C.
  - Each sample Science Reasoning section will be approximately 5 minutes in length.
  - One of the practice sessions will be timed in a standard environment to promote student familiarity of the test format/environment.
- Approximately four days out of two weeks will be used to review the test-taking skills by practicing all four of the intervention strategies prior to the post/final assessment of the Experimental Groups.

- On days one and three, a section of the Science Reasoning test will be administered to the students in standardized format; the seating will be organized and structured, the directions will be read, and the students will be timed.
- On days two and four, class will be used to evaluate student performance and the nature of the questions in the exam following each practice. This analysis will provide students an opportunity to evaluate their application of test-taking skills and fine-tune their skills.

- Week 6: A full sample Science Reasoning section was used as a post-test for both the Control and Experimental Groups. A copy of the post-test is included as Appendix B. Pre-test scores will be used to compare the effects of the intervention methods. A copy of the pre- and post-test can be found in the appendix.

Methods of Assessment/Solutions Expected

The goal of this project is to measure the effects of the intervention method on test-taking skills on samples of students that are randomly placed in groups according to their academic placement at the target school. The effects of the intervention methods will be measured through the test scores obtained by the students on a sample Science Reasoning Section that has 40 questions; this needs to be completed in the time frame of 35 minutes in a standardized environment. The data will entail pre-and post-test scores for both the Control and Experimental Groups. A comparison will be made by subtracting the post-test scores from the pre-test scores, and the difference will outline any improvements in student performance in the standardized norm-referenced exams. This comparison will be conducted for both the control and Experimental Groups in order to negate any influence from the chemistry curriculum (such as the content learned and/or the chapter exams students have taken).

CHAPTER 4

PROJECT RESULTS

Historical Description of the Intervention

The objective of this research was to improve students' scores on norm-referenced standardized tests. This was done by teaching students test-taking strategies that would help them recall prior knowledge and apply it to new situations. Initially, the students in all four classes were given a pre-test. Then, four strategies were taught; Question-Answer Relationship Strategy (QARS), Plus/Minus System, Multiple Choice Methods, and True/False method. Each intervention was slated to be taught in a one week period, one fifteen minute block of time dedicated to teaching the skill, and an additional fifteen minutes later in the week to practice the skill. At the end of the four weeks necessary to teach the skills, students received two additional days of practice each with a day of analysis of the practice exam, used to help students evaluate their success with the skills. During week six, students were given a complete Science Reasoning post-test.

During the time period in which the intervention was applied, the target school had numerous days with shortened schedules, causing the entire study to take twelve weeks as opposed to six. Students often went a week without being taught or practicing a skill. Students then needed a mini-review of the skills previously learned. This meant that some days were nearly completely devoted to the research in the Experimental class, while the Control class was taught content. After the ten weeks, the Experimental class was slightly behind the Control class and needed time to catch up so that they arrived in the same place at the end of the semester. In the first semester of 2000, ten chapters were covered. In the first semester of 2001 (the research semester), students also finished ten chapters, though in less detail then previous years.

Presentation and Analysis of Results

Figures 14-32 (p. 25-35), portray results for all test given by Teacher One and Teacher Two to both Experimental and Control Groups. Explanation, comparisons, and analysis of project results will be found beginning on p. 36.

Student No.	Pre-Test Score	Post-Test Score	Difference	% Change
1	7	20	13	32.5
2	16	17	1	2.5
3	9	16	7	17.5
4	14	18	4	10
5	18	16	-2	-5
6	20	15	-5	-12.5
7	22	14	-8	-20
8	19	15	-4	-10
9	18	28	10	25
10	13	15	2	5
11	21	23	2	5
12	19	16	-3	-7.5
13	21	23	-2	-5
14	10	15	5	12.5
15	19	13	-6	-15
16	20	17	-3	-7.5
17	12	11	-1	-2.5
18	29	30	1	2.5
19	11	13	2	5
20	10	16	6	15
Mean	16.4	17.55	0.95	2.375
Median	18	16	1	2.5
Minimum	7	11	-8	-20
Maximum	29	30	13	32.5
Std. Dev.	5.500239229	4.93617154	5.423875096	13.55968774

Figure 14: Teacher One Control Group

Student No.	Pre-Test Score	Post-Test Score	Difference	% Change
1	23	27	4	10
2	23	21	-2	-5
3	17	23	6	15
4	15	20	5	12.5
5	19	12	-7	-17.5
6	13	10	-3	-7.5
7	25	22	-3	-7.5
8	18	22	4	10
9	12	18	6	15
10	7	20	13	32.5
11	17	14	-3	-7.5
12	19	20	1	2.5
13	11	11	0	0
14	14	18	4	10
15	18	19	1	2.5
16	13	16	3	7.5
17	13	23	10	25
18	14	22	8	20
Mean	16.16666667	18.77777778	2.611111111	6.527777778
Median	16	20	3.5	8.75
Minimum	7	10	-7	-17.5
Maximum	25	27	13	32.5
Std. Dev.	4.630461799	4.5961052	5.123634841	12.8090871

Figure 15: Teacher One Experimental Group

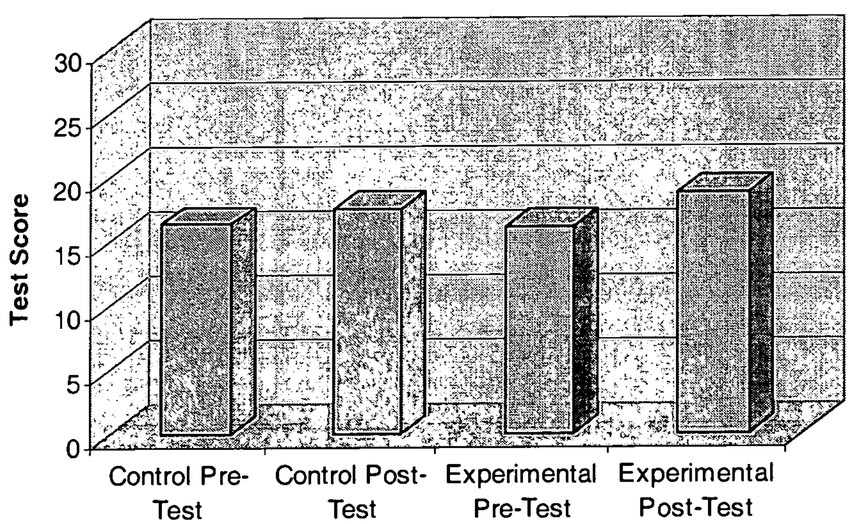


Figure 16: Average pre- and post-test scores, Teacher One (Experimental and Control Groups).

Teacher One	Number of Students in the Control Group	Number of Students in the Experimental Group
Students with no change in test score	0	1
Students with a decrease in test score	9	5
Students with an increase in test score	11	12
Students with 0% > Score Increase ≥ 5%	5	2
Students with 5% > Score Increase ≥ 10%	1	4
Students with greater than 10% increase in the test score	5	6

Figure 17: Student Performance of Teacher One – a comparison between pre- and post-test.

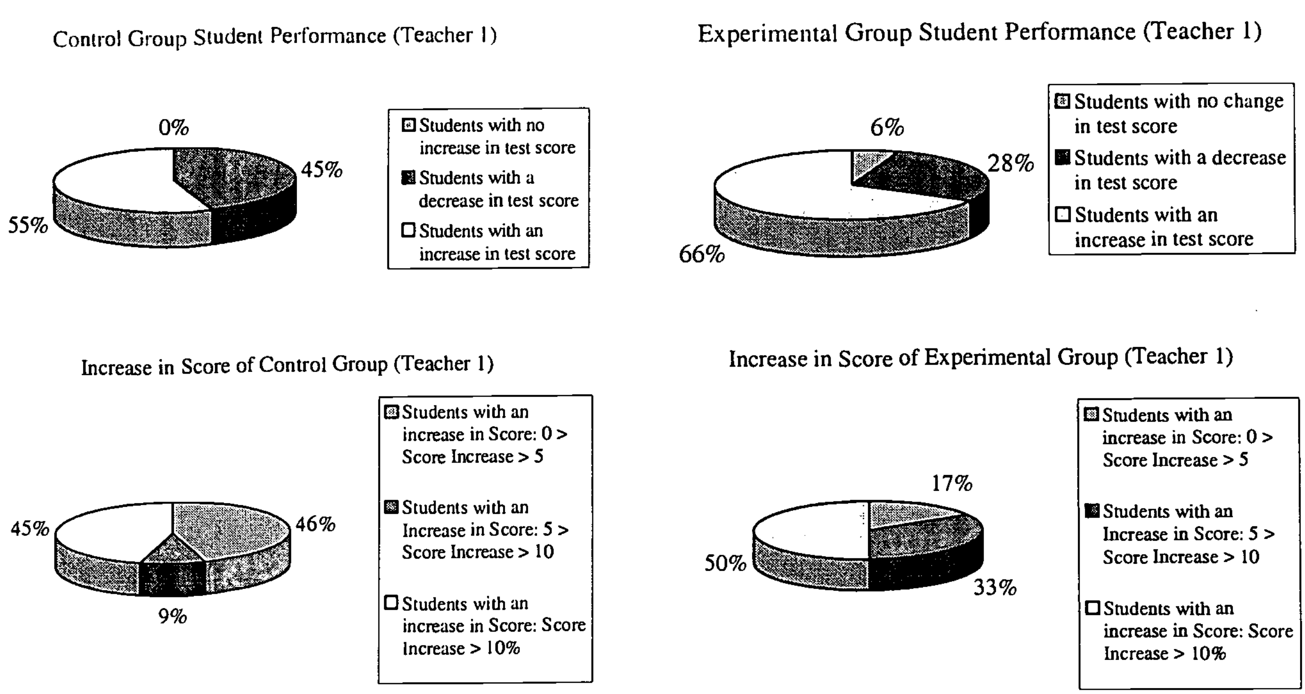


Figure 18: Student performance of Teacher One (post-test minus pre-test score)

Student No.	Pre-Test Score	Post-Test Score	Difference	% Change
1	13	12	-1	-2.5
2	20	18	-2	-5
3	23	19	-4	-10
4	12	8	-4	-10
5	14	13	-1	-2.5
6	9	3	-6	-15
7	18	9	-9	-22.5
8	18	15	-3	-7.5
9	12	18	6	15
10	22	24	2	5
11	17	13	-4	-10
12	13	8	-5	-12.5
13	22	23	1	2.5
14	16	21	5	12.5
15	9	14	5	12.5
16	18	22	4	10
17	13	15	2	5
18	22	19	-3	-7.5
Mean	15.82352947	15	-0.944444444	-2.361111111
Median	16	15	-1.5	-3.75
Minimum	9	3	-9	-22.5
Maximum	23	24	6	15
Std. Dev.	4.40504439	5.916079783	4.263003943	10.65750986

Figure 19: Teacher Two Control Group

Student No.	Pre-Test Score	Post-Test Score	Difference	%Change
1	17	27	10	25
2	9	17	8	20
3	11	10	-1	-2.5
4	17	19	2	5
5	13	20	7	17.5
6	17	20	3	7.5
7	12	19	7	17.5
8	18	24	6	15
9	19	23	4	10
10	20	21	1	2.5
11	18	22	4	10
12	16	19	3	7.5
13	23	25	2	5
14	12	20	8	20
15	13	16	3	7.5
16	22	24	2	5
17	23	24	1	2.5
Mean	20	25.5	4.117647059	10.29411765
Median	20	20	3	7.5
Minimum	9	10	1	-2.5
Maximum	23	27	10	25
Std. Dev.	4.25907336	4.016510045	3.038962671	7.597406678

Figure 20: Teacher Two Experimental Group

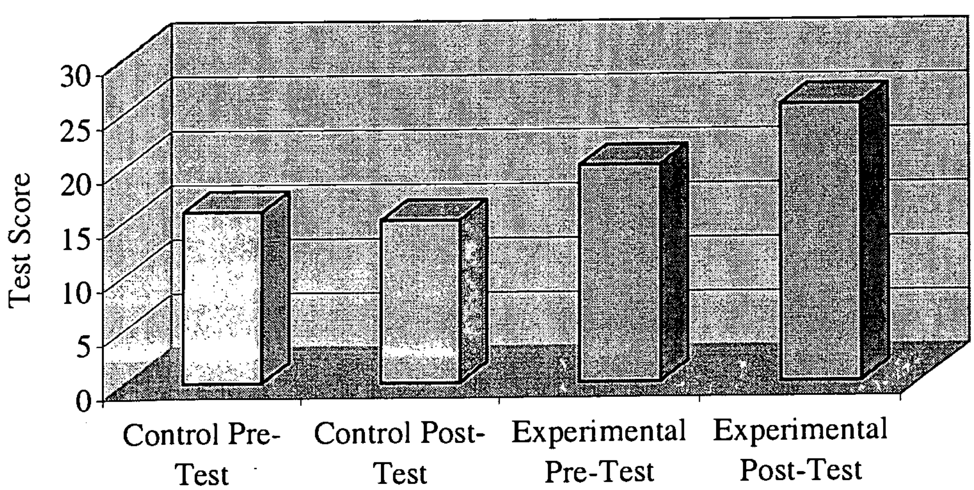
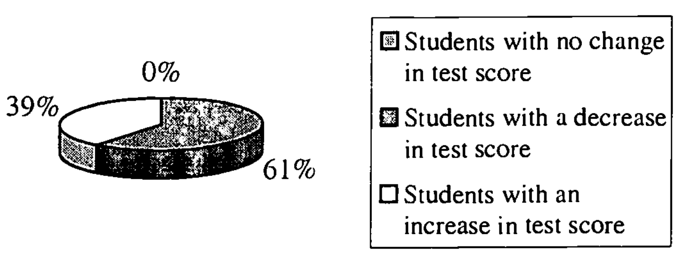


Figure 21: Average pre- and post-test scores, Teacher Two (Experimental and Control Groups).

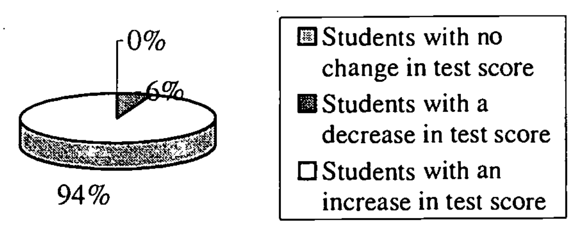
Teacher Two	Number of Students in the Control Group	Number of Students in the Experimental Group
Students with no change in test score	0	0
Students with a decrease in test score	11	1
Students with an increase in test score	7	16
Students with 0% > Score Increase ≥ 5%	3	5
Students with 5% > Score Increase ≥ 10%	1	5
Students with greater than 10% increase in the test score	3	6

Figure 22: Student Performance of Teacher Two – a comparison between pre- and post-test.

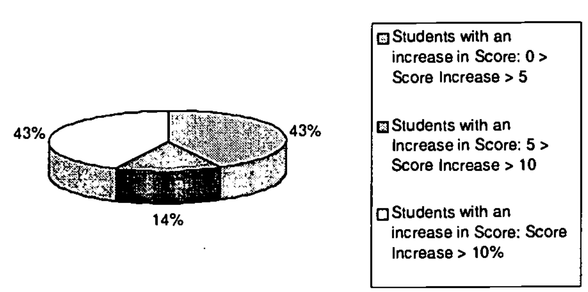
Control Group Student Performance (Teacher 2)



Experimental Group Student Performance (Teacher 2)



Increase in Score in Control Group



Increase in Score in Experimental Group

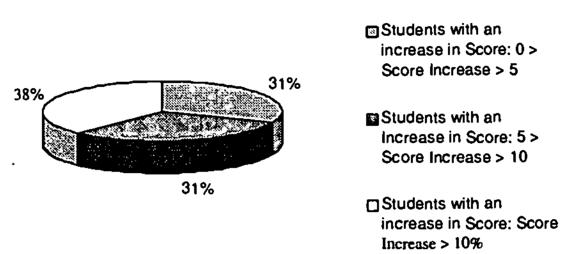


Figure 23: Student performance of Teacher Two (post-test minus pre-test score)

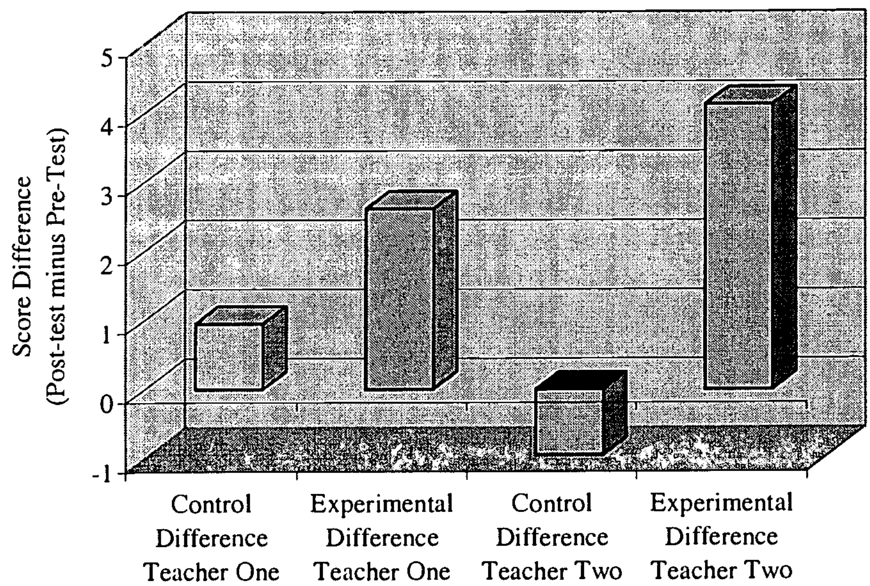


Figure 24: Average score difference -- post-test minus pre-test (Teacher One and Teacher Two)

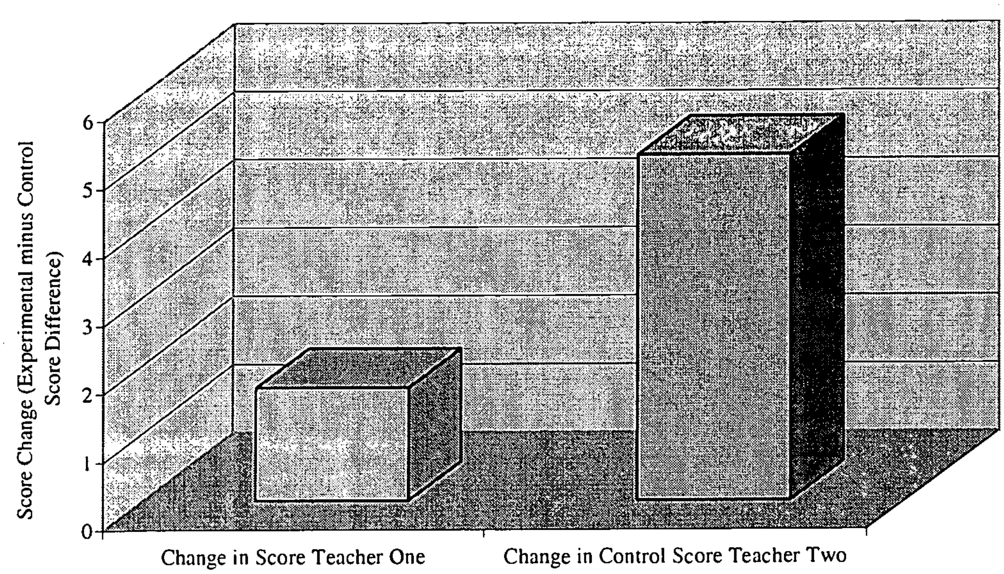


Figure 25: Test score change (experimental minus control test scores)

Student No.	Pre-Test Score	Post-Test Score	Difference	% Change
1	7	20	13	32.5
2	16	17	1	2.5
3	9	16	7	17.5
4	14	18	4	10
5	18	16	-2	-5
6	20	15	-5	-12.5
7	22	14	-8	-20
8	19	15	-4	-10
9	18	28	10	25
10	13	15	2	5
11	21	23	2	5
12	19	16	-3	-7.5
13	21	23	-2	-5
14	10	15	5	12.5
15	19	13	-6	-15
16	20	17	-3	-7.5
17	12	11	-1	-2.5
18	29	30	1	2.5
19	11	13	2	5
20	10	16	6	15
21	13	12	-1	-2.5
22	20	18	-2	-5
23	23	19	-4	-10
24	12	8	-4	-10
25	14	13	-1	-2.5
26	9	3	-6	-15
27	18	9	-9	-22.5
28	18	15	-3	-7.5
29	12	18	6	15
30	22	24	2	5
31	17	13	-4	-10
32	13	8	-5	-12.5
33	22	23	1	2.5
34	16	21	5	12.5
35	9	14	5	12.5
36	18	22	4	10
37	13	15	2	5
38	22	19	-3	-7.5
Mean	16.28947368	16.25714286	0.052631579	0.131578947
Median	17.5	16	-1	-2.5
Minimum	7	3	-9	-22.5
Maximum	29	30	13	32.5
Std. Dev.	4.991386607	5.42614974	4.937158437	12.34289609

Figure 26: Control Group (Teacher One + Teacher Two)

Student No.	Pre-Test Score	Post-Test Score	Difference	% Change
1	23	27	4	10
2	23	21	-2	-5
3	17	23	6	15
4	15	20	5	12.5
5	19	12	-7	-17.5
6	13	10	-3	-7.5
7	25	22	-3	-7.5
8	18	22	4	10
9	12	18	6	15
10	7	20	13	32.5
11	17	14	-3	-7.5
12	19	20	1	2.5
13	11	11	0	0
14	14	18	4	10
15	18	19	1	2.5
16	13	16	3	7.5
17	13	23	10	25
18	14	22	8	20
19	17	27	10	25
20	9	17	8	20
21	11	10	-1	-2.5
22	17	19	2	5
23	13	20	7	17.5
24	17	20	3	7.5
25	12	19	7	17.5
26	18	24	6	15
27	19	23	4	10
28	20	21	1	2.5
29	18	22	4	10
30	16	19	3	7.5
31	23	25	2	5
32	12	20	8	20
33	13	16	3	7.5
34	22	24	2	5
35	23	24	1	2.5
Mean	16.31428571	19.65714286	3.342857143	10.29411765
Median	17	20	3	7.5
Minimum	7	10	-7	-2.5
Maximum	25	27	13	25
Std. Dev.	4.390976232	4.358513354	4.249171941	7.597406678

Figure 27: Experimental Group (Teacher One +Teacher Two)

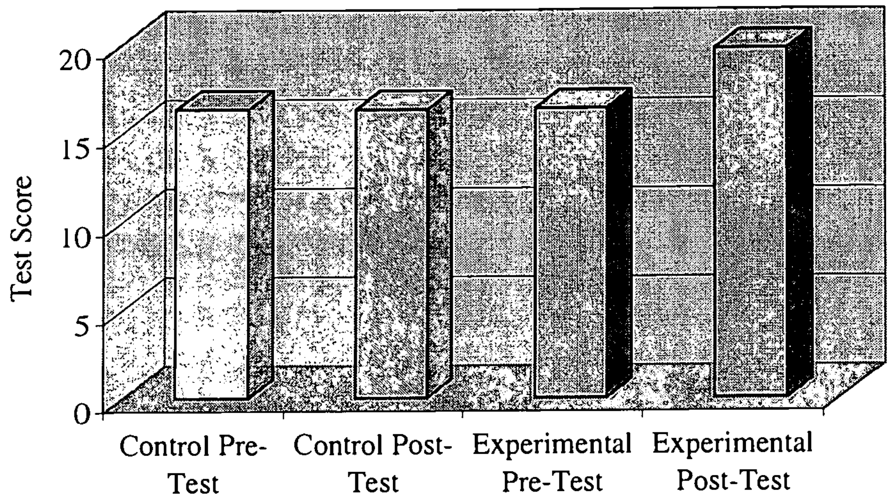


Figure 28: Combined test-scores (experimental and Control Groups).

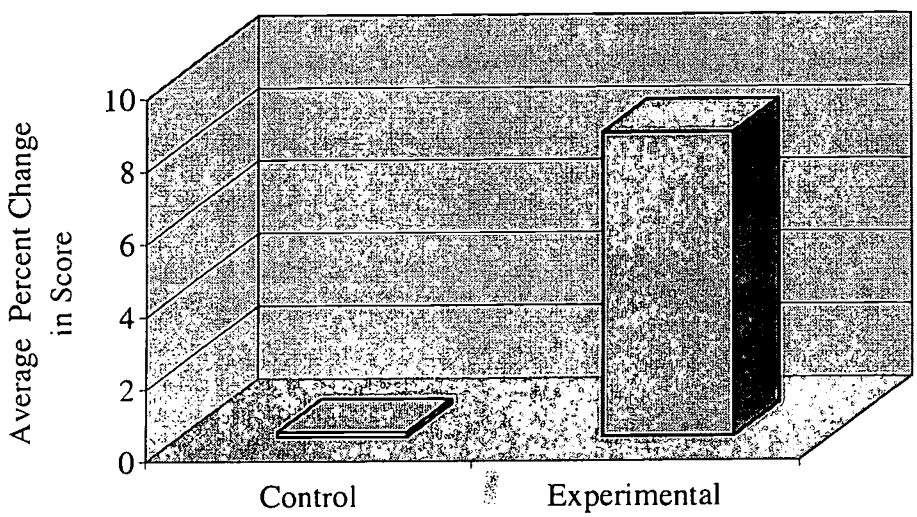


Figure 29: Score difference (post-test minus pre-test).

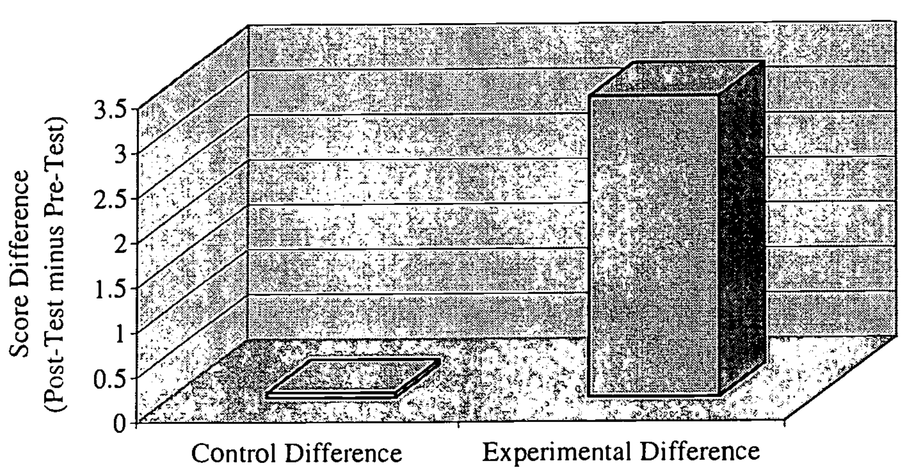


Figure 30: Score difference (post-test minus pre-test).

Teacher One + Teacher Two	Number of Students in the Control Group	Number of Students in the Experimental Group
Students with no change in test score	0	1
Students with a decrease in test score	18	6
Students with an increase in test score	18	28
Students with 0% > Score Increase ≥ 5%	8	7
Students with 5% > Score Increase ≥ 10%	2	9
Students with greater than 10% increase in the test score	8	12

Figure 31: Student performance of Teacher Two – a comparison between pre- and post-test.

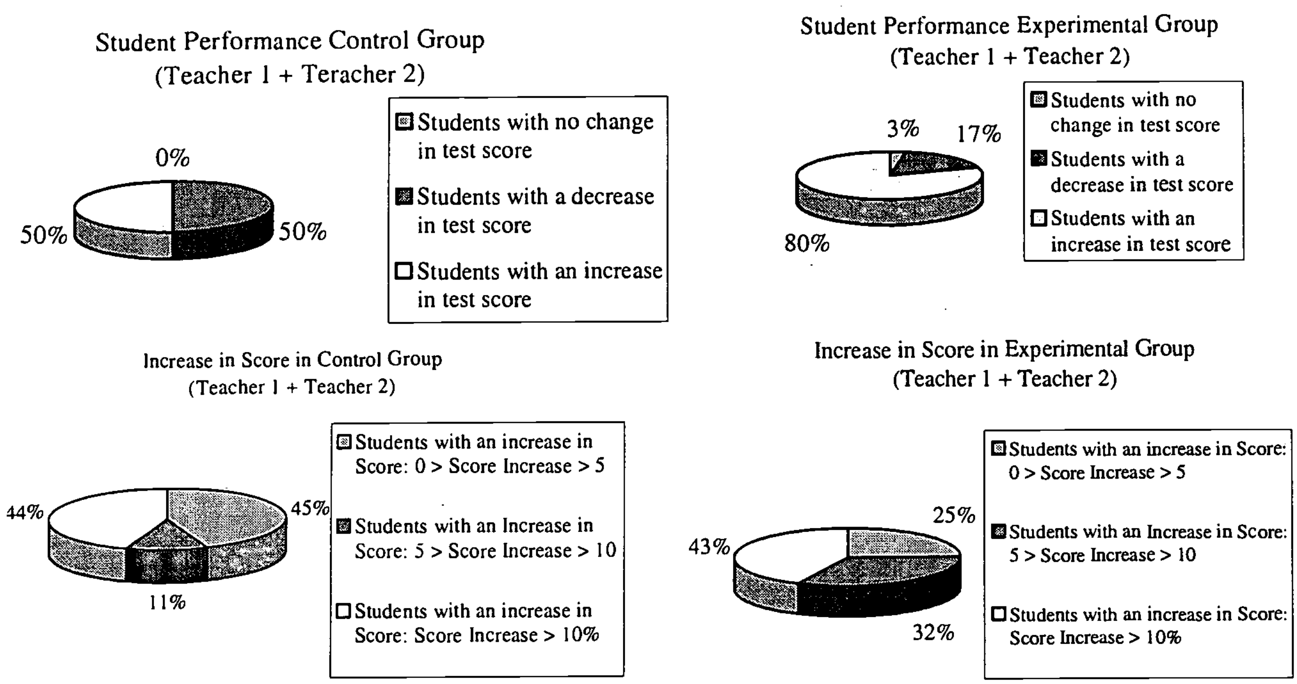


Figure 32: Student performance of Teacher One + Teacher Two (post-test minus pre-test score).

The experimental design for this study involved a randomized comparative experiment that randomly divided subjects into two groups and exposed each group to a different treatment. As a result, the two-sample test t-test was used to compare the differences of the exam scores of the experimental and the Control Groups. The two-sample *t*-Test of the results from this study indicates the p-value of 0.0025 ($p = 0.0025$), which corresponds to 99.5% confidence level. The t-test values were obtained by using the following equation.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

The following table identifies the variables and the corresponding number from the data.

Variable	Variable Identity	Numerical Data
$\bar{x}_1$	Mean value of experimental test score difference	3.3429
$\bar{x}_2$	Mean value of control test score difference	0.0526
s_1^2	Standard deviation value for the experimental test score difference	$(4.2492)^2$
s_2^2	Standard deviation value for the control test score difference	$(4.9371)^2$
n_1	Sample size (# of students) in the Experimental Group	35
n_2	Sample size (# of students) in the Control Group	38

The t value is equal to 3.0585. This value, plus the $n - 1$ degree of freedom which is equal to 34, correlates to $p = 0.0025$, having a confidence level of 99.5%.

As can be seen in Figures 14 and 16 (p. 25 and p. 26), Teacher One showed a pre-test average of 16.4 and a post-test average of 17.55 in the Control Group. This equates to a difference of 0.95 which is a 2.375% increase in average from pre- to post-test. Teacher Two experienced a pre-test average of 15.82 and a post-test average of 15 in the Control Group. This was a -0.94 or 2.36% decrease in the pre- to post-test average (Figures 19 and 21, p. 28 and p. 29). In the Experimental Group, Teacher One experienced a pre-test average of 16.17 and a post-test average of 18.78. This was a 2.61 or 6.52% increase from pre- to post-test (Figures 15 and 16, p. 26). Teacher Two demonstrated a pre-test average of 20 and a post-test average of 25.5. This equates to a 4.11 or 10.29% increase from pre- to post-test average (Figures 20 and 21, p. 29). Figure 24 (p. 31) summarizes the differences in pre- and post-test scores in both the control and Experimental Groups for both teachers. Teacher One encountered a 1.66 point difference between the Control and Experimental Group from pre- to post-test, while Teacher Two showed a 5.06 point difference (Figure 25, p. 31). In terms of the improvement on individuals, Teacher One saw 11 students in the Control Group improve; 5 increased their score by up to 5%, 1 increased their score between 5% and 10%, and 5 increased by more than 10%. In the Experimental Group, Teacher One saw 12 students improve their score; 2 increased their score by up to 5%, 4 increased their score between 5% and 10%, and 6 increased by more than 10% (Figures 17 and 18, p. 27). Teacher Two also showed an increase in 7 of the scores of the Control Group; 3 increased their score by up to 5%, 1 increased their score between 5% and 10%, and 3 increased by more than 10%. In the Experimental Group, 16 students improved their scores; 5 increased their score by up to 5%, 5 increased their score between 5% and 10%, and 6 increased by more than 10% (Figures 22 and 23, p. 30).

When combining the data from both Teachers 1 and 2, the following results were obtained. In the Control Group, scores averaged 16.29 for the pre-test and 16.26 for the post-test. This is a 0.053 point difference, which is an increase of 0.013% (Figure 26, p. 32). As for the Experimental Group, the average score for the pre-test was 16.31, while the post-test average was 19.66. This is a 3.34 point difference, which is an increase of 10.29% (Figure 27, p. 33). These results are summarized in Figures 28 through 30 (p. 34). Overall, in the Control Group, 18 students improved their scores; 8 increased their score by up to 5%, 2 increased their score between 5% and 10%, and 8 increased by more than 10%. The changes in the Experimental Group are much more dramatic. Twenty-eight students increased their scores; 7 increased their score by up to 5%, 9 increased their score between 5% and 10%, and 12 increased by more than 10% (Figures 31 and 32, p. 35).

The difference in the test score was obtained through subtracting the pre-test score from the post-test score. The average of the score difference for the both the Control and the Experimental Groups was obtained by adding the score difference of each student and dividing this sum by the total number of students. The average test-score difference between the control and the Experimental Groups for each teacher contained large standard deviations. For Teacher One, the average score difference in the Control Group was 0.95 ± 5.42 and in the Experimental Group was 2.61 ± 5.12. The standard deviation of 5.42 and 5.12 between the average score differences of 0.95 and 2.61 respectively, showed that the score difference between the Control and Experimental Groups was not significant. Rather, difference between the two groups was due to randomness of the data. For instance, according to the standard deviation value of 5.42, the average score difference of .92 can have the maximum value of 6.37, which is greater than 2.61. This overlaps the average score difference between the Experimental and Control Groups, thus making the results insignificant. For Teacher Two, the average score difference in Control Group was -0.94 ± 4.26 and in the Experimental Group was 4.11 ± 3.04. Like the Teacher One, the standard deviations for both averages of score difference showed that the difference between the score of the pre- and post-test scores was not distinct or significant.

Although the individual results of teachers did not indicate significance in the difference of test scores, the combined result of Teachers One and Two showed significance. The difference in test score between the Experimental and Control Groups through two-sample t-test showed that the results were significant and that this difference was not due to random events. This was due to the increase of the sample size, the higher n value, when the results of the two teachers were combined. The p-value obtained from the two-sample t-test, which is 0.0025, was less than the p-value of 0.05, the value at which significance is determined. Thus, based on the p-value of 0.0025 that has 99.5% confidence level, the average score difference between Control and Experimental Groups was distinct and significant.

Discussion

The intervention strategies used in this study proved to be effective. A randomized comparative experiment randomly divided students into two groups, the Control versus the Experimental Group, and exposed each group to a different treatment was conducted. Upon evaluating the data through a two-sample t-test, the results indicate significance in distinguishing Experimental and Control Group performance. The Experimental Group received the intervention strategy of test-taking skills, while the Control Group was not taught the test-taking skills. Each group was tested with the pre-test and post-test, and the difference in the score of the two test sample was used to measure the significance of the results. The two-sample t-test value of 3.0585 with 34 degrees of freedom corresponds to a p-value of 0.0025 and 99.5% confidence level. Thus, the test-taking strategies taught to the students had a significant effect upon student performance in the standardized norm-referenced exams. The intervention strategies were premised upon empowering the students to better utilize their knowledge and educational experience from their regular schooling versus teaching the test to the students. The strategies include the Question-Answer Relationship Strategy (QARS), Plus/Minus system, Multiple Choice Method, and True/False Methods. These strategies were selected and implemented to help students gain familiarity with the format of the standardized norm-referenced exam as well as the nature of the questions that these questions asked. Furthermore, these strategies allowed students to better manage their time while effectively and accurately assessing and responding to the questions. This aid to the students proved to be valuable, resulting in improved scores. For instance, for Teacher One, 66.7% of the Experimental students showed an improvement in their scores. For Teacher Two, 94.1% of the students showed an improvement in their scores. For Teacher One, two students increased their exam scores greater than or equal to 25%, while for Teacher Two, one student increased her score by 25%. Thus, teaching test-taking skills to students has a positive effect on student performance on the standardized norm-referenced exams.

The uses of the standardized norm-referenced exams to assess the performance of the students and schools have placed pressure on schools to improve their performances on these exams. This pressure has led many of the schools and teachers to align their curriculum around the contents of the exams, rather than teaching their usual curriculum, which is based upon a comprehensive overview of all disciplines. This study's purpose and findings shows how a regular curriculum can be implemented at the same time that students are taught necessary skills to effectively apply their knowledge on a standardized exam format. With little sacrifice of class time in implementing intervention strategies, students can show a significant improvement in their test scores. This effect counters the efforts of those teachers and schools who focus their curriculum on standardized exams.

All of the intervention methods proved beneficial to the students. All test-taking strategies were readily used by the students to analyze and respond to test questions. However, the Plus-Minus System appeared most interesting and popular to the students. This time-management, test-taking skill provided students with an effective means to better manage their time and allocate their effort efficiently to address most, if not all, of the questions on the standardized exam. Most students in the sample groups appeared to possess the necessary knowledge; they just needed need time to analyze all of the questions and access their knowledge to properly respond to questions. The Question-Answer Relationship Strategy (QARS) helped the students to better implement the Plus-Minus test-taking skills. This strategy helped students to better identify the nature or type of questions on the exams. As a result, students were able to better manage their time; they answered the questions that require less time first before they engaged in questions that required more time and effort. When encountering exam questions that were too difficult, the Multiple Choice Method helped students the most. The process of elimination (as well as recognizing absolute qualifiers), allowed the students to eliminate distracters in the choices and isolate those responses that were correct. The least beneficial strategy was the True/False Method. The nature of the standardized norm-referenced exam eliminates the need for this strategy; the standardized norm-referenced exams usually do not have many true/false questions. Although this method appeared to offer little benefit, the rest of the intervention methods proved to be highly effective, as is evident in the results of this study.

The biggest challenge in implementing the test-taking intervention strategy was finding and maintaining time for instruction and practice. In the midst of curriculum emphasis and pressure in a school with high academic drive, many difficulties were encountered, necessitating delays in practice of the test-taking skills. Firm discipline and accountability will be helpful in consistently implementing the intervention strategy. This may produce more reliable data, where student scores are higher and more consistent with a lower standard deviation. Thus, the significance or degree of confidence from a statistical analysis may also improve.

Conclusions

Project results clearly demonstrate that specific efforts to improve student performance on standardized norm-referenced tests can be beneficial; however, one sacrifice in implementing this intervention method was the loss of instructional time. Content emphasis was less in the Experimental Groups than the Control Groups; however, the class grade distribution between Experimental and Control Groups was approximately the same. Perhaps students in the Experimental Groups could have outperformed the students in the Control Group if additional instructional time was provided for these students. An additional study could be performed to address this concern; however, this result may provide an insight into the current findings. The students of the Experimental Groups, on average, may be more intelligent than the Control Group and the samples that were assumed to be random may have not been randomly established. As a result, the outcome might have possessed a systematic error. However, this study was conducted on the premise that counselors at the target school randomly place students in the appropriate levels based upon prior school performance. This should establish random samples so that a controlled experiment may be conducted for comparison.

REFERENCES

Angaran, J. (1999). Reflection in an age of assessment. Educational Leadership, 58(6), 71-72.

Baresic, J. & Gilman, D. (2001). How does the pendulum swing on standardized testing? Educational Digest, 66(5), 12-16

Bond, L. A. (1995). Critical issue: Rethinking assessment and its role in supporting educational reform. North Central Regional Educational Reform [On-Line]. Available: http://www.ncrel.org/sdrs/areas/issues/ methods/assment/as700.htm

Clarke, K. (1993). Why students are feeling so testy. U.S. Catholic, 65(7), 27.

Frierson, H. T. Jr (1984). Assessing intervention effects on repeated standardized test performance: Examining raw percentage scores. (Report No. HE018752). North Carolina. (ERIC Document Reproduction Service No. ED 262683).

Hillard, A. G. (2000). Excellence in education versus high-stakes standardized testing. Journal of Teacher Education, 51(4), 293-304.

Illinois School Report Card, 2001.

Jongsma, E. A. & Warshauer, E. (1975). The effects of instruction in test-taking skills upon student performance on standardized achievement tests. Final report. (Report No. TM004887). (ERIC Document Reproduction Service No. ED 114 408).

Lord, M. (2000, April 3). High-stakes testing: It's backlash time. U.S. News & World Report, 128(13), 54.

McLean, J. E. & Lockwood, R. E. (1996). Why we assess student perfromance – and how: The competing measures of student performance. Thousand Oaks, CA: Corwin Press, Inc.

Mehrens, W. A. (1989). Preparing students to take standardized achievement tests. (Report No. TM014145). District of Columbia. (ERIC Document Reproduction Service No. EDO-TM-89-2).

Popham, J. W. (2000a). Modern educational measurement; practical guidelines for educational leaders. Needham, MA: A Pearson Education Company.

Popham, J. W. (2000b). Selling higher test scores. American Prospect, 11(22), 29-31.

Popham, J. W. (2000c, March). Teaching the test? Educational Leadership, 58(6),16-20.

Popham, J. W. (1999). Why standardized tests don't measure educational quality. Educational Leadership, 58(6), 8-15.

Sacks, P. (1997). Standardized testing: Meritocracy's crooked yardstick. Change, 29, 24-31.

White, K., et al. An evaluation of training in standardized achievement test taking and administration. final report of the 1981-82 Utah state refinement to the ESEA Title I Evaluation and Reporting System. (Report No. TM840033). Utah. (ERIC Document Reproduction Service No. ED 238931).

Whitefield, B. & Santee, P. (1994). Using standardized test results as an instructional guide, The Clearing House, 67(6) 320-322.

APPENDICES

Appendix A
Pre-Test

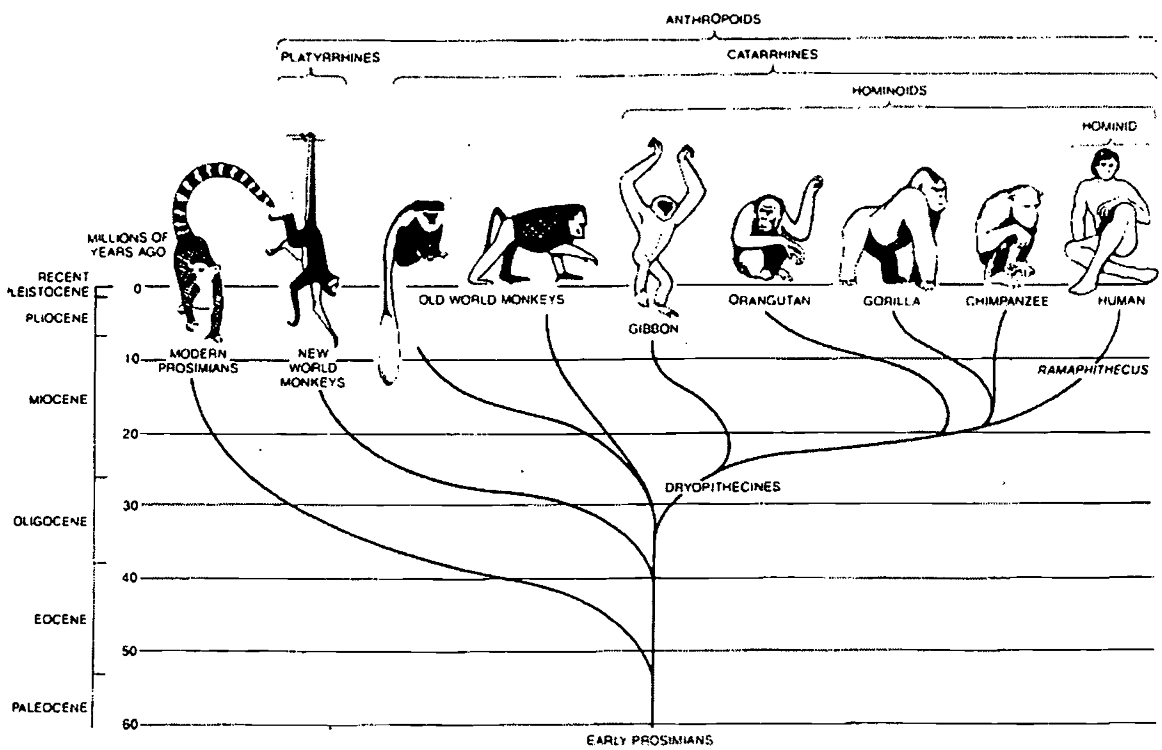
Model Science Reasoning Test

35 Minutes—40 Questions

DIRECTIONS: This test consists of several distinct passages. Each passage is followed by a number of multiple-choice questions based on the passage. Study the passage, and then select the best answer to each question. You are allowed to reread the passage. Record your answer by blackening the appropriate space on the answer sheet.

Passage I

The diagram below shows the probable evolutionary relationships of the primates.



From Helena Curtis. *Biology*. 4 ed. Worth Publishers. New York. 1983.

GO ON TO THE NEXT PAGE.

Appendix A continued

1. The most recent common ancestor of human beings and the gorilla lived about:
   A. 10 million years ago.
   B. 19 million years ago.
   C. 25 million years ago.
   D. 33 million years ago.

2. The diagram proposes that the largest group that descended from Dryopithecines is:
   F. the gibbons.
   G. The Catarrhines.
   H. the Hominoids.
   J. the great apes.

3. Two organisms are sometimes said to be closely related if they have a recent common ancestor. By this criterion, which of the following pairs are most closely related?
   A. Old World and New World monkeys
   B. Modem prosimians and New World monkeys
   C. Human and orangutan
   D. Orangutan and gibbon

4. What does the diagram tell us about the relationships of the lemurs, which are the modern prosimians?
   F. Lemurs have existed unchanged for 60 million years.
   G. Lemurs are the ancestors of all primates.
   H. Of all primates, lemurs are most similar to the ancestors of the whole group.
   J. Lemurs have evolved to a greater extent than any of the other primates.

5. *Australopithecus* is a genus of primates. more recent and humanlike than *Ramapithecus*. Of the taxonomic groups named on the chart, which is the smallest to which *Australopithecus* belongs?
   A. Anthropoids
   B. Hominoids
   C. Catarrhines
   D. Hominids

Passage II

Experiments are done in a stream to determine how the size of particles eroded and deposited is affected by the velocity of the water.

Experiment 1

To study erosion, sediments composed of particles of various sizes are placed in the bottom of a stream at many different points. At each point, the velocity of the stream is measured. The minimum velocity needed to lift the particles off the stream bottom is determined for each size particle.

Material	Particle size (mm)	Minimum velocity (cm/s)
Clay	0.001	700
Silt	0.01	180
Fine sand	0.1	60
Coarse sand	1	90
Pebbles	10	210
Cobbles	100	600

Experiment 2

To study deposition, the sediments are dropped into the stream from the surface at various points in the stream. The minimum stream velocity that will prevent the particles from sinking to the bottom of the stream is measured for each size particle.

Material	Particle size (mm)	Minimum velocity (cm/s)
Clay	0.001	20
Silt	0.01	20
Fine sand	0.1	20
Coarse sand	1	30
Pebbles	10	100
Cobbles	100	300

6. A landslide dumps a mixture of particles of all sizes into a stream flowing at 40 cm/s. What kinds of material will be deposited in the stream bed?
   F. All sizes of particles
   G. Clay, silt, sand, and pebbles
   H. None; all will wash away
   J. Pebbles and cobbles

7. A newly formed stream starts to flow at 30 cm/s over a land area where the soil contains particles of all sizes. After a number of years, what kinds of particles will remain in the stream bed?
   A. All kinds
   B. Clay, silt, and fine sand
   C. Coarse sand, pebbles, and cobbles
   D. Clay and cobbles only

8. At a construction site, a quantity of coarse sand finds its way into a stream flowing at 40 cm/s. What happens to it?
   F. It falls to the bottom of the stream and stays there.
   G. It is carried away by the stream and falls to the bottom if the stream speeds up to 90 cm/s.

BEST COPY AVAILABLE

GO ON TO THE NEXT PAGE.

H. It is carried away by the stream and falls to the bottom if the stream slows down to 25 cm/s.
J. It falls to the bottom of the stream and is then picked up and carried into quieter water.

9. A stream flows at 130 cm/s over soil containing a mixture of all kinds of particles. What kinds of particles will drop to the bottom at a point where the velocity has slowed down to 25 cm/s?
   A. Pebbles and cobbles
   B. Coarse and fine sand only
   C. Coarse sand only
   D. Clay, silt, pebbles, and cobbles

10. Which of the following hypotheses might be advanced to account for the fact that a very large stream velocity is required to lift clay off the bottom of the river?
   F. Clay particles clump together to form lumps the size of cobbles.
   G. Clay particles are much larger than sand grains.
   H. Rivers always slow down when they flow over clay.
   J. Clay particles have a high density, so they drop to the bottom very easily.

11. If similar experiments were done with other materials in other rivers, which of the following outcomes would always be found?
   A. The velocity figures in Experiment 1 would always be larger than the corresponding values in Experiment 2.
   B. The velocity figures would never be larger than 600 cm/s in either experiment.
   C. The velocity figures would never be smaller than 20 cm/s in either experiment.
   D. Medium-size particles are always picked up by the slowest part of the river and carried furthest.

GO ON TO THE NEXT PAGE.

Appendix A continued

Passage III

The rate of oxygen absorption in fruits is a measure of the ripening process. The graphs below show how this rate is affected, in bananas and avocados, by exposure to various concentrations of the gas ethylene (ppm is parts per million).

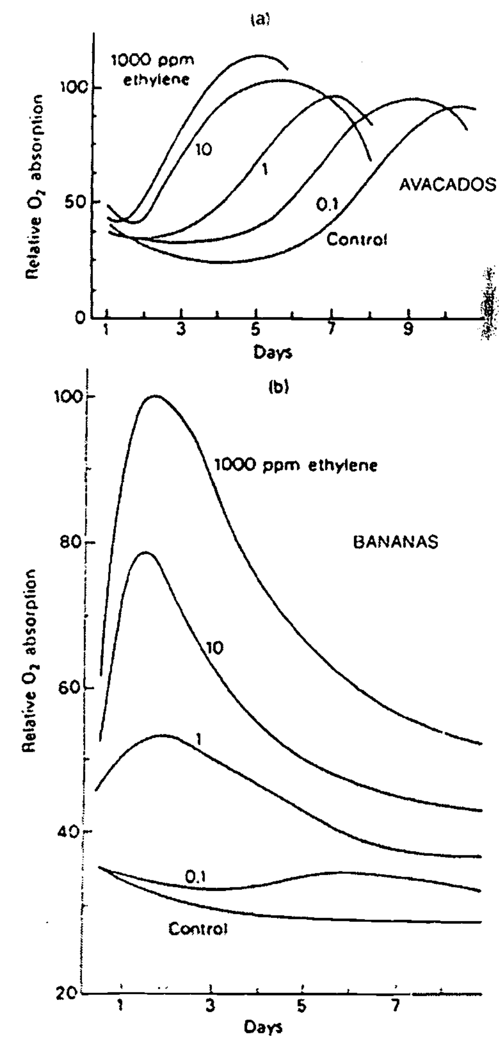


Figure 5.17 Oxygen uptake by developing fruits, and the effect of ethylene gas upon the normal respiratory trends. (a) For fruits which normally show the respiratory climacteric phenomenon (e.g., avacado pear). (b) For non-climacteric fruits such as banana.
(From *The Biochemistry and Physiology of Plant Growth Hormones*, McGraw Hill, 1971.)

12. If the fruits are not exposed to ethylene:
    - **F.** after the first few days, the rate of oxygen absorption is greater in avocados than in bananas.
    - **G.** absorption of oxygen decreases steadily in both fruits.
    - **H.** in both fruits, oxygen absorption rises to a peak and then drops off.
    - **J.** in avocados, oxygen absorption increases greatly after the first few days.

13. The effect of increasing the exposure to ethylene is to:
    - **A.** increase oxygen uptake in both fruits.
    - **B.** speed up oxygen uptake in both fruits.
    - **C.** increase oxygen uptake for bananas only.
    - **D.** speed up oxygen uptake for avocados only.

14. In what circumstance would both fruits reach about half of their maximum level of oxygen absorption?
    - **F.** On the fourth day with 1 part per million of ethylene.
    - **G.** On the second day with 10 parts per million of ethylene.
    - **H.** On the sixth day with 10 parts per million of ethylene.
    - **J.** On the first day with 50 parts per million of ethylene.

15. A reasonable assumption from the graphs is that:
    - **A.** ethylene causes bananas to greatly increase the total amount of oxygen they absorb in ripening.
    - **B.** concentrations of ethylene above 10 parts per million have little effect on the ripening time of avocados.
    - **C.** the total amount of oxygen absorbed by avocados in ripening is not affected by the ethylene concentration.
    - **D.** bananas are affected more strongly than avocados by the presence of ethylene.

16. The information provided suggests that if a fruit wholesaler wanted to use ethylene to control the ripening of his produce, it would be necessary to:
    - **F.** keep the concentration of ethylene below 10 ppm.
    - **G.** separate the different kinds of fruit into different containers.
    - **H.** test each batch separately with no ethylene.
    - **J.** get the fruit to market promptly after exposure to ethylene.

GO ON TO THE NEXT PAGE.

Appendix A continued

Passage IV

A chemist is investigating the effect of various kinds and amounts of solutes on the boiling point of a solution.

Experiment 1

Solutions are made of various amounts of glucose dissolved in 1 liter of water, and the boiling point of each solution is measured.

Glucose concentration (g/L)	Boiling point (° C)
0	100.0
100	100.3
200	100.6
300	100.9
400	101.2
500	101.5

Experiment 2

Solutions are made of various solutes (substances dissolved) in water, all with a concentration of 300 g/L. The boiling point of each solution is measured.

Substance	Molecular weight	Boiling point (° C)
Acetaldehyde	44	103.5
Glycerol	92	101.7
Glucose	180	100.9
Sucrose	342	100.5

Experiment 3

Solutions are made of various solutes dissolved in benzene (the solvent), which boils at 80.1° C. All concentrations are 300 g/L.

Solute	Molecular weight	Boiling point (° C)
Butyric acid	88	89.8
Triethylamine	101	88.6
Naphthalene	178	84.8
Cholesterol	387	82.3

17. If 200 g of glucose are dissolved in 500 mL of water, the boiling point of the solution will be:
    A. 100.3° C.
    B. 100.6° C.
    C. 101.2° C.
    D. 101.6° C.

18. Three hundred grams of a substance with a molecular weight of 65 are dissolved in 1 liter of water. The boiling point of the solution will be about:
    F. 103.5° C.
    G. 102.4° C.
    H. 101.7° C.
    J. 100.3° C.

19. For a given concentration and molecular weight of solute, how does the elevation of the boiling point depend on the kind of solvent?
    A. It is the same for all solvents.
    B. It is the same for water and for benzene.
    C. It is more for water than for benzene.
    D. It is more for benzene than for water.

20. The experiments indicate that the boiling point elevation is directly proportional to:
    F. the molecular weight of the solute.
    G. the number of molecules of solute per gram of solvent.
    H. the boiling point of the solvent used.
    J. a combination of solute concentration and molecular weight.

21. In Experiment 3, the solution of butyric acid in benzene boiled at a higher temperature than cholesterol in benzene. A possible explanation is:
    A. the molecules of cholesterol are larger, so they lower the boiling point of the benzene.
    B. the mass of cholesterol in solution was larger than the mass of butyric acid.
    C. cholesterol reacts chemically with benzene, but butyric acid does not.
    D. the butyric acid solution contains more molecules of solute than the cholesterol solution.

22. In trying to determine the nature of a newly discovered substance, a chemist might use experiments of this kind to discover its:
    F. chemical formula.
    G. concentration.
    H. molecular weight.
    J. boiling point.

GO ON TO THE NEXT PAGE.

Appendix A continued

Passage V

The diagram below represents the forms of energy consumption in an old building, in the design for a new building, and in the actual new building after it was built.

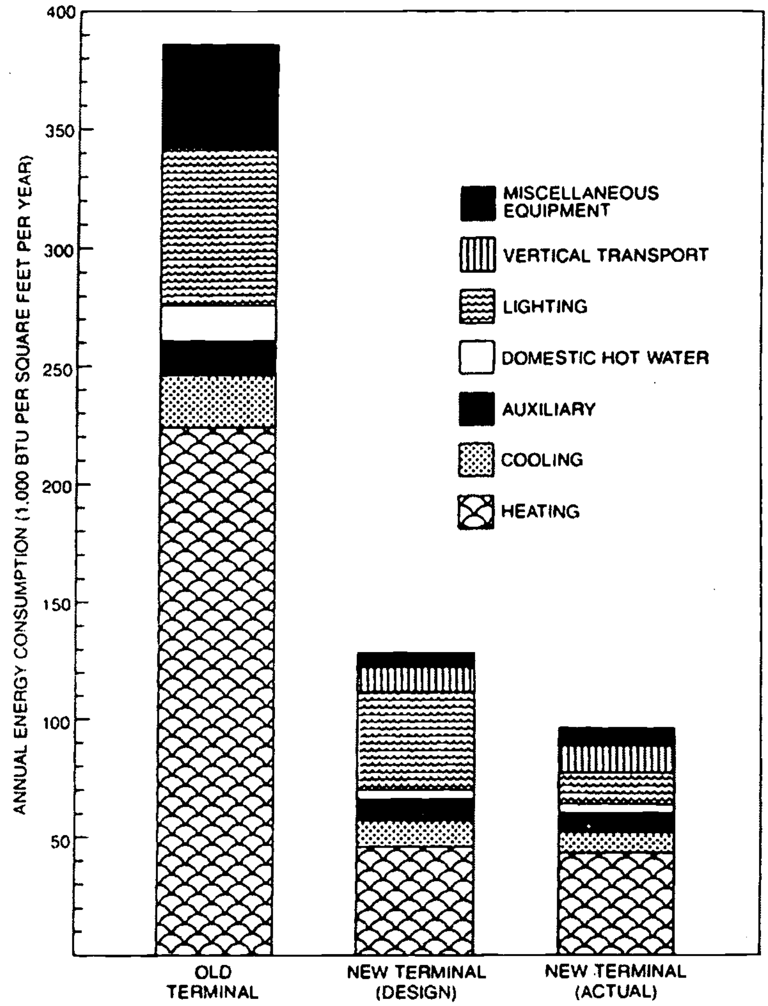


Illustration by Gabor Kiss from "Energy-efficient Buildings," by Arthur H. Rosenfeld and David Hafemeister.

GO ON TO THE NEXT PAGE.

23. How effective was the new design in improving the efficiency of energy use?
    - A. It reduced the amount of energy used by about 15%.
    - B. It was a great improvement, but did not accomplish all that was expected of it.
    - C. It was so good that the new building actually performed better than expected.
    - D. It was unnecessary, since the new building was so much better than the old one.

24. What service was provided in the new building that was not available in the old one?
    - F. Elevators
    - G. Fluorescent lighting
    - H. Air conditioning
    - J. Electronic energy control

25. The difference between the new design and the performance of the actual building might have been due to miscalculation of the energy saving provided by:
    - A. insulation.
    - B. fluorescent lighting.
    - C. air conditioning.
    - D. improved boilers.

26. What approach to future problems of design of heating efficiency is suggested by these results?
    - F. Since heating is the largest part of the energy cost even in the new building, this design did little to improve heating efficiency.
    - G. Design emphasis should be placed on heating because this is the area where the major savings can be made.
    - H. Better designs must be sought because this one did not produce a result better than the one achieved in the actual new building.
    - J. Little additional research is needed because this design provided the maximum possible saving of heating fuel.

27. Which of the following questions had to be answered before a decision was made as to whether or not to use the new design in constructing the new building?
    - A. Should the new building be equipped with air conditioning?
    - B. Will the new design improve the efficiency of hot water heating?
    - C. Does the actual new building accurately reflect the gains suggested by the design?
    - D. Does the new design cost so much more than the old that the fuel saving would not make up the difference?

Passage VI

The purpose of this experiment is to study the rate at which the eyes of guppies become light-adapted.

Experiment 1

Three guppies are kept in daylight conditions at 24° C. They are fed 50 water fleas (*Daphnia*) once a day. The number of *Daphnia* captured in 5 minutes is counted. In six trials, the following results are obtained:

Number of *Daphnia* captured: 35 32 32 36 34 35

Experiment 2

The guppies are kept in the dark at 24° C for a full day. Then a light is turned on. After a measured time delay, 50 *Daphnia* are added to the tank, and the number captured in 5 minutes is counted. The experiment is repeated with various time delays.

Time delay (minutes)	Number captured in 5 minutes
2	0
4	18
6	24
8	33
10	33
12	32

Experiment 3

A similar experiment is done with the guppies, which have been kept at various temperatures. A uniform time delay of 8 minutes is used before the food is added.

Temperature (° C)	Number captured in 5 minutes
15	12
18	18
21	26
24	35
27	34
30	35
33	22
36	9

28. What assumption underlies the design of these experiments?
    - F. Guppies are most active when illumination is high.
    - G. The ability of guppies to find food depends on their ability to see it.
    - H. Temperature affects the ability of guppies to find food.
    - J. The eyes of guppies are just like the eyes of people.

BEST COPY AVAILABLE

GO ON TO THE NEXT PAGE.

29. What was the purpose of Experiment 1?
    A. To establish a criterion as to when the guppies' eyes are light-adapted
    B. To control any possible effect of temperature
    C. To condition the guppies to respond to the presence of *Daphnia*
    D. To keep the guppies in healthy condition

30. What was the purpose of the time delays in Experiment 2?
    F. To see how long it would take for the guppies to find their food
    G. To allow for differences between guppies in their feeding ability
    H. To find out how long it takes for the eyes of the guppies to become completely light-adapted
    J. To measure the time rate at which guppies find their food under standard conditions

31. What evidence is there that guppies depend solely on their eyesight to find food?
    A. Experiment 3 shows that they cannot see well at low temperatures.
    B. Experiment 2 shows that, in the first 2 minutes after being kept in the dark, they cannot find any.
    C. Experiment 1 shows that they find food very efficiently in daylight.
    D. Experiment 2 shows that the rate at which they find food diminishes as the food supply dwindles.

32. A time delay of 8 minutes was selected for Experiment 3 because this is the amount of time required for:
    F. the guppies to consume most of the food.
    G. the *Daphnia* to become adapted to the tank.
    H. the guppies' eyes to become light-adapted at 24° C.
    J. the water in the tank to reach a steady temperature.

33. Which hypothesis could NOT account for the results of Experiment 3?
    A. Light adaptation is delayed at unusually high temperatures.
    B. Guppies are damaged at very high temperatures, and are thus unable to feed well.
    C. At high temperatures, *Daphnia* become immobile and more difficult to find.
    D. The rate of adaptation to light increases uniformly with temperature.

Passage VII

Two scientists disagree about an important point in the theory of evolution.

Scientist 1

Darwin's theory of evolution by natural selection has a serious flaw. It cannot account for the origin of completely new organs. Consider, for example, the wings of insects. These originate in the emhryo as outgrowths of the external skeleton of the thorax. How could such outgrowths have evolved? According to all we know about evolution, they must have started as tiny outgrowths that grew bigger generation by generation until they became functional wings. However, a small, incipient wing is useless; an insect cannot fly with 1% of a wing. Since those tiny wing buds would have no use, they could not contribute to the ability of the insects to survive, and thus could not contribute to the evolution of wings. It has been suggested that these small outgrowths were originally used as gliding surfaces, like the skin flaps of a flying squirrel. This does not solve the problem, however, for a good-sized surface would be needed for gliding; 1% of a surface is as useless for gliding as for flying. Somehow, such a surface must have gradually developed by some means other than natural selection. Natural selection could improve a gliding surface, and turn it into a functional, flapping wing, but it could not invent the surface in the first place. The insect wing is an example of preadaptation, the development of a new structure that subsequently evolves into something useful. Evolution gives us many such examples. Preadaptation surely exists, but it cannot be explained by the theory of natural selection. Explanation of the origin of new, useful structures is an unfinished task of biological science.

Scientist 2

The problem raised by Scientist 1 is not new. Darwin provided the answer a century ago, although at that time there was no real evidence. It all hinges on the meaning of the term *preadaptation*. It does not mean that a new feature arises and then develops for many evolutionary years while it has no function. What it does mean is that an existing structure, developed for one function, can acquire a totally different use. This is undoubtedly how insect wings evolved. Recent studies show that the wings of insects have an important function other than flight: they are temperature-regulating structures. Many insects control their body temperatures by sitting in the sun to warm up. In ancient, primitive insects, an expansion of the thorax exposed more surface. This expansion was promoted by selection, since it allowed for faster and more efficient warmup. Lateral projections provided additional surfaces for temperature regulation. When these projections grew big enough, they acquired a new function; they became useful as gliding surfaces. The ability to glide has developed many times, and it has

GO ON TO THE NEXT PAGE.

Appendix A continued

sometimes evolved into real, controlled flight. Selection converted the clumsy gliding surfaces of ancient insects into the complex and effective insect wings of today. Preadaptation is simply a shift of function, and is perfectly incorporated into the theory of evolution by natural selection.

34. One of the scientists, but not the other, would agree that:
    - F. natural selection results in improved functioning of insect wings.
    - G. natural selection can explain the origin of the wings of insects.
    - H. insects use their wings to help in temperature control.
    - J. insect wings evolved from surfaces previously used for gliding.

35. What evidence does Scientist 2 advance to support her argument?
    - A. Tests show that insects use their wings for temperature control.
    - B. Darwin provided a competent explanation for preadaptation.
    - C. In the development of an insect, the wings arise as outgrowths of the thoracic skeleton.
    - D. Many insects use projections from the thorax for gliding.

36. According to Scientist 2, what principle has Scientist 1 overlooked in his argument?
    - F. Natural selection can greatly improve the functioning of an organ.
    - G. In the evolution of many different flying organisms, gliding always precedes real flying.
    - H. New organs can arise even if they have no function at first.
    - J. In the course of evolution, an organ can acquire a new function.

37. Which of the following discoveries would support the hypothesis of Scientist 2?
    - A. Fossils of insects with primitive wings capable of gliding, but not of flying
    - B. A living species of insect that can glide, but cannot fly
    - C. A fossil insect with small lateral projections from the thorax
    - D. A living insect in which juvenile stages have short wings used only for temperature regulation

38. How should the argument of Scientist 1 be evaluated in terms of its contribution to science?
    - F. It is a foolish argument because the question was settled a long time ago.
    - G. It is useful to raise the question because it challenges scientists to find evidence for shifts of function.
    - H. It wastes valuable time because it has long been established that evolution results from natural selection.
    - J. It challenges scientists to find an answer to a perplexing problem for which there is no satisfactory answer.

39. What would the two scientists expect to find in the fossil record?
    - A. Scientist 1, but not Scientist 2, would expect to find that functional gliding surfaces appeared suddenly.
    - B. Both scientists would expect to find gradual evolution of wings, starting with small projections.
    - C. Scientist 2, but not Scientist 1, would expect to find gradual evolution, starting with small projections.
    - D. Both scientists would expect to find that there are no fossil insects without some sort of gliding surface.

40. In a recent series of experiments, paper models of insects were exposed to sunshine and to aerodynamic tests in a wind tunnel. Size and wing length were varied to test the concept of:
    - F. natural selection.
    - G. shift of function.
    - H. preadaptation.
    - J. evolution.

END OF TEST 4
STOP! DO NOT TURN THE PAGE UNTIL TOLD TO DO SO.

Appendix B
Post-Test

Model Science Reasoning Test

35 Minutes—40 Questions

DIRECTIONS: This test consists of several distinct passages. Each passage is followed by a number of multiple-choice questions based on the passage. Study the passage, and then select the best answer to each question. You are allowed to reread the passage. Record your answer by blackening the appropriate space on the answer sheet.

Passage I

The diagram below summarizes the chemical processes that produce acid rain.

NO is nitric oxide.
NO_2 is nitrogen dioxide.
SO_2 is sulfur dioxide.
SO_3 is sulfur trioxide.
CO is carbon monoxide.
CO_2 is carbon dioxide.

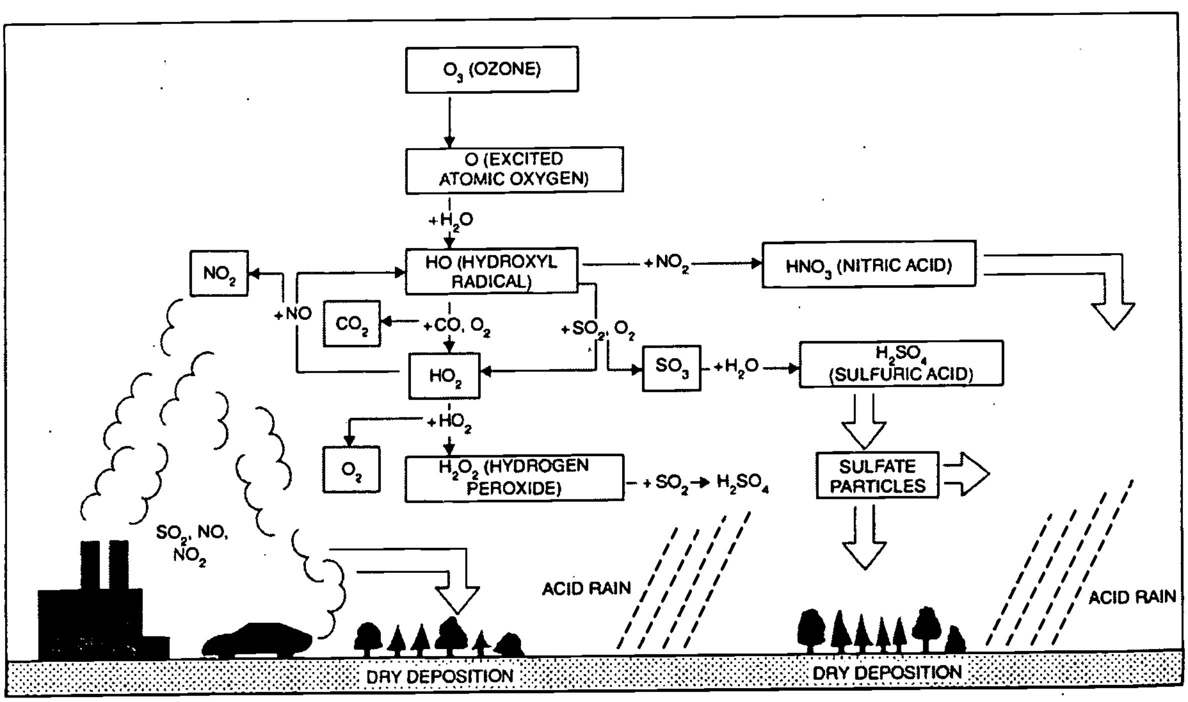


Illustration by Bob Conrad from "The Challenge of Acid Rain," by Volker A. Mohnen. Copyright © August 1988 by *Scientific American, Inc.*

Appendix B continued

1. What is in smokestack gases and automobile exhaust that ultimately produces acid rain?
   A. Ozone
   B. Carbon dioxide
   C. Oxides of sulfur and nitrogen
   D. Nitric and sulfuric acids

2. Acid rain forms as a result of the reaction of hydrogen peroxide with:
   F. sulfur dioxide.
   G. water.
   H. nitrogen dioxide.
   J. sulfuric acid.

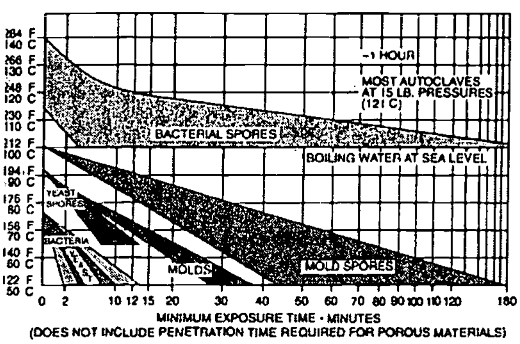
3. Carbon monoxide is removed from the air when it reacts with:
   A. nitric oxide to form nitric acid.
   B. sulfur dioxide to form sulfuric acid.
   C. oxygen to form water.
   D. oxygen to form carbon dioxide.

4. What gas combines directly with atmospheric water to form a polluting acid?
   F. Nitrogen dioxide
   G. Sulfur trioxide
   H. Sulfur dioxide
   J. Ozone

5. What substance that makes an important contribution to acid rain does not originate in the oxides of the exhaust gases?
   A. Sulfuric acid
   B. Ozone
   C. Nitrogen dioxide
   D. Sulfur trioxide

Passage II

In the sterilizing process, instruments and cultures are exposed to high temperatures for a definite length of time. The diagram below displays the combinations of temperature and time required to kill various kinds of microorganisms. The six graph areas represent the living stages of bacteria, yeasts, and molds, and the spore stages of these kinds of organisms.



6. The kind of microorganism that is most difficult to kill is:
   F. mold spores.
   G. bacterial spores.
   H. yeasts.
   J. yeast spores.

7. If a laboratory technician keeps instruments in boiling water for 3 hours, the purpose of the procedure is to kill:
   A. mold spores
   B. bacterial spores.
   C. all spores.
   D. all organisms.

8. What procedure could be used to kill off mold spores in a culture, but leave the yeast spores still viable?
   F. Hold the culture at 80°C for 20 minutes.
   G. Keep the culture at 90°C for 8 minutes and then at 85°C for another 5 minutes.
   H. Keep the culture at 70°C for 10 minutes.
   J. No combination of time and temperature can do this.

9. The chart suggests that, by controlling time and temperature, a technician might be able to:
   A. kill off bacterial spores while leaving live bacteria viable.
   B. kill all bacterial spores without destroying all the mold spores.
   C. kill off certain kinds of bacterial spores and leave other kinds still viable.
   D. destroy all living bacteria without killing off the living yeasts.

10. What general biological rule might be suggested by the contents of this graph?
    F. Microorganisms form spores to enable them to survive all kinds of unfavorable conditions.
    G. Molds are more sensitive than bacteria to temperature.
    H. Spore formation in microorganisms is a mechanism that protects the species against high temperatures.
    J. Spores are a vital mechanism for the reproduction of certain microorganisms.

BEST COPY AVAILABLE

GO ON TO THE NEXT PAGE.

Appendix B continued

Passage III

Experiments are done to study some of the factors that determine the rate of a reaction. When sulfuric acid acts on potassium iodate, elemental iodine is released. Iodine signals its presence by turning starch blue.

Experiment 1

A test solution is made of sulfuric acid and soluble starch in water. If potassium iodate is added, iodine accumulates at some definite rate. When the iodine reaches a certain concentration, the solution suddenly turns blue. Various concentrations of potassium iodate solution are used, and the time required for the mixture to turn blue is measured.

Potassium iodate concentration (%)	Time (seconds)
10	18
9	20
8	22
7	24
6	26
5	29
4	32

Experiment 2

To determine the effect of temperature on reaction rate, a 5% solution of potassium iodate is added to the test solution at various temperatures.

Temperature (° C)	Time (seconds)
5	36
15	31
25	27
35	24
45	22

11. Starch was added to the solution because:
    A. it speeds the reaction that produces iodine.
    B. it provides a test for the presence of elemental iodine.
    C. it slows down the reaction so that the time becomes easily measurable.
    D. it prevents the sulfuric acid from destroying the potassium iodate.

12. Experiment 1 shows that:
    F. elemental iodine turns starch blue.
    G. at higher iodate concentration, iodine is liberated more quickly.
    H. the rate of the reaction depends on the concentration of sulfuric acid used.
    J. the release of elemental iodine occurs suddenly.

13. Experiment 2 is an example of a general rule that:
    A. higher concentrations speed reactions.
    B. higher concentrations slow down reactions.
    C. higher temperatures speed reactions.
    D. higher temperatures slow down reactions.

14. Experiment 1 was done at a temperature of about:
    F. 10° C
    G. 20° C
    H. 30° C
    J. 40° C

15. By studying the results of this experiment, what can be concluded as to the time the reaction would take at a temperature of - 15° C?
    A. It would take about 48 seconds.
    B. It would take longer than 36 seconds, but it is impossible to predict how long.
    C. It is not possible to make any prediction because the results of the experiment are too scattered.
    D. It might take a long time, or the whole thing might freeze and stop the reaction.

16. About how long would it take for the starch to turn blue if a 10% solution of potassium iodate was used at 45° C?
    F. 15 seconds
    G. 18 seconds
    H. 22 seconds
    J. 29 seconds

GO ON TO THE NEXT PAGE.

Appendix B continued

Passage IV

The graphs below represent the percentages of fat and of water in the human body, by age and sex.

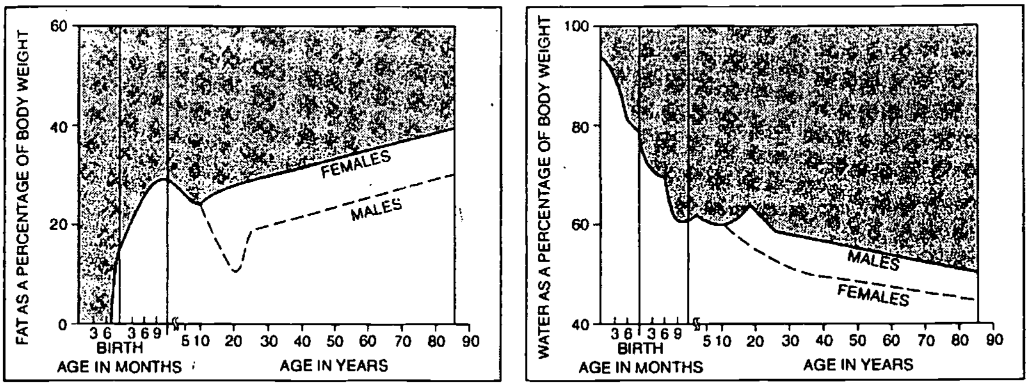


Illustration by Bob Conrad from "Fatness and Fertility," by Rose E. Frisch. Copyright © March 1988 by *Scientific American, Inc.*

17. During the adolescent years, the most notable change is:
- **A.** a decrease in fat percentage for boys.
- **B.** an increase in fat percentage for girls.
- **C.** a decrease in water percentage for boys.
- **D.** an increase in water percentage for girls.

18. The percent of fat in the body increases most rapidly during:
- **F.** middle age.
- **G.** adolescence.
- **H.** babyhood.
- **J.** the prenatal period.

19. At age 60, the amount of water in the body of a 150-pound man is:
- **A.** the same as that in a 150-pound woman.
- **B.** twice as much as that in a 150-pound woman.
- **C.** about the same as that in a 140-pound woman.
- **D.** about the same as that in a 160-pound woman.

20. The percentage of body weight made up of both water and fat is:
- **F.** greater for females after adolescence.
- **G.** greater for males after adolescence.
- **H.** approximately the same for both sexes after adolescence only.
- **J.** approximately the same for both sexes at all ages after birth.

21. What hypothesis about the role of sex hormones during adolescence might be advanced from the graphs?
- **A.** Both male and female sex hormones cause an increase in the percent of fat in the body.
- **B.** Male hormones cause a reduction in the percent of fat and an increase in the percent of water.
- **C.** Female hormones have a much greater influence than male hormones on the percent of water.
- **D.** Male hormones cause the growth of male secondary sex characteristics.

BEST COPY AVAILABLE

GO ON TO THE NEXT PAGE.

Passage V

Seeds are tested for their ability to produce substances that kill microorganisms. Each seed is placed on cultures of two bacteria (*Staphylococcus* and *Escherichia*) and two molds. Seeds are classified on a scale of 0 (no effect) to 5 (strong effect), according to the amount of microorganism-free space that develops around the seed.

Experiment 1

Seeds of two members of the Lily family are tested against four different microorganisms:

Microorganism	Lily Family	
	Garlic	Daylily
Staphylococcus	4	0
Escherichia	5	4
Bread mold	2	2
Penicillium mold	3	0

Experiment 2

The same experiment is repeated using seeds of two members of the Composite family:

Microorganism	Composite Family	
	Dandelion	Thistle
Staphylococcus	5	5
Escherichia	4	5
Bread mold	4	3
Penicillium mold	2	2

Experiment 3

The experiment is then done with two members of the Legume family:

Microorganism	Legume Family	
	Soybean	Alfalfa
Staphylococcus	0	0
Escherichia	4	2
Bread mold	2	3
Penicillium mold	3	4

22. Which of the microorganisms is most susceptible to attacks by the chemicals produced by seeds?
- **F.** *Staphylococcus*
- **G.** *Escherichia*
- **H.** Bread mold
- **J.** *Penicillium* mold

23. Of the following, which kind of seed is more effective against molds than against bacteria?
- **A.** Alfalfa
- **B.** Daylily
- **C.** Thistle
- **D.** Garlic

24. To find an antibiotic that will protect oranges against *Penicillium* mold, a scientist would concentrate on:
- **F.** seeds of the thistle and its close relatives.
- **G.** a variety of members of the Composite family.
- **H.** members of the Legume family.
- **J.** seeds of the daylily and its relatives.

25. What conclusion can be reached about bread mold?
- **A.** It can survive by attacking seeds.
- **B.** It is highly resistant to chemical poisoning.
- **C.** It cannot be destroyed by seeds of the Composite family.
- **D.** It is moderately susceptible to attack by many kinds of seeds.

26. What hint might a scientist trying to find an antibiotic to control *Staphylococcus* infections get from these experiments?
- **F.** Looking for seeds that produce such an antibiotic would be a waste of effort.
- **G.** It would be inadvisable to concentrate on seeds of the Legume family.
- **H.** It would be wise to concentrate on *Penicillium* mold and its close relatives.
- **J.** The scientist should not waste time trying the bread mold and its close relatives.

27. Which of the following ecological hypotheses is supported by the evidence of these experiments?
- **A.** Molds are better able to survive than bacteria wherever the two kinds of microorganisms compete.
- **B.** The Legume family produces valuable fodder crops because its seeds have a high survival rate.
- **C.** The bacteria *Escherichia* and *Staphylococcus* may be highly damaging to leguminous crops.
- **D.** The Composite family has so many successful sturdy weeds because its seeds destroy microorganisms.

BEST COPY AVAILABLE

GO ON TO THE NEXT PAGE

Appendix B continued

Passage VI

Experiments are done to test the optical properties of lenses immersed in media having different indices of refraction.

Experiment 1

A lens made of flint glass, index of refraction 1.720, is tested. A beam of parallel light rays is sent into the lens, and the distance from the lens to the point of convergence of the beam is measured. This is the focal length of the lens. This focal length is measured with the lens immersed in media of various indices of refraction.

Medium	Index of refraction	Focal length (cm)
Air	1.00	8
Folinol	1.24	13
Water	1.33	20
11% Sugar solution	1.50	39
Carbon disulfide	1.62	95
Methylene iodide	1.74	*

*Rays do not converge at all.

Experiment 2

Another lens is tested. It is made of the same kind of glass as in Experiment 1, but this lens is thicker, more strongly curved.

Medium	Index of refraction	Focal length (cm)
Air	1.00	5
Folinol	1.24	8
Water	1.33	12
11% Sugar solution	1.50	24
Carbon disulfide	1.62	60
Methylene iodide	1.74	*

Experiment 3

A lens made of a new plastic is then tested. This lens is identical in size and shape to the glass lens in Experiment 2.

Medium	Index of refraction	Focal length (cm)
Air	1.00	13
Folinol	1.24	34
Water	1.33	360
11% Sugar solution	1.50	*
Carbon disulfide	1.62	*
Methylene iodide	1.74	*

28. The index of refraction column is the same in all three experiments because:
- **F.** all three lenses have the same basic properties.
- **G.** the same liquids are used in all three experiments.
- **H.** the temperatures at which the experiments are performed are carefully controlled.
- **J.** the color of the light source is not allowed to change from one experiment to another.

29. As the index of refraction of the medium increases, what happens to the rays of light emerging from the lens?
- **A.** They converge more strongly in all cases.
- **B.** They converge more strongly on leaving the glass lenses, but not the plastic lens.
- **C.** They converge less strongly in all cases.
- **D.** They converge less strongly on leaving the plastic lens, but not the glass lens.

30. Making a lens thicker and more strongly curved:
- **F.** shortens the focal length.
- **G.** increases the focal length.
- **H.** increases the index of refraction.
- **J.** decreases the index of refraction.

31. A reasonable hypothesis that can be derived from Experiments 1 and 2 is that:
- **A.** a lens will not focus light if its index of refraction is lower than that of the medium it is in.
- **B.** methylene iodide tends to spread light out so that it does not come to a focus.
- **C.** the focal length of a lens depends entirely on the index of refraction of the medium it is in.
- **D.** the thicker a lens, the less the convergence it produces on the light that passes through it.

32. Measurements of the kind made in these experiments would NOT be useful in efforts to find:
- **F.** the index of refraction of a liquid.
- **G.** the way a prism in a fluid would bend light rays.
- **H.** the concentration of a sugar solution.
- **J.** the chemical composition of an unknown liquid.

33. The index of refraction of the plastic lens in Experiment 3 must be:
- **A.** less than 1.33.
- **B.** between 1.33 and 1.50.
- **C.** more than 1.33.
- **D.** more than 1.50.

BEST COPY AVAILABLE

GO ON TO THE NEXT PAGE.

Appendix B continued

Passage VII

Two scientists disagree on the question of the origin of petroleum.

Scientist 1

There have been many theories suggesting a nonorganic origin of petroleum. but none of them have been successful. It is now accepted almost universally by geologists that petroleum comes from the decay of living things. Petroleum formation occurs in enclosed oceanic basins, such as the Black Sea. There must be an extremely large and continuous supply of marine organisms, adding their corpses to an accumulation at the bottom of the sea. They are quickly buried in sediment, so quickly that they do not have time to decay. In the enclosed basin, there is little circulation, so there is no supply of fresh, oxygenated water. In the absence of oxygen, there is little decay. The organic matter of the corpses degenerates into hydrocarbons, which accumulate as oil and gas. Since oil is lighter than water, it rises. As the deposits are covered with more sediments, the oil and gas rise into them and accumulate there. Petroleum geologists know that oil is often found in salt domes, formed by the evaporation of seawater.

Scientist 2

The current theory about the origin of petroleum postulates a very unlikely combination of circumstances. It needs an enclosed basin, exceptionally rich in marine life, with sediments pouring rapidly into it from the surrounding countryside. Although this combination might occur occasionally, it is too rare to account for the enormous earth areas underlain by petroleum. In my opinion, oil has been present deep in the earth since its origin. Meteorites, comets, and satellites are rich in hydrocarbons. The earth formed by agglomeration of these kinds of objects. After the earth formed, the hydrocarbons seeped upward, accumulating in porous sedimentary rocks. However, oil and gas are sometimes found seeping out of igneous rocks, which have no fossils at all, if these rocks have been thoroughly fractured by deep earthquakes. Oil wells now drill down to only about 15,000 feet. A recent explorational drilling found an oily sludge at 20,000 feet. If we could get to 30,000 feet, we would find an enormous pool of oil underlying the whole crust of the earth.

34. Both scientists agree that petroleum:
   F. forms at the bottom of the sea.
   G. seeps upward into sedimentary rocks.
   H. is present in great quantities below 30,000 feet.
   J. has always been present on earth.

35. Which of the following discoveries would greatly weaken the argument of Scientist 2?
   A. A vast oil deposit is found in sedimentary strata 20,000 feet deep.
   B. A meteorite is analyzed and found to contain few hydrocarbons.
   C. The sludge discovered at 20,000 feet turns out to be contamination from drilling oil.
   D. A large accumulation of oil is found in highly fractured igneous rock.

36. According to Scientist 2, what strategy would be most likely to increase world supplies of petroleum?
   F. Drill wells to greater depths.
   G. Inccrease exploration of offshore sedimentary strata.
   H. Drill wells in igneous rocks.
   J. Develop techniques of extraction from meteorites.

37. Exploration of the Persian Gulf reveals that it is an enclosed body of water rich in marine life. According to the hypothesis of Scientist 1, what additional condition would be necessary in order for petroleum deposits to develop?
   A. Vertical circulation to carry oxygen downward
   B. High concentration of salt in the water
   C. An accumulation of meteorites
   D. Rapid deposition of sediments

38. Scientist 2 considers that oil seepage from igneous rocks is damaging to Scientist 1's theory because igneous rocks:
   F. are easily fractured by earthquakes.
   G. never contain fossils.
   H. are always located deep in the crust.
   J. contain many meteorites.

39. What evidence given by Scientist 1 was not refuted by Scientist 2?
   A. There have been many theories of a nonorganic origin of petroleum, and all of them have failed.
   B. Petroleum is very often found associated with salt domes.
   C. All petroleum deposits are in porous sedimentary rock.
   D. Meterorites come to earth in the ocean just as often as on land.

40. The chief objection that Scientist 2 has to the theory of Scientist 1 is that it:
   F. postulates the formation of petroleum in a highly unusual set of conditions.
   G. cannot account for the accumulation of petroleum in sedimentary rocks.
   H. arbitrarily rejects the theory of nonorganic origin.
   J. places a limit on the amount of petroleum that can be extracted from the earth.

END OF TEST 4

STOP! DO NOT TURN THE PAGE UNTIL TOLD TO DO SO.

Science Reasoning

DIRECTIONS: This test consists of a science passage obtained from the science reasoning section of ACT exam. Each science reasoning section consists of seven sections with 40 questions that should be completed in 35 minutes. Thus, this passage should take no more than 5 minutes. Each passage is followed by a number of multiple-choice questions based on the passage. Study the passage below, and then select the best answer to each question. You are allowed to reread the passage. Record your answer by blackening the appropriate space on the answer sheet.

The charts below show the composition of the average American diet as it exists (solid) and as recommended by the National Research Council (cross-hatched). The kilocalorie is a measure of the energy content of food.

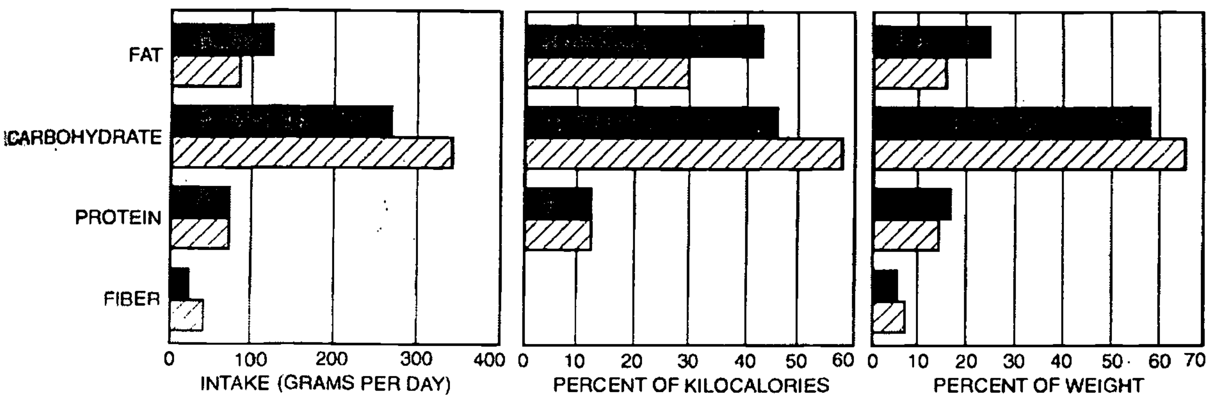


Illustration by Andrew Christie from "Diet and Cancer," by Leonard A. Cohen. Copyright © November 1987 by Scientific

1. If the recommendations of the National Research Council were followed, people would eat:
   A. more protein and less fiber.
   B. more protein and less fat.
   C. more carbohydrate and less fat.
   D. D. more fiber and less protein.

2. Comparison of the charts shows that:
   A. most of our food energy comes from proteins.
   B. we now get much more of our energy from carbohydrates than from fats.
   C. we cannot increase our energy intake by eating more fiber.
   D. the quantities of fats and carbohydrates in our present diet are approximately equal.

3. According to these recommendations, what comment can be made about the present American diet?
   A. It is overloaded with carbohydrates.
   B. It has too much fiber.
   C. It does not have enough fat.
   D. It contains the proper amounts of proteins.

4. If the recommendations for a changed diet were followed, our diet would have about:
   A. four times as much carbohydrate as fat.
   B. two and a half times as much carbohydrate as fat.
   C. equal amounts of carbohydrate and fat.
   D. nearly twice as much carbohydrate as fat.

5. Comparison of the percent by weight of the different nutrients in the diet and the percent of energy each supplies shows that:
   A. 1 gram of fat supplies more energy than 1 gram of carbohydrate.
   B. 1 gram of carbohydrate supplies more energy than 1 gram of protein.
   C. 1 gram of protein supplies about three times as much energy as 1 gram of fiber.
   D. 1 gram of carbohydrate supplies more energy than 1 gram of fat.

Science Reasoning

DIRECTIONS: This test consists of a science passage obtained from the science reasoning section of ACT exam. Each science reasoning section consists of seven sections with 40 questions that should be completed in 35 minutes. Thus, this passage should take no more than 5 minutes. Each passage is followed by a number of multiple-choice questions based on the passage. Study the passage below, and then select the best answer to each question. You are allowed to reread the passage. Record your answer by blackening the appropriate space on the answer sheet.

A bacteriologist is investigating the use of glucose by a type of bacterium as a source of energy in spore formation.

Experiment 1

The bacteria are grown in a nutrient solution containing a supply of glucose. When the glucose has been largely depleted, the contents of each cell shrink away from the cell wall and form a spore, which is highly resistant to environmental damage of all kinds.

Experiment 2

A culture of the bacteria is grown in a medium containing little glucose. The bacteria use the glucose as they grow, but do not form spores when the glucose has been depleted.

Experiment 3

A culture is grown in a medium containing ample glucose, but the cells are removed while there is still plenty of glucose in the medium. They are placed in distilled water, and form spores in about 13 hours.

Experiment 4

As in Experiment 3, cells are transferred from a glucose-rich medium to distilled water. If glucose is added to the water 5 hours later, the cells never form spores. If glucose is added 10 hours after the transfer, spores form 3 hours later.

6. Comparison of Experiments 1 and 2 shows that:
   A. glucose is necessary for the bacteria to grow.
   B.. the process of spore formation needs a good supply of glucose.
   C. bacteria can protect themselves against unfavorable conditions by forming spores.
   D. spore formation is inhibited by large concentrations of glucose.

7. A reasonable hypothesis from Experiment 3 is that:
   A. distilled water promotes the formation of spores.
   B. distilled water retards the formation of spores, but does not prevent it.
   C. bacterial cells store enough glucose to form spores.
   D. bacterial cells are able to form spores without any source of glucose.

8. Experiment 1 indicates that spore formation is stimulated by deprivation of glucose. Considering the results of Experiment 4, how long must this deprivation continue?
   R Less than 5 hours
   G. Somewhere between 5 and 10 hours
   H. More than 10 hours
   J. At least 13 hours

9. In Experiment 4, adding glucose to the distilled water after 10 hours:
   A. causes spores to form 3 hours later.
   B. delays the formation of spores for 3 hours.
   C. speeds up the formation of spores by 5 hours.
   D. has no effect at all on the formation of spores.

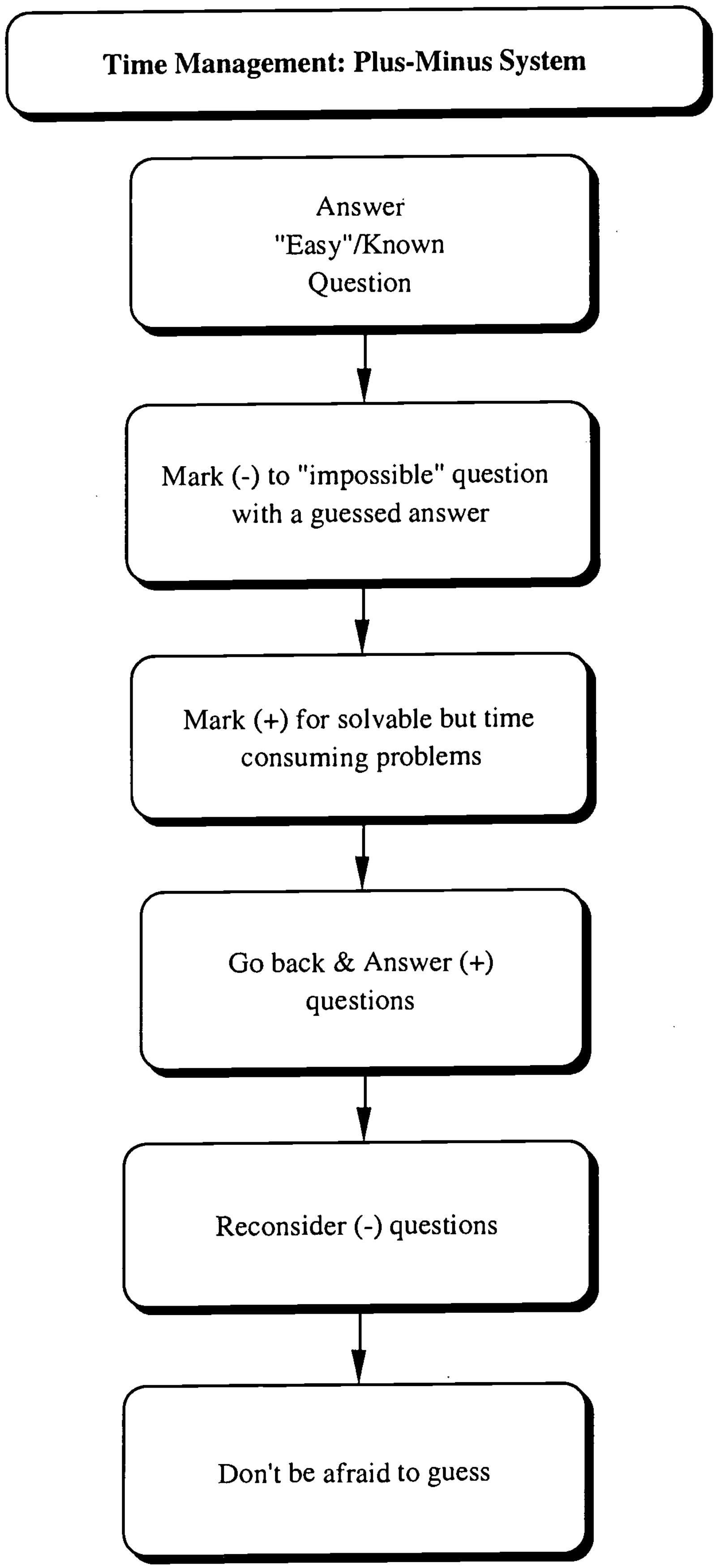
10. Which of the following experiments would NOT be useful in efforts to learn more about the way bacteria use sugars in spore formation?
    A. Repeat Experiment 4 adding glucose to the water at various times after transferring the bacteria to distilled water.
    B. Repeat Experiments 3 and 4 using bread molds instead of bacteria as the spore-forming organism.
    C. Repeat Experiments 1 and 2 using other kinds of sugar than glucose as energy sources.
    D. Repeat Experiment 2 using different concentrations of glucose

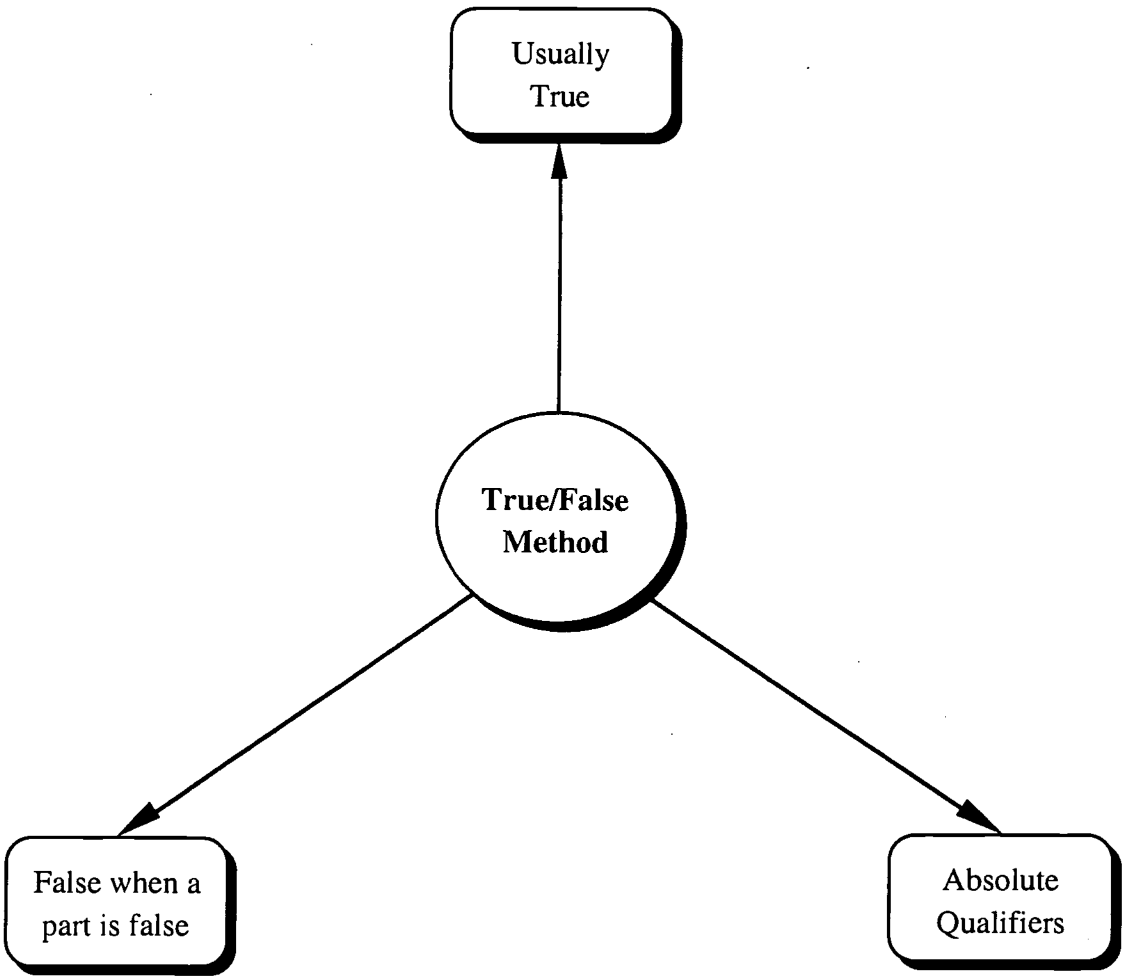
11. According to these experiments, what condition must be met in order for this type of bacterium to form spores?
    A. A good supply of glucose in the medium, followed by a period in which there is little glucose
    B. A steady supply of glucose in high concentration
    C. A prolonged period of glucose deprivation
    D. A sudden increase in the concentration of glucose in the medium

Appendix D
Intervention Method

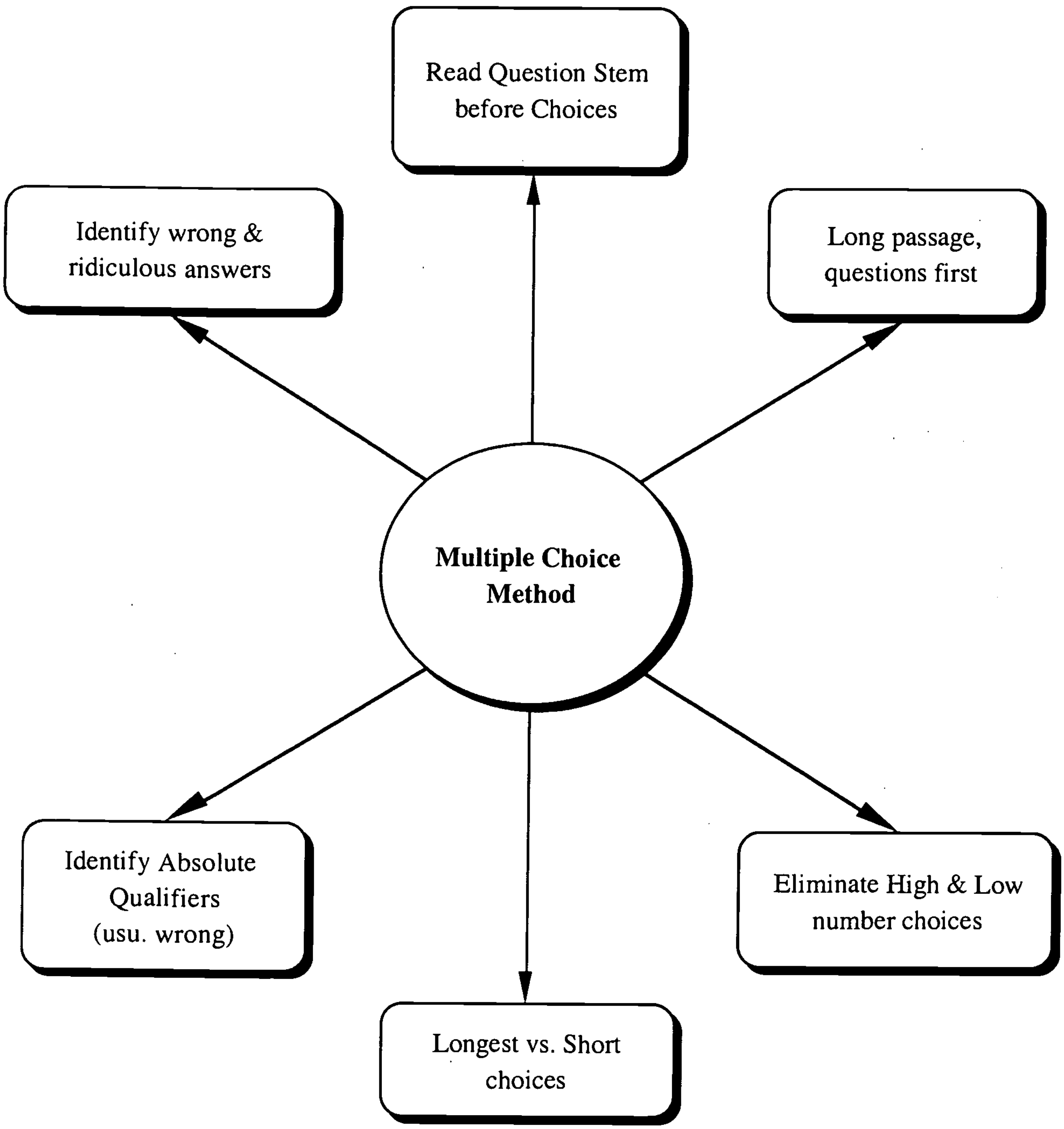
The Question-Answer Relationship (QARS)

Kind of Question	Textually Explicit or Implicit	Characteristics	Key Words/Notes
Right There	Explicit	Answer stated directly in the text	Name...... List ...... In what year ......
Think & Search	Implicit	Answer requires students to put together several of the text.	Compare & contrast Cause & Effec What caused.... What are some of the ..
Author & You	Implicit & Explicit	Answer requires that students mesh their thoughts about a topic with information the author presents	Evaluate based on info.. Why do you think.....
On Your Own	Implicit	Answer comes from application of the information from the text	Your interpretation What do you think....


Time Management: Plus-Minus System
Answer "Easy"/Known Question
Mark (-) to "impossible" question with a guessed answer
Mark (+) for solvable but time consuming problems
Go back & Answer (+) questions
Reconsider (-) questions
Don't be afraid to guess


Usually True
True/False Method
False when a part is false
Absolute Qualifiers

Appendix D continued



U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC

TM034550

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Circumventing the Pressures of Standardized Norm-Referenced Tests

Author(s): Camel, Christine; Chung, Tom

Corporate Source: Saint Xavier University

Publication Date: ASAP

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1	Level 2A	Level 2B
[X]	[ ]	[ ]
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, please

Signature: Christine Camel T. Chung	Printed Name/Position/Title: Student/FBMP
Organization/Address: Saint Xavier University 3700 W. 103rd St. Chgo, IL	Telephone: 708-802-6219 FAX: 708-802-6208 E-Mail Address: crannell@sxu.edu Date:

William Crannell, Ed.D.

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:	ERIC/REC 2805 E. Tenth Street Smith Research Center, 150 Indiana University Bloomington, IN 47408