

Utah State University

DigitalCommons@USU

---

All Graduate Theses and Dissertations

Graduate Studies

---

5-1979

## ACT Score Declines: Looking for the Source

Thomas Edward Atkin  
*Utah State University*

Follow this and additional works at: <https://digitalcommons.usu.edu/etd>



Part of the [Psychiatry and Psychology Commons](#)

---

### Recommended Citation

Atkin, Thomas Edward, "ACT Score Declines: Looking for the Source" (1979). *All Graduate Theses and Dissertations*. 5561.

<https://digitalcommons.usu.edu/etd/5561>

This Thesis is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact [digitalcommons@usu.edu](mailto:digitalcommons@usu.edu).



ACT SCORE DECLINES: LOOKING FOR THE SOURCE

by

Thomas Edward Atkin

A thesis submitted in partial fulfillment  
of the requirements for the degree

of

MASTER OF SCIENCE

in

Psychology

Approved:

UTAH STATE UNIVERSITY  
Logan, Utah

1979

## ACKNOWLEDGEMENTS

I would like to thank Drs. William Dobson and Michael R. Bertoch for their continued aid throughout my academic career.

Special thanks to friend and mentor, Dr. Keith T. Checketts, for enduring and guiding, ever so gently, this frustrated novelist.

Most especially thanks to Pooh, Tigger, Piglet, and little Roo. But, most especially to you, Pooh. Things just wouldn't be as thingish without you. Eeyore.

Thomas Edward Atkin

## TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS .....	ii
LIST OF TABLES .....	v
LIST OF FIGURES .....	vi
ABSTRACT .....	vii
Chapter	
I. INTRODUCTION .....	1
II. REVIEW OF LITERATURE .....	4
Describing the Decline .....	5
Explaining the Decline .....	8
Problems with the test .....	9
Changes in the testing population .....	10
Changes in society .....	13
Changes in schools .....	14
Summary .....	16
III. RESEARCH DESIGN .....	18
Hypotheses .....	18
Sample .....	18
Measures .....	19
Procedure .....	19
Statistical Analysis .....	21
IV. RESULTS .....	22
Skyview High School Sample .....	22
Logan High School Sample .....	23
Logan High School Subsample .....	23



## TABLE OF CONTENTS (Continued)

Chapter	Page
V. DISCUSSION .....	32
Evaluation of Findings .....	32
Logan High School data .....	32
Skyview High School data .....	33
Summary .....	33
Limitations .....	33
Recommendations for Further Research .....	34
BIBLIOGRAPHY .....	36

## LIST OF TABLES

Table	Page
1. Two-Way Analysis of Variance Repeated Measures Results on ACT English Subtest Scores for Logan High School Male/Female Sample . . . . .	29
2. Two-Way Analysis of Variance Repeated Measures Results on ACT Mathematics Subtest Scores for Logan High School Male/Female Sample . . . . .	29
3. Two-Way Analysis of Variance Repeated Measures on ACT Social Studies Subtest Scores for Logan High School Male/Female Sample . . . . .	30
4. Two-Way Analysis of Variance Repeated Measures Results on ACT Natural Science Subtest Scores for Logan High School Male/Female Sample . . . . .	30
5. Two-Way Analysis of Variance Repeated Measures Results on ACT Composite Scores for Logan High School Male/Female Sample . . . . .	31

## LIST OF FIGURES

Figure	Page
1. Mean scores for Logan High School sample . . . . .	25
2. Mean scores for Skyview High School sample . . . . .	26
3. Mean scores for Logan High School males . . . . .	27
4. Mean scores for Logan High School females . . . . .	28

## ABSTRACT

ACT Score Declines: Looking for the Source

by

Thomas Edward Atkin, Master of Science

Utah State University, 1979

Major Professor: Dr. Keith T. Checketts  
Department: Psychology

During the last 15 years there has been a well documented decline in achievement test scores. Declines have been documented on both the American College Testing Program (ACT) and the Scholastic Aptitude Test (SAT). An increased interest in the score declines and their possible explanations has also taken place in the last 10 years. The explanation for the declines fit under four main headings: (1) problems with the tests, (2) changes in the testing population, (3) changes in society, and (4) changes in the schools. The purpose of this study was to begin with the test itself and check for possible score drift between two forms of the ACT, a 1977 ACT form and a form 5 to 7 years older than that.

The sample population consisted of 242 juniors at Logan High School, Logan, Utah, and 153 juniors at Skyview High School, Smithfield, Utah, during the 1976-77 school year. These samples represent 83% of the Logan

High junior population and 50% of the Skyview High junior class. All subjects took the two forms of the ACT mentioned above.

In analyzing the data from the major samples with a two-tailed t test, it was found that there was a significant difference in the two tests. The largest difference was found on the mathematics subtest; the difference being that students scored higher on the older forms, not only on the mathematics subtest, but on the other subtests, also. A two-way analysis of variance with repeated measures was used to check for variance due to sex of subjects and form of test and their interactions. The findings were consistent with those of the t test. It was found that the mathematics scores varied due to the form of the test. The only sex difference was that females' standard scores were higher than males' on the English subtests of both ACT forms.

This study found a significant difference between an older ACT form and a 1977 ACT form as measured by subjects' scores on both. The findings of this research indicate that the test itself may be a partial cause of the current ACT score declines.

(46 pages)

## CHAPTER I

### INTRODUCTION

During the last 15 years, there has been a steady downward trend in achievement test scores. The declines are evident and very similar in almost all achievement tests administered to students in the fifth grade and up (Harnischfeger & Wiley, 1975). Among the standardized achievement tests which have demonstrated a decline in scores have been the Scholastic Aptitude Test (SAT) and the American College Testing Program (ACT). These tests are the two most widely used college entrance examinations in the United States. The concern has mounted as the trend has continued. Along with the research that has begun in order to find the reasons behind the declines, a great deal of pure speculation and conjecture is evident in the popular press.

The score declines pose some very important questions that demand carefully researched answers. Do these declines reflect a true drop in the aptitudes of students or are they an artifact of the tests themselves? If the scores reflect a true decline in students' aptitudes and college readiness, what is or is not happening to students today to cause this? Societal conflicts, the effects of the youth culture, growing divorce, drugs, television, and family size, along with school curriculum and the tests themselves have been pointed to as at least contributing to the decline.

It is the purpose of this study to examine one of these areas more closely, the test itself. Several authors (Munday, 1976; Harnischfager & Wiley, 1975, 1976; Ferguson, 1976), have stated that the logical place to start looking for the causes of the declines is with the tests themselves.

Munday (1976), in discussing the ACT and the possibility of test technical factors contributing to the decline, states:

There have been no changes over the last 15 years in methods used for scaling and equating of the ACT tests. Scaling and equating are two technical procedures used to insure that the difficulty of one test form is equal to that of others. The methods used by ACT are explicit and straightforward, and are unlikely to bias test scores one direction or the other. There is every reason to assume that the SAT experience is similar. Each year three new forms of the ACT tests are developed and equated to a form used the previous year. This form is called the anchor form, and because a new anchor form is used yearly, it is possible some downward drift may have occurred. It is unlikely, however, that this would have occurred year after year. (p. 2)

The major objective of this study is to investigate the ACT for score drifts by comparing scores obtained by the same sample on two forms of the ACT. One of these forms was in use during the 1976-77 testing year and the other was used approximately 5 to 7 years prior to the 1976-66 testing year.

Following, in Chapter II, is a review of relevant literature concerning the test score declines. Empirical research findings as well as articles from the popular press will be used to: (1) document the magnitude and type of declines experienced and (2) put forth the many explanations to which the declines have been related. Chapter III is a statement of the methods and

procedures of this study. Results as related to the hypothesis are reported in Chapter IV, and a discussion of the findings, limitations of the study and recommendations for future research in this area are contained in Chapter V.



## CHAPTER II

## REVIEW OF LITERATURE

The literature on the score declines can be divided into two areas; those that describe the declines and those that attempt to explain the declines. The decline in scores was not noticed until the late 60's and then it was found that the data being kept to that point was inadequate to accurately map the declines (Harnischfeger & Wiley, 1975). Before research attempting to explain the data could be undertaken, accurate descriptions of the declines had to be made. Harnischfeger and Wiley (1976) found that:

Up to the mid-1960's, achievement test scores steadily increased. Since then, scores have been declining in all tested achievement areas for grades five through twelve, with more dramatic drops occurring in recent years and being most evident for higher grades. The declines have been most pronounced in verbal tests and, therein, for college-bound females. (p. 5)

The present review is most concerned with the findings on the American College Testing (ACT) assessment and, because of its similarities in testing population and those aptitudes which it tests, the Scholastic Aptitude Test (SAT). The American College Testing (ACT) assessment is the second most widely used college admissions test in the United States. The test is comprised of four subtests: (1) English, (2) mathematics, (3) social studies readings, and (4) natural sciences. Scores are reported separately for each of the subtests and a composite score, the mathematical

mean of the four subtests, is also reported. The range of scores on the ACT is from 1 to 36 (Ferguson, 1976). The SAT is the most widely used college admissions test in the United States today. In 1926, the first regularly used college admissions test of the objective type was introduced; this was the first SAT (Harnischfeger & Wiley, 1976). The modern SAT has been constructed and administered by the Educational Testing Service since 1948. It is comprised of two subtests, verbal and mathematics. Scaled scores from 200 to 800 are reported for both subtests as well as a composite score.

#### Describing the Decline

Tolman (1976) reports a decline of 1.6 standard score points on the composite score of the ACT over the 10 year period from 1964-65 to 1974-75. He also reports that the 1974-75 testing year represents the most dramatic decline in overall scores. For the 1964-65 testing year, the mean composite score on the ACT was 19.9. This score had declined to 18.7 by the 1973-74 testing year (Munday, 1976), and 18.3 by 1974-75 (Tolman, 1976).

Of the four ACT subtests, social studies readings shows the greatest declines over the 10 year period from 1964-65 to 1974-75 declining from a high in 1964-65 of 20.6 to a low of 17.1 in 1974-75, a drop of 2.5 standard scores. Mathematics usage dropped 1.5 standard scores from 19.6 to 18.1 over the same period. English usage showed a

similar decline from 18.7 to 17.3, 1.4 standard scores. The natural science subtest was the only subtest that did not show a decline; in fact it showed a slight increase going from 20.4 to 20.8 in the same 10 year period (Tolman, 1976; Munday, 1976; Ferguson, 1976).

McCandless (1975), Munday (1976), Ferguson (1976), and Harnischfeger and Wiley (1975, 1976) all report that declines on the SAT and ACT are consistent with one another. Harris (1976) reports that during a period of 12 years from 1962-63 to 1974-75, verbal standard scores on the SAT have declined 41 points. Munday (1976) reports a larger drop of 44 points from 478 in the 1962-63 testing year to 434 in 1974-75. There is a similar difference in reports for mathematics scores with Harris (1976) reporting a 29 point decline for the same 12 year period, while Munday (1976) reports a 30 point drop from 502 to 472. Also consistent with the ACT, the 1974-75 testing year showed the largest drops of any of the testing years studied (Harris, 1976).

Both the ACT and SAT have shown differences in decline for males and females. In general, the decline for females have been more substantial than for males. Ferguson (1976) and Munday (1976) report that ACT composite scores in the 10 year period from 1964-65 to 1974-75 have declined 1.2 standard scores for males, while female scores have dropped 2.1 standard scores. The social studies scores declined the most for both groups: 2.3 standard scores for males and 4.5 for females. Though women still score

higher on the English test, their standard scores declined 1.9 points as compared to 1.1 points for males. This was reversed on the mathematics test where male declines were slightly larger than females' even though males continue to score higher on the mathematics test. The only area on the ACT resisting the declines was the natural sciences subtest. Females' scores on the natural science subtest remained relatively constant at 19.7 while males' increased from 21.0 to 22.0. While male declines were not always smaller than the females' on the ACT (male declines were slightly larger on the mathematics subtest), as reflected in the composite scores, female score declines have been larger than those for males. SAT scores have shown similar trends with overall scores for females dropping more rapidly than for males (Rorak, 1977; Edson, 1976; Harnischfeger & Wiley, 1975, 1976).

In summary, several points may be made about the ACT and SAT score declines. The SAT and ACT score declines are well documented over the past 10 to 15 years (Harnischfeger & Wiley, 1975, 1976). The SAT and ACT show similar declines over this time period. The declines over any one year are not cause for alarm, but a trend has definitely been established which is cause for concern (Munday, 1976). The decline in scores for both the ACT and SAT have been more pronounced for females than for males (Ferguson, 1976). The decline is not consistent across subject areas. The verbal scores on the SAT are dropping more rapidly than scores for the mathematics subtest. On the ACT, the social studies

reading scores have declined dramatically, while natural science scores have increased slightly (Ferguson & Maxey, 1976). Declining scores are observable across all geographical regions of the United States, lending support to there being real declines not isolated to any particular culture or group of people (Munday, 1976).

### Explaining the Decline

In attempting to explain the score declines, hypotheses can be listed under four main headings: (1) problems with the test, (2) changes in the testing population, (3) changes in society, and (4) changes in the schools (Harris, 1976).

In reviewing the findings of the College Entrance Examination Board's special advisory panel appointed to look into the SAT declines Fields (1977b) quotes from their findings:

Searching for the causes of the SAT score decline over the last six or seven years is essentially an exercise in conjecture. So much has happened that may have effected this record that there is no way of telling what did. The only evidence is circumstantial. (p. 1 & 13)

This statement is repeated in meaning often in the literature (Angoff, 1975; Edson, 1976; Harnischfeger & Wiley, 1975, 1976; Ferguson, 1976; Ferguson & Maxey, 1976; Hechinger, 1974; etc.). Some authors have put forth educated guesses, but for the most part, most "causes" for the score declines are still hypothetical.

Problems with the test. Harnischfeger and Wiley (1975, 1976) suggest that, though looking for drifts in SAT scaling is a logical place to start, the decline cannot be attributed to test artifacts. Munday (1976) agrees and calls attention to the similarity of the SAT and ACT equating procedures. Other authors (Harris, 1976; Ferguson, 1976; Angoff, 1975) agree that test technical factors such as equating, scoring and calibrating are not responsible for the score declines. Breland (1976), however, states that test specific factors such as scaling and test construction may be contributing to the decline.

Munday (1976) checked ACT subject content for changes and found no substantial variability in predictive validity which would change if content were to vary appreciably. Ferguson (1976), Ferguson and Maxey (1976), Harnischfeger and Wiley (1975, 1976), Harris (1976), and Edson (1976), all support Munday's findings of no content change on the ACT or SAT.

A related question is--Are the tests still measuring important skills--in other words, have the tests kept pace with changing curricula? A drop in predictive validity would be expected if this were not the case (Munday, 1976). Harnischfeger and Wiley (1976) address the question of the relationship between the SAT and school curricula by stating:

We must continue to emphasize that the test was not designed nor intended to measure school performance. The SAT was designed to measure long term developed verbal and mathematical reasoning abilities, attributes that are related to college performance. (1. 17)



The question of changing curricula will be dealt with in greater detail later in this paper.

Another criticism of the tests is that they lack instructional validity. According to Feldhusen, Hymes, and Ames (1976), instructional validity subsumes content validity. A test is instructionally valid when learning objectives have been specified, when prior learning opportunities have been provided to learn those objectives, and when the behavioral conditions of learning match the behavioral demands of the test items. Feldhusen et al. (1976) conclude that the tests do not have instructional validity. While worthwhile learning may be going on in the schools, the tests, because of their decline in instructional validity, are failing to assess what is being learned. Instructional validity on a national assessment such as the ACT or SAT would be extremely difficult to apply (Feldhusen et al., 1976).

Changes in the testing population. One group of explanations for the score declines focuses on the testing population and possible changes in its makeup. Throne (1976) states that the entire problem may be explained by carefully analyzing the very definition of standardized tests; a standardized test being a measure of performance under standardized (uniform) conditions. Throne claims that administering standardized tests normed on samples not presently representative of the current testing population is a violation of the standardized test, and this is what Throne

claims is happening. Throne states that there are too many variables at work to ever be able to assume that the norm group is representative of the actual testing population.

Other authors support the theory of a change in the pool of test takers. Ferguson (1976) reports an increase in the standard deviations since 1965 across all subtests on the ACT. This points to a greater heterogeneity among the testing population. Several theories have been raised to explain this change in students tested. The female percentage of ACT test takers has increased from 45% in the 1964-65 testing year to 53% in the 1974-75 year. In addition, the percentage of women scoring in the lower score intervals of the test has increased (Ferguson, 1976; Munday, 1976; Harnischfeger & Wiley, 1976). Harris (1976) notes the same trend in the SAT experience. Harris goes on to state that, though this is an established trend, the increase in female test takers in the test population cannot explain more than the very smallest portion of the score decline. In addition, it is not just females' scores that are dropping. Males' scores have also been declining.

Another question along the same line has to do with minorities. Has there been an increase in traditionally low scoring minority students taking the tests? Edson (1976), in examining SAT statistics, found that the percentage of minority students taking the test has remained constant between 1972 and 1975, at 14%. ACT data (Ferguson, 1976) supports these



findings in that between the 1970-71 and 1974-75 testing years, the minority population remained stable at 11%. Burns (1976) conducted a study on minority students and the SAT and found that over the 2 testing years of the study, the minority students' scores, though predictably lower than the norms, did not decline at a different rate than did any other group's scores. In summary, Burns states that other explanations will have to be found to explain the declines.

Another popular theory dealing with the test population has to do with more academically weak students taking the tests (Munday, 1976). From 1958 to 1968, the high school drop-out rate declined from 50% to 25%. Since then it has remained constant (Harnischfeger & Wiley, 1976). This coupled with the increased accessibility of colleges to students traditionally denied the opportunity of a college education, has made for larger numbers of lower ability students taking the ACT (Rever & Kojaku, 1975). More open college admissions has also been blamed for lessening student motivation and interest in school (Fields, 1977a). A 1973 Gallup Poll, "The Importance of Education to Success," showed that the endorsement concept linking higher education and success increased with age. In other words, students currently in school do not strongly support this concept linking higher education with success (Sapone & Giuliano, 1976).

Champagne and Roberts (1976) have shown that as interest in school increases, as measured by the Pennsylvania Department of Education's Educational Quality Assessment, SAT scores go up. From 1965 to 1973 the study

population's SAT scores declined in accord with the national trend. After implementing a Title III program aimed at increasing interest in school, the trends reversed themselves. From 1973 to 1975, while nationally the biggest declines ever were being experienced, this population's SAT scores jumped 85.2 points from 817.2 to 902.4. Russo and Checketts (1978) also found that the higher the student's educational aspirations, the higher his or her ACT score.

In summary, a greater heterogeneity exists in today's test population than existed 15 years ago. Changes in attitude on the parts of the students as well as many other variables such as pupil retention, more females and minorities being tested and more open college admissions explain the greater heterogeneity. Preliminary research shows that the change in the students tested has had some effect on the test scores. However, before definitive explanation about what kind or how much impact this set of variables has on the score declines can be made, further research is needed (Ferguson, 1976).

Changes in society. Little research is available regarding how our changing society effects the student and thereby, the test scores. How the "youth culture" and "drug culture" have effected the scores is hard to say. Since working mothers, single parents and growing divorce all have traumatic effects on a child's life, it seems plausible that they would also effect test scores (Sapone & Giuliano, 1976).

Television viewing has increased continually during the time period of the score declines (Fields, 1977a). Russo and Checketts (1978) found no

relationship between television watching habits and ACT scores. Television watching habits continue as an explanation, however, in the popular press.

The period of greatest decline, 1972 to 1975, was preceded by 5 or 6 years of disillusionment for the students in this country. Vietnam, Watergate, political assassinations, race riots; all had a traumatic effect on American youth. How much this trauma effected student's motivation and, hence, their test scores, is impossible to say (Fields, 1977a). As with the other societal changes, they are questions that need carefully researched answers.

Another societal variable is the size of families of test takers. Zajonc (1976) has plotted family size and birth order as it effects intellectual growth and has found some parallels. Zajonc's study draws the conclusion that intelligence generally suffers as family size and birth order increases. This phenomenon was observed across cultural and socioeconomic strata. Parallels were drawn between the children of World War II "baby boom" and the current downward trend in SAT scores. If Zajonc's theories are corrent, he predicts the downward trends in test scores should reverse themselves somewhere between 1978 and 1982.

In summary, societal changes have been vast and far-reaching over the past 1 1/2 decades. How these changes have effected the students and their test scores is still largely a matter of conjecture (Fields, 1977a).

Changes in schools. Munday (1976) collected test data from statewide assessment programs in both Iowa and Minnesota. These data

represented nearly all the students in these two states over a 20 year period. The data were similar to those found on the ACT, leading Munday to conclude that more poorly developed academic abilities are at least a partial explanation for the score declines.

One of the most promising explanations has to do with curricular changes in the high schools which make it possible for students to take fewer traditional college preparatory courses while taking more non-traditional course work. Harnischfeger and Wiley (1975, 1976) show a total drop in courses per pupil of 13%. The most noticeable declines took place in regular grade specific English classes (8%), foreign languages (7.5%), math (7%) and physical sciences (7%), with an accompanying drop in practical training courses (e.g., business, home economics, vocational training) of over 30%. These figures represent changes taking place between the 1970-71 and 1972-73 school years. The figures suggest fewer academic classes being taken, but also show that the students are not making up that deficit by taking more practical training classes. Russo and Checketts (1978) suggest that work study programs which have recently been implemented in secondary schools could feasibly detract from the more academic-cognitive course work traditionally pursued by students. He found that the number of academic courses taken explained 25% of the variability in ACT scores in his study. These findings do not necessarily mean students are learning less. It may be that they are simply learning from a different pool of knowledge (Harnischfeger & Wiley, 1975).

Another area pointed to by Harnischfeger and Wiley (1975) as having possible negative effects on achievement test scores is large academic class size. The literature in this area is inconsistent (Russo, 1978). Russo reviews the literature on class size and summarizes by stating, "it would appear that while most teachers and educators would be in general agreement that smaller class sizes are more conducive to achievement, the research literature is far from conclusive" (p. 24).

In summary, the literature in this area is inconclusive. Some authors (Harris, 1976) see the question of the schools' effects on the ACT as irrelevant. There are those (Sapone & Giuliano, 1976) who would exonerate the schools from all blame, while others (Harnischfeger & Wiley, 1975) see the schools as having had an effect on the achievement test score decline.

### Summary

The decline in achievement test scores has been amply documented (Munday, 1976; Harnischfeger & Wiley, 1975). The first part of this review was devoted to fully documenting and describing these declines. The second part of this review was spent putting forth some of the popular explanations for the declines. Many of these explanations result from speculation, prejudice, and poorly kept data (Harnischfeger & Wiley, 1975). To better understand the explanations, they were divided into four major areas: (1) test-related factors, (2) changes in the testing population, (3) changes in society

and (4) changes in the schools. What research there is in the area is inconclusive. More research is needed to determine the source (or sources) of the declines. It is the purpose of the present study to begin looking for the source of the declines at the most logical place (Harnischfeger & Wiley, 1975; Munday, 1976), the test itself.

### CHAPTER III

#### RESEARCH DESIGN

##### Hypotheses

The major objective of this study is to determine if there is any difference in what was the current 1977 form of the ACT and a 5 to 7 year old form of the ACT. The following hypotheses were formulated on the basis of the preceeding objectives:

1. There is no difference in scores obtained on a 1977 form of the ACT and an older (5 to 7 years) ACT form when administered to the same sample of students.
2. There is no difference in males' and females' scores obtained on a 1977 form of the ACT and an older (5 to 7 years) form.

##### Sample

Subjects for this study were selected from the 1976-77 junior classes of Logan High School, Logan, Utah, and the junior class of Skyview High School, Smithfield, Utah. The Skyview sample was comprised of 153 volunteers representing 50% of the Skyview junior class. These subjects had previously registered for the April, 1977, national test date.

The Logan sample was comprised of 242 of the 292 juniors enrolled at Logan High School for the 1976-77 school year. This figure represents



83% of the total class. The only students not included in the study from Logan were those who were absent on one or both testing dates. Only those students taking both the old and new ACT forms were included.

### Measures

Two forms of the American College Testing Program's ACT were used in this study. One was the then current 1977 form; the other form was one used 5 to 7 years previous to the 1977 form. These two forms are assumed to be equal as a result of the equating procedures used by ACT to insure continued equality of test forms (Munday, 1976). The American College Testing program has been developed for use in the admission, placement and counseling of students. The test consists of four separate subtests: English, mathematics, natural sciences and social studies. Scores are reported for all four subtests and an arithmetic mean of the four scores is reported as a composite score. The odd-even reliability coefficients obtained from a sample of over 900 students was .90 (English), .89 (mathematics), .83 (natural science), .86 (social studies), and .95 (composite) (Buros, 1965).

### Procedure

Data for this research were collected on the April 2, 1977, national ACT test date and on two dates during the following 2 weeks. The Skyview sample was given the 1977 ACT form on the regular national test date, but



were segregated from the other test takers. The older form was administered 4 days later on Wednesday, April 6, during a regular school day. The students were instructed that they were part of a study and that they would be given two forms of the ACT, one being the one currently in use and the other being an experimental form. The students were told that it was not known by the proctors which test they were taking on the national test date and which test they took the following Wednesday during school hours. They were told that Utah State University would honor the highest of the two scores when considering students for scholarship selection, but the scores from the current form would be used for admissions. All other instructions and procedures adhered to those prescribed in the ACT proctor's manuals.

Subjects from Logan High School were tested in a similar manner, the only exception being that both test dates were during school hours. The first testing was on Wednesday, April 6, and the second was on Monday, April 11, 1977. The Logan High School sample was divided into two equal groups, one group receiving the current ACT form first, the other, the old form first. The sample was split to negate the effects of order of testing. All instructions and procedures were the same as those followed at Skyview High School. Both groups were told that it was not known by the proctors which test they were taking and that the tests were indistinguishable, outwardly, from one another.

### Statistical Analysis

The data from the Skyview sample were analyzed using a t test for comparing means. The data from Logan High School was also analyzed with a t test. In addition, a male/female sample was generated from the Logan data at random to check for sex as a possible source of variance. This group was comprised of 160 subjects, 80 males and 80 females, representing 66% of the total sample. A two-way analysis of variance with repeated measures on one dimension was used to analyze the data from this subgroup. A null hypothesis was rejected if its statistical probability exceeded the .05 level.

## CHAPTER IV

## RESULTS

To answer the question, "Is there a difference between scores obtained on a then 5 to 7 year old ACT form and the then current 1977 ACT form," a t test was run on the data from the samples from both Logan and Skyview High School. A subsample was generated from the Logan High School sample in order to check for variance due to the sex of subjects. Subjects' data returned from ACT was not identifiable by sex, so the entire sample could not be divided on that basis. One hundred sixty subjects were identified by sex (80 males and 80 females) and an analysis of variance with repeated measures on one dimension was run on this subsample.

Skyview High School Sample

The Skyview data showed significance in three areas. In mathematics and social studies reading, subjects scored significantly higher on the old form (22.05 and 19.08, respectively) than they did on the 1977 form (mathematics, 21.03 and social studies, 18.01). The natural science subtest scores showed the Skyview subjects scoring significantly higher (3.14 standard scores) on the 1977 test than on the older test. Except for the natural science scores, the subjects scored better on the older ACT form than the 1977 ACT (English, 0.15 standard scores; composite, 0.31 standard scores). The data from the Skyview High School sample are shown in Figure 1.

### Logan High School Sample

The Logan data yielded two significant scores. The students at Logan scored significantly higher on the older ACT form than on the 1977 ACT in mathematics (3.30 standard scores) and on the composite scores (1.10 standard scores). Although no other scores reached significance, on all three of the other subtests, subject's older ACT form scores were higher than their 1977 ACT scores. The data from the Logan High School sample is shown in Figure 2.

### Logan High School Subsample

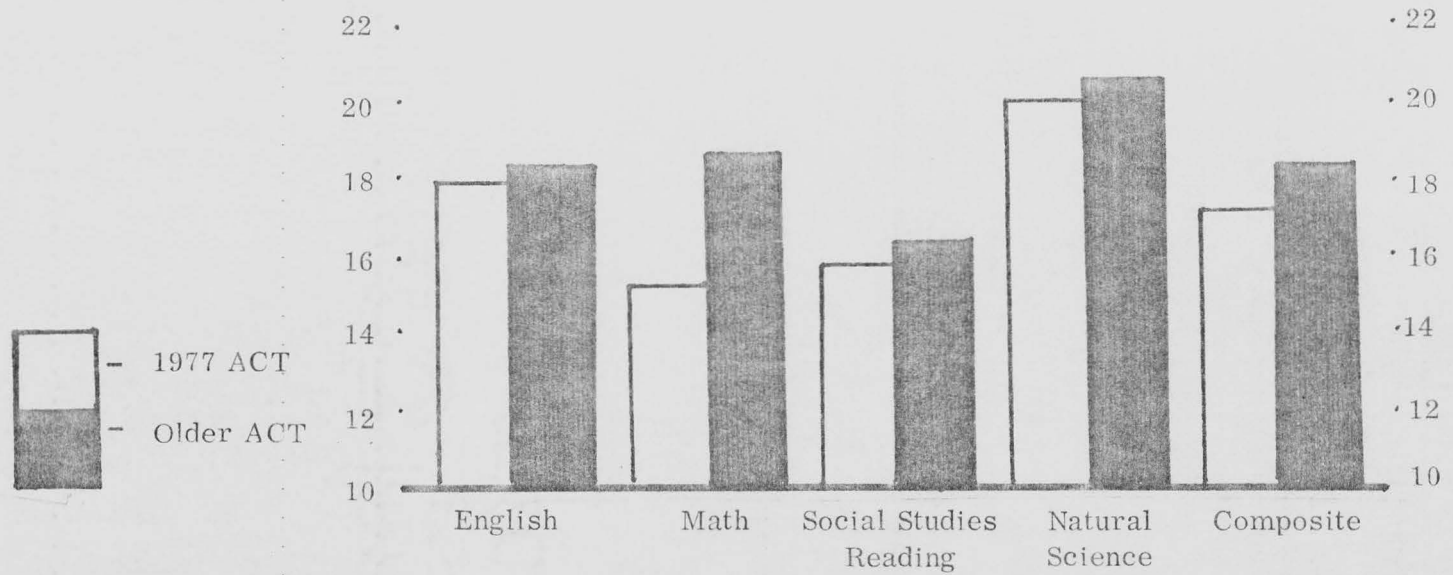
To answer the question, "Is there a difference due to sex between scores on the 1977 ACT and a 5 to 7 year older ACT," a two-way analysis of variance with repeated measures was run on a randomly selected group of 80 male and 80 female students at Logan High School. The means and difference scores are found in Figures 3 and 4 and the results of the two-way analysis of variance are found in Tables 1 through 5.

The means and difference scores found in Figures 3 and 4 show that both the male and female subsamples scored higher on the older ACT form than they did on the 1977 form. The exception was the female group's social studies subtest score showing an average 0.23 higher standard score on the 1977 form.

Three F values reached significance. Females scored significantly higher on the English subtest than did males on both tests, as was predicted

(Harnischfeger & Wiley, 1975). The other two significant F values were the mathematics "type of test" value, and the composite "type of test" value.

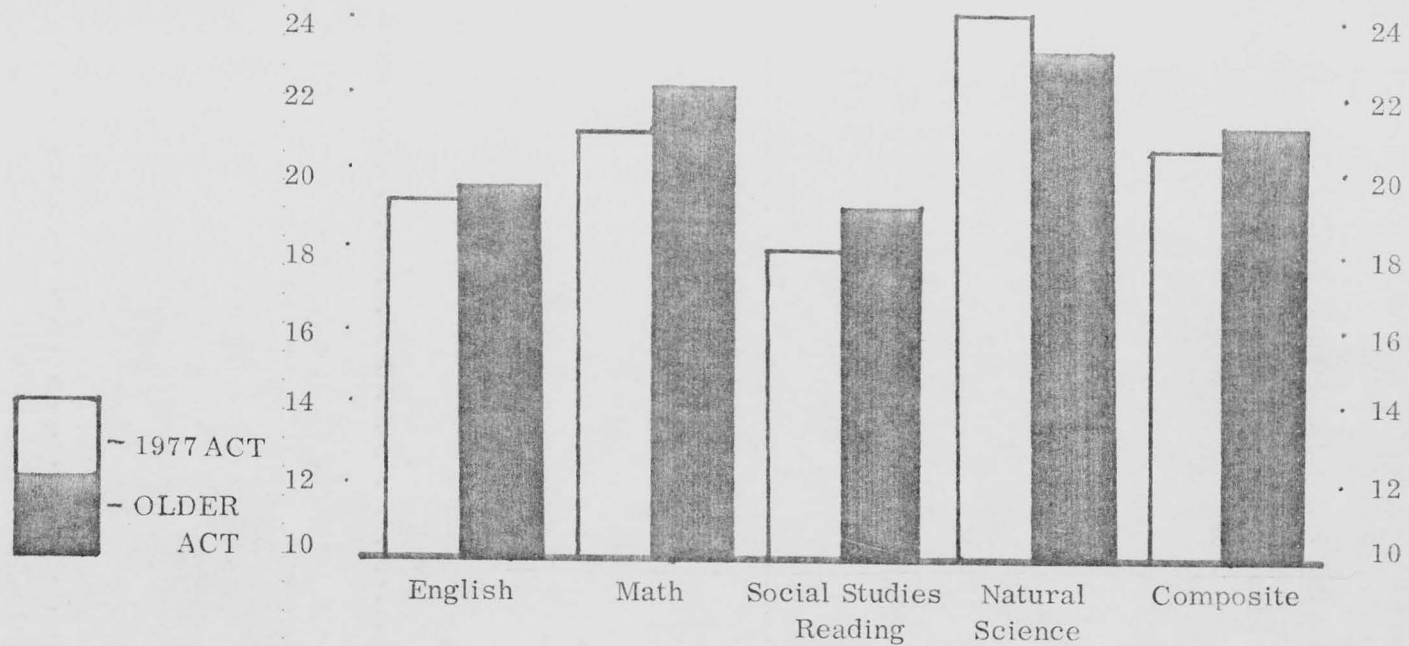
Both F's were significant at the .01 level. None of the interaction F values reached significance indicating that differences in the tests were not sex specific but due to differences in the two test forms.



1977 ACT Means	17.05	15.35	15.84	19.96	17.16
Older ACT Means	17.27	18.65	16.28	20.32	18.26
Difference	-0.22	-3.30	-0.44	-0.36	-1.10
t Value	-1.12	-11.20*	-1.40	-1.33	-7.38*
Two-Tailed Probability	0.265	0.000	0.164	0.184	0.000

\*Significant

Figure 1. Mean scores for Logan High School sample.

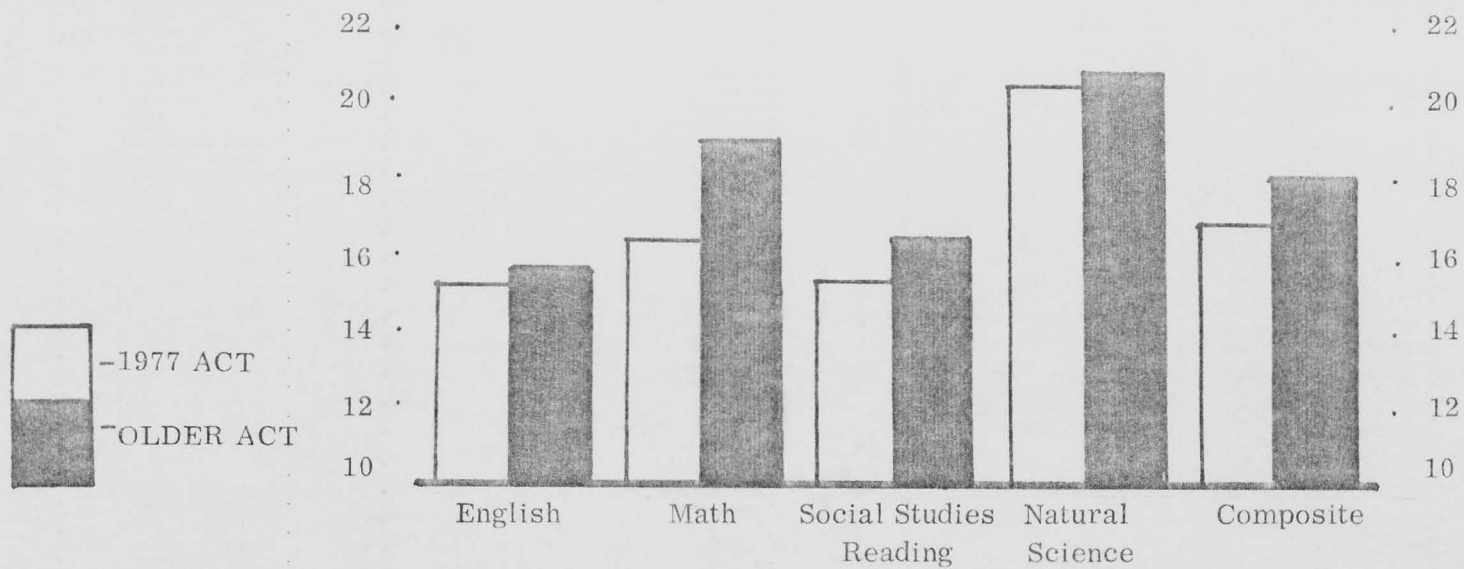


1977 ACT Means	19.25	21.03	18.01	24.25	20.75
Older ACT Means	19.40	22.05	19.08	23.19	21.06
Difference	-0.15	-1.02	-1.07	1.06	-0.31
t Value	-.066	-3.00*	-.262*	3.14*	-1.74
Two-Tailed Probability	0.513	0.002	0.010	0.002	0.084

\*Significant

Figure 2. Mean scores for Skyview High School sample.





1977 ACT Means	15.15	16.27	15.36	20.54	16.94
Older ACT Means	15.62	18.99	16.30	20.87	18.10
Difference	-0.47	-2.72	-0.94	-0.33	-1.16

Figure 3. Mean scores for Logan High School males.



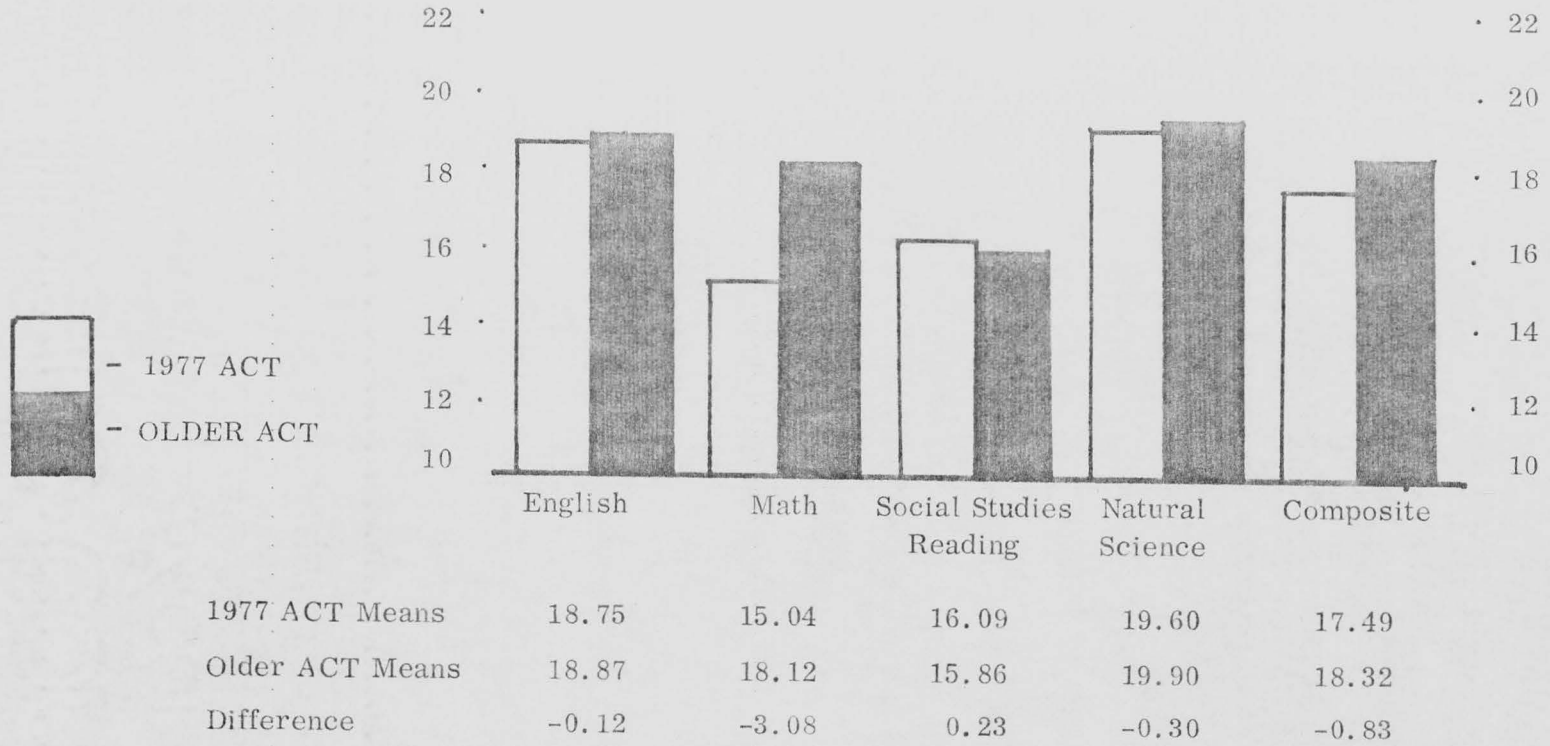


Figure 4. Mean scores for Logan High School females.

Table 1

Two-Way Analysis of Variance Repeated Measures Results on ACT  
 English Subtest Scores for Logan High School  
 Male/Female Sample

Source	df	M.S.	F
Sex	1	938.45	17.99*
Test	1	7.20	1.63
Sex X Test	1	2.45	0.56
Subjects (Sex)	158	52.17	
Subjects X Test (Sex)	158	4.41	

\*Significant (F larger than 3.91 = Significance at the .05 level  
 F larger than 6.81 = Significance at the .01 level)

Table 2

Two-Way Analysis of Variance Repeated Measures Results on ACT  
 Mathematics Subtest Scores for Logan High School  
 Male/Female Sample

Source	df	M.S.	F
Sex	1	88.20	0.93
Test	1	672.80	69.73*
Sex X Test	1	2.81	0.29
Subjects (Sex)	158	94.38	
Subjects X Test (Sex)	158	9.65	

\*Significant (F larger than 3.91 = Significance at the .05 level  
 F larger than 6.81 = Significance at the .01 level)

Table 3

Two-Way Analysis of Variance Repeated Measures Results on ACT

Social Studies Subtest Scores for Logan High School

Male/Female Sample

Source	df	M.S.	F
Sex	1	1.65	0.02
Test	1	10.15	0.82
Sex X Test	1	27.03	2.19
Subjects (Sex)	158	92.10	
Subjects X Sex (Sex)	158	12.35	

\*Significant (F larger than 3.91 = Significance at the .05 level  
F larger than 6.81 = Significance at the .01 level)

Table 4

Two-Way Analysis of Variance Repeated Measures Results on ACT

Natural Science Subtest Scores for Logan High School

Male/Female Sample

Source	df	M.S.	F
Sex	1	73.15	0.96
Test	1	8.13	1.04
Sex X Test	1	0.03	0.04
Subjects (Sex)	158	76.40	
Subjects X Test (Sex)	158	7.82	

\*Significant (F larger than 3.91 = Significance at the .05 level  
F larger than 6.81 = Significance at the .01 level)

Table 5

Two-Way Analysis of Variance Repeated Measures Results on ACT

Composite Scores for Logan High School

Male/Female Sample

Source	df	M. S.	F
Sex	1	12.01	0.19
Test	1	80.00	29.30*
Sex X Test	1	2.11	0.77
Subjects (Sex)	158	63.24	
Subjects X Test (Sex)	158	2.73	

\*Significant (F larger than 3.91 = Significance at the .05 level  
 F larger than 6.81 = Significance at the .01 level)

## CHAPTER V

## DISCUSSION

Evaluation of Findings

Previous studies (Harris, 1976; Harnischfeger & Wiley, 1975) have concluded that the achievement test score declines have had nothing to do with the tests themselves. Data from this study do not support these conclusions in relationship to the ACT.

Logan High School data. In examining the Logan High School data, significant differences were found in the scores obtained on the mathematics subtest between the two forms of the ACT; the difference between means being 3.30 standard scores higher on the older form. The difference between the mean composite scores was also significant. Although no other differences were significant, the mean scores were higher on the older form than on the then current form. The data from the Logan High School sample indicate a possible score drift between the two forms of the ACT used in this research. The greatest amount of downward drift is on the mathematics subtest.

Females had higher scores on the English subtest than males, but both males and females had higher scores on the older form of the test. Male and female scores do not differ significantly on the other subtests. These data indicate that the drifts are not sex specific, but may be due to differences in the tests themselves.

Skyview High School data. The data from the Skyview High School differed in magnitude, but the direction of difference was consistent with Logan's data. The one exception was the natural science subtest scores where students scored higher on the 1977 ACT form. This is consistent with the literature which indicates a slight increase in natural science scores. The data from the Skyview sample, except for the natural science score, is supportive of the Logan High School data.

Summary. Of the hypotheses listed on page 18, the first, "There is no difference in scores obtained on a 1977 form of the ACT and an older (5 to 7 years) ACT form when administered to the same sample of students," was rejected due mainly to the findings on the mathematics subtest. The second, "There is no difference in males' and females' scores obtained on a 1977 form of the ACT and an older (5 to 7 years) form," was rejected, also. However, even though the females' standard scores were significantly higher on the English subtest than the males, both males' and females' scores were higher on the older test.

There appears to be an effect due to the test forms. A downward drift was noted from the older form to the 1977 ACT form. This drift resulted from a large drift on the mathematics subtest. This drift would also indicate a need on the part of ACT to re-examine their equating procedures.

#### Limitations

It was the intention of this study to investigate the effect that the order of testing had upon the scores of the two ACT forms. During the study,

however, data critical to that analysis was not captured, and with it, the ability to deal with the question of order. It was experienced during testing that students attempted to identify the test forms on the basis of dates, sexist statements and the number of metric references within the test. In light of the data showing higher scores on the non-critical test, there are three possible explanations; the students could not identify the important test, the students identified the tests incorrectly, or there is a difference in the two test forms. In administering the older form first to half the Logan High School sample and the 1977 form first to the other half, the effects of order should have been minimized. Due to the lack of data there is no way to directly test the effect that order of testing had on the results.

Another limitation regards the Skyview sample. All of these subjects were volunteers and all of them received the 1977 form first on the regular Saturday test date. Even though precautions were taken and instructions given in such a manner as to decrease the likelihood, it would be scientifically naive to assume that some of the students did not ascertain the order in which the forms were given. These data should be looked at, not as conclusive, but as supportive for the more highly controlled Logan High School sample.

#### Recommendations for Further Research

Research and literature to date has mainly described the declines and offered opinions and speculations as to the causes. Too many possible causes for the score declines have gone untried, while speculation and opinion have



taken the place of objective research. A study of this same basic design, but with better control of the data should be undertaken to check effects due to order of testing. This type of research could also be run on the SAT and the other achievement tests which have shown declines.

## BIBLIOGRAPHY

- Angoff, W. H. Why the SAT scores are going down. English Journal, 1975, 10-11.
- Babcock, B. B. Should we really wonder why the SAT scores are going down? English Journal, 1975, 64 (3), 10-11.
- Breland, H. M. The SAT score decline: A summary of related research. Princeton, New Jersey: Educational Testing Service, January 1976.
- Burns, R. W. Minorities, instructional objectives and the SAT. Educational Technology, June 1976, 39-41.
- Buros, O. K. (ed.) The sixth mental measurements yearbook. Gryphon Press, 1965.
- Champagne, D. W., & Roberts, R. J. An exercise in freedom: A place where test scores appear to be rising. Educational Technology, July 1976, 18-24.
- Edson, C. H. Why scholastic aptitude test scores are falling. OSSC Bulletin, Center for Educational Policy and Management, University of Oregon, 1976.
- Feldhusen, J. F., Hymes, K., & Ames, C. A. Is a lack of instructional validity contributing to the decline of achievement test scores? Educational Technology, July 1976, 13-16.
- Ferguson, R. L. The decline in ACT test scores: What does it mean? Educational Technology, June 1976, 21-27.
- Ferguson, R. L., & Maxey, E. J. Trends in the academic performance of high school and college students. Iowa City, Iowa: ACT Research Report, No. 70, 1976.
- Fields, C. M. Study stirs new criticism of exams. Chronicle of Higher Education, September 6, 1977a, 15, 12.
- Fields, C. M. Why the big drop in SAT scores? Chronicle of Higher Education, September 6, 1977b, 15, 1, 13.

- Harnischfeger, A., & Wiley, D. E. Achievement test score decline: Do we need to worry? CERMELE, Chi., December 1975.
- Harnischfeger, A., & Wiley, D. E. The marrow of achievement test score declines. Educational Technology, June 1976, 5-14.
- Harris, W. V. The SAT score decline: Facts, figures and emotions. Educational Technology, July 1976, 5-12.
- Haugh, O. The standardized test to be or not to be. English Journal, 1975, 64 (3), 53-55.
- Hechinger, F. M. SAT scores: What happened to the best and brightest? Saturday Review World, 1974, 1, 65.
- Kapfer, P. G., Kapfer, M. B., & Woodruff, A. D. Declining test scores: Interpretations, issues and relationship to life-based education. Educational Technology, July 1976, 5-12.
- Maxey, J. E., Wimpey, L. M., Ferguson, R. L., & Hanson, G. R. Trends in the academic abilities, background characteristics and educational and vocational plans of college-bound students: 1970-71 to 1974-75. Iowa City, Iowa: ACT Research Report, No. 74, May 1976.
- McCandless, S. A. The SAT score decline and its implications for college admissions. Paper presented at the 1975 Western Regional Meeting of the College Entrance Examination Board, January 1975.
- Munday, L. A. Declining admissions test scores. ACT Research Report, No. 71, January 1976.
- Rever, P. R., & Kojaku, L. K. Access, attrition, test scores and grades of college entrants and persisters: 1965-1973. Paper presented at the National Conference of the American Association of Collegiate Registrars and Admissions Officers. Minneapolis, Minnesota, April 1975.
- Rippey, R. M. The test score decline: If you don't know where you're going, how do you expect to get there? Educational Technology, July 1976.
- Rorak, A. C. This fall freshmen may have the lowest S.A.T. scores ever. Chronicle of Higher Education, September 6, 1977, 15, 12.

- Russo, T. J. Three ordered sets of factors and their relationship to ACT scores. Unpublished doctoral dissertation, Utah State University, Logan, Utah, 1978.
- Russo, T. J., & Checketts, K. T. Three sets of ordered variables and their association to the American College Test scores. Journal of Educational Research, March/April 1978, 71, 198-202.
- Sapone, C. V., & Giuliano, J. R. The test score decline: Are the public schools the scapegoat? Educational Technology, June 1976, 43-44.
- Scully, M. G. Drop in aptitude test scores is largest on record. Chronicle of Higher Education, September 15, 1975, 1, 18-19.
- Tolman, R. R. A look at achievement test score declines. BSCS Newsletter, Biological Sciences Curriculum Study, 1976.
- Throne, J. M. Has the key to the mystery in standardized tests been discovered? Educational Technology, July 1976, 17-18.
- Zajonc, R. B. Family configuration and intelligence. Science, 1976, 192, 227-235.

