

A TEST OF FREDERIC LORD'S PREMISE RELATIVE TO FORMULA
SCORING

by

Elizabeth Hawkes Koball

Dissertation submitted to the Graduate Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

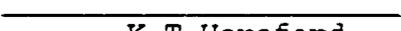
in

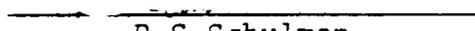
Educational Research

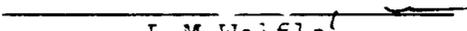
APPROVED:

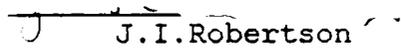

L. h. Cross, Chairman


R. B. Frary


K. T. Hereford


R. S. Schulman


L. M. Wolfle


J. I. Robertson

July, 1988

Blacksburg, Virginia

10-26-88
HGD

A TEST OF FREDERIC LORD'S PREMISE RELATIVE TO
FORMULA SCORING

by

Elizabeth Hawkes Koball

Committee Chairman: Lawrence H. Cross
Educational Research

(ABSTRACT)

Although formula scoring has been used since the early 1900s, it was not until 1975 that Frederic Lord offered the following potential psychometric justification for its use: If under formula-scoring directions an examinee omits only those items which would result in completely random guesses under number-right scoring directions, then the formula score will be a more efficient estimator of the examinee's standing on the trait measured. Whenever the number of omissions is greater than zero, the formula score will be more reliable than the number-right score.

The purpose of this study was to test the premise that examinees omit only those items for which they have no knowledge when taking a test under formula-scoring directions. Several studies had been carried out previously to test this premise, and the design used in this study was a synthesis of the previous designs.

Included in this study was an investigation of examinees' responses, under formula-scoring directions,

to items that were constructed to be obscure. Also examinees responded to questions about their attitudes towards formula-scored tests and their strategies when taking formula-scored tests.

Because of the results of the test of Lord's premise, also included in this study was a further investigation of omissiveness, the tendency to omit items under formula-scoring directions. Item difficulty and item omissions were examined relative to Lord's premise. A variable, called L for convenience, was computed for each item in order to find to what extent responses to test items support Lord's premise. Finally, the possibility of misinformation producing a counter effect to inappropriate omissions relative to Lord's premise was investigated.

ACKNOWLEDGEMENTS

The author wishes to express her genuine appreciation to Dr. Lawrence H. Cross, her advisor, for his assistance, direction, and support. She also greatly appreciates the assistance and insights provided by Dr. Robert B. Frary.

The author wishes to express a special thank you to Dr. James I. Robertson, Jr., for allowing this project to take place in his Civil War History class, and for his assistance in creating the tests that were used.

Further, appreciation is expressed to the other members of her committee: Dr. Karl T. Hereford, Dr. Robert S. Schulman, and Dr. Lee M. Wolfle, for their editing and suggestions for improvement.

The author finds words inadequate to express thanks to her husband, _____, for his tremendous support and encouragement. Finally, she thanks her close friend, _____ for her meticulous editing and her steadfast support.

TABLE OF CONTENTS

	Page
ABSTRACT.....	ii
ACKNOWLEDGEMENTS.....	iv
TABLE OF CONTENTS.....	v
LIST OF TABLES.....	viii
LIST OF FIGURES.....	xi

Chapter

1. INTRODUCTION.....	1
RESEARCH QUESTIONS.....	7
RESEARCH HYPOTHESIS.....	8
2. LITERATURE REVIEW.....	11
3. METHODS.....	17
DATA COLLECTION.....	17
Preliminary Testing Phase.....	17
Measures Obtained in the Preliminary Testing Phase.....	18
Experimental Testing Phase.....	21
ANALYSIS.....	24
Test of Lord's Premise.....	24
Preliminary Analysis.....	25
Primary Analysis.....	25
Risk-taking Behavior on Obscure Items.....	29
Preliminary Analysis.....	29
Primary Analysis.....	30

Chapter	Page
Response Strategies.....	31
4. RESULTS.....	32
TEST OF LORD'S PREMISE.....	32
Preliminary Analysis.....	32
Primary Analysis.....	38
RISK-TAKING BEHAVIOR ON THE OBSCURE ITEMS.....	46
RESPONSE STRATEGIES.....	52
RECAP OF THE RESULTS OF THE TESTS OF LORD'S PREMISE.....	60
POST HOC ANALYSES.....	61
Omissiveness.....	61
Analyses Based on Subsets of Final Examination Items.....	72
Investigation of Lord's Premise by Items.....	73
5. CONCLUSIONS	80
REFERENCES.....	86
APPENDICES.....	88
A. Instructions for Midterm Examinations.....	88
B. Instructions for Fall Final Examination.....	90
C. Fall Final Examination.....	92
D. Type I Obscure Questions.....	102
E. Type II Obscure Questions.....	104

Chapter	Page
F. Addendum to Fall Final Examination..	106
G. Winter Final Examination.....	108
H. Addendum to Winter Final Examination.....	119
I. Oral Instructions for Winter Final Examination.....	122
VITA.....	124

LIST OF TABLES

Table	Page
1. Analysis of Variance Design for Testing the Effect of Question Set and Order on CNR and CFG Scores.....	27
2. Repeated Measures Analysis of Variance Design for Testing the Effect of Omissiveness on CNR and CFG Scores...	28
3. Descriptive Statistics and Reliability Estimates for Experimental Test by Form.....	33
4. Descriptive Statistics of CNR by Question Set and Order of Instructions	34
5. ANOVA Summary for Question Set and Order Effect on CNR Scores.....	35
6. Descriptive Statistics of CFG by Question Set and Order of Instructions	36
7. ANOVA Summary for Question Set and Order Effect on CFG Scores.....	37
8. Descriptive Statistics of CNR and CFG by Category of Omissiveness for Forms I and II Combined.....	39
9. Repeated Measures ANOVA Summary Table of CNR and CFG by Category of Omissiveness for Forms I and II Combined.....	40
10. Descriptive Statistics of CNR and CFG by Category of Omissiveness for Forms III and IV Combined.....	43
11. Repeated Measures ANOVA Summary Table of CNR and CFG by Category of Omissiveness for Forms III and IV Combined....	44
12. Frequencies of Responses to Type I Obscure Items.....	47
13. Frequencies of Responses to Type II Obscure Items.....	48

Table	Page
14. Cross Tabulation of Risk I with Responses to Strategy Question on Fall Final Examination.....	50
15. Cross Tabulation of Risk II with Responses to Strategy Question on Fall Final Examination.....	51
16. Cross Tabulation of Gender by Omissiveness Category.....	54
17. Cross Tabulation of Advice from Class in Test Taking Skills with Category of Omissiveness.....	55
18. Cross Tabulation of Responses to Item 61 with Category of Omissiveness.....	56
19. Cross Tabulation of Responses to Item 62 with Category of Omissiveness.....	57
20. Cross Tabulation of Responses to Item 65 with Category of Omissiveness.....	58
21. Cross Tabulation of Responses to Item 69 with Category of Omissiveness.....	59
22. Cross Tabulation of Omissions on Final Examination with Category of Omissiveness Obtained from Midterm Examination for Forms I and II.....	63
23. Cross Tabulation of Omissions on Final Examination with Category of Omissiveness Obtained from Midterm Examination for Forms III and IV.....	64
24. Descriptive Statistics of CNR and CFG for Consistent Omitters and Non-omitters for Forms I and II Combined..	66
25. Repeated Measures ANOVA Summary Table of CNR and CFG by Category of Omissiveness for Forms I and II.....	67
26. Descriptive Statistics of CNR and CFG for Consistent Omitters and Non-omitters for Forms III and IV.....	69

Table

Page

27.	Repeated Measures ANOVA Summary Table of CNR and CFG by Category of Omissiveness for Forms III and IV.....	70
28.	Item Information.....	76

LIST OF FIGURES

Figure	Page
1. Plot of Mean CNR and CFG by Category of Omissiveness for Forms I and II....	41
2. Plot of Mean CNR and CFG by Category of Omissiveness for Forms III and IV..	45
3. Plot of Mean CNR and CFG by Category of Omissiveness for Forms I and II, Based on Consistency of Omissiveness..	68
4. Plot of Mean CNR and CFG by Category of Omissiveness for Forms III and IV, Based on Consistency of Omissiveness..	71
5. Histogram of Variable L.....	75

Chapter 1

INTRODUCTION

When multiple-choice tests were introduced in the 1920s, widespread concern developed over the extent to which guessing could influence scores. The following hypothetical situation illustrates the basis for this concern. Suppose two examinees take a 100-item, five-choice, multiple-choice test. Further suppose that each knows the answer to 50 items and is ignorant concerning the other 50. One examinee marks the remaining 50 at random, and the other decides not to respond to these items. The former would be expected to answer correctly about 10 more items than his reticent counterpart. Scores based on the number of right answers would then lead to a false conclusion concerning the difference in the two examinees' abilities. Naturally, there was a desire to prevent this outcome.

One approach to the problem just outlined would be to encourage all examinees to guess whenever they did not know an answer or lacked time to read and consider an item. This practice would tend to produce equivalent scores for cases such as the two examinees described above. However, early proponents of multiple-choice tests expressed concern that this approach would have a

deleterious effect on public confidence in multiple-choice testing. They feared that test scores would then be viewed as reflecting luck rather than knowledge, and that instructions to guess always in the absence of knowledge would encourage carelessness and engender poor attitudes among examinees. (See Davis, 1964, or Lindquist, 1968, for a recounting of this argument.)

An alternative to advising students to guess always in the absence of knowledge is to admonish examinees not to guess and to include a score deduction for gains due to guessing. Thus, the "correction-for-guessing" was introduced. It was believed that if students were informed that points would be lost for incorrect guesses, they would refrain from answering when in doubt. By eliminating guessing, the scores would be a more valid measure of an examinee's knowledge. Further, by eliminating random guessing, a source of error variance would be removed and so the scores would be more reliable.

The conventional correction-for-guessing score, also called formula score, consists of subtracting a fraction of the number of incorrect responses from the number of correct responses on a multiple-choice test. The formula is: $S = R - W/(C-1)$, where:

R = number right; W = number wrong; and

C = number of choices for each item.

It should be noted that, in order for formula scoring to have the intended effect, examinees must be advised of its use and be informed that random guessing cannot be expected to enhance scores.

The logic of this scoring rule can best be illustrated by considering its effect in another hypothetical situation. Suppose an examinee responds at random to an 80 item test consisting of five-choice items. The expected outcome would be 16 items correct and 64 items wrong. The formula score would be

$$S = 16 - 64/4 = 0.$$

Based on the premise that an examinee has either certain knowledge or no knowledge, what results is, in effect, a theorem. If an examinee marks an answer correctly, omits, or guesses at random, then the formula provides an unbiased estimate of the number known.

Until 1975, when Frederic Lord advanced another theorem, no other basis for justifying the use of formula scoring had been advanced, though it was clear from the outset that examinees have varying degrees or states of knowledge with respect to multiple-choice items. Indeed, under formula scoring, it is to an examinee's advantage to guess whenever it is possible to eliminate even one wrong choice from consideration. To illustrate this, consider a third hypothetical situation. Suppose an

examinee can correctly eliminate exactly one wrong choice on each item of an 80 item test consisting of five-choice items. The expected outcome would be 20 items correct and 60 items wrong. The formula score would be:

$$S = 20 - 60/4 = 5.$$

In fact, the expected score increases as the number of wrong choices eliminated increases.

Therefore, the instructions for a formula-scored test should advise the examinees of the appropriate strategy, namely, to omit only when the alternative is a completely random guess among all choices. Instructions were often not this forthright in the early years of formula scoring. The instructions led some, but not all, examinees to believe that they should guess under no circumstance. That the instructions could be misleading was eventually recognized. Today formula-scored tests generally advise examinees to guess if they can eliminate at least one distractor.

Because of the potential for unfairness and the lack of an effective psychometric justification for its use, formula scoring was frequently attacked from within the psychometric community. Thus it was of considerable interest when Lord (1975) advanced a potentially valid justification for formula scores. Based on the premise that omissions under formula scoring would be replaced by

responses that are truly random under number-right scoring, Lord proved that formula scores will be more efficient ability estimates than the number-right scores. More efficient estimates lead to more reliable scores. A potentially valid justification for formula scoring was provided. To confirm this justification (for any particular testing situation) it would only be necessary to show that examinees' omissions under formula scoring would, indeed, be replaced by random guesses among all choices under number-right scoring.

To illustrate Lord's premise consider one more hypothetical situation. This time suppose that there are four examinees taking the 80 item multiple choice examination consisting of five-choice items. Suppose that each examinee knows the correct choice for 24 items; can eliminate no distractors for 20 items; can eliminate exactly one distractor for the other 32 items; and is misinformed on four items, having classified the answer and one other option as distractors for each item. Under number-right scoring instructions all three examinees respond to all of the 80 items. The expected score for each then is

$$24 + (1/5)20 + (1/4)32 = 36.$$

Under formula scoring instructions, however, suppose the first examinee omits no responses, the second exami-

nee omits only those items for which he can eliminate no distractors, the third examinee omits all items for which he can eliminate no more than one distractor, and the fourth examinee omits all items of which he is not certain. The expected formula score for the first examinee is

$$36 - (1/4)44 = 25.$$

The expected formula score for the second examinee is

$$32 - (1/4)28 = 25.$$

As Lord's premise suggests, no matter how many of the "no knowledge" items are omitted, the expected formula score is not affected. One can see, further, that were the correction-for-guessing formula applied to the responses obtained under number-right scoring instructions, that an unbiased estimate of the expected formula score would be obtained.

The third examinee's expected formula score, however, is

$$24 - (1/4)4 = 23$$

because he has omitted some items for which he has some knowledge. His actual formula score is less than the estimate obtained from the test taken under number-right scoring instructions. Interestingly, the fourth examinee, who omits even more often than the third, has an expected score higher than the third examinee, but still

lower than his optimal score. His expected score would be 24. Although he has omitted some items for which he has partial knowledge, the loss of points has been offset by the omission of items on which he is misinformed.

Lord challenged the psychometric community to investigate the extent to which the third and fourth situations occur in "real life." At the same time, he suggested that to the extent they do occur, examinees might be taught appropriate behavior with respect to omitting. However, to the extent that the third or fourth situation does occur, there is potential for bias in using formula scores for estimating ability.

Research questions

Lord's challenge raises several questions which were the object of this investigation. These are:

1. Are there examinees for whom the premise of Lord's theorem is not true? In particular, if examinees have a tendency to omit responses to items on formula-scored, multiple-choice tests, does the assumption hold that they will omit only those questions that would have resulted in random guesses among all choices under number-right scoring instructions?

2. What percentage of examinees will guess when no choices can be eliminated as incorrect? What percentage

will guess when exactly one distractor can be eliminated with near certainty?

3. Are there particular attitudes or beliefs about formula scoring that tend to accompany the tendency of examinees to omit responses under formula scoring instructions? The tendency to omit items will be referred to hereafter as omissiveness

Research hypothesis

The purpose of this study was to test Lord's premise that examinees' omissions under formula scoring instructions would be replaced by random guesses among all choices under number-right scoring instructions. The * research hypothesis follows: examinee's omissions under formula scoring instructions would not be replaced by random guesses among all choices under number-right scoring instructions.

The primary approach was to obtain two scores for each examinee: one under number-right scoring instructions and the other under formula scoring instructions. The correction-for-guessing formula was applied to the scores obtained under both sets of directions. Those obtained under number-right instructions are referred to as corrected-number-right (CNR) scores. Those obtained under formula scoring instructions are referred to as

corrected-for-guessing (CFG) scores. This investigation focused on the extent to which CNR and CFG scores are different for each examinee. If Lord's premise is correct, there should be no differences due to omissions.

Secondary investigations

Secondarily, examinees were presented with items, under formula scoring instructions, for which the probability of eliminating even one distractor was very small, in order to see the extent to which examinees were willing to guess when faced with no information. Further they were presented with items for which the probability of eliminating exactly one distractor correctly was nearly certain, to see the extent to which examinees would be willing to guess under this condition.

Finally, examinees were questioned about their beliefs and strategies relative to guessing. The purpose of this was to investigate the relationship between omissiveness and some factors unrelated to knowledge.

Given that multiple-choice tests are a frequently used tool to measure both achievement and aptitude, it is essential that the resulting scores be as reliable and valid as possible. To the extent that random guessing decreases, one can expect greater reliability. However it is essential that the scoring and the accompanying

instructions which are designed to decrease random guessing do not incur a bias unrelated to the knowledge being tested. This is especially important since scores on multiple-choice tests are a major criteria in the selection process for entrance into many of America's institutions of higher learning and for the awarding of scholarships to these institutions.

Chapter 2

LITERATURE REVIEW

Previous to the publication of Lord's theorem, it would seem that his challenge had been anticipated. Many studies of the relationship of an examinee's omission of items to factors other than knowledge were undertaken. In general, these studies demonstrated that some examinees omit items even when they have some knowledge that would lead to a correct response. (See Cronbach, 1946; Sherriffs and Boomer, 1954; Price, 1964; Slakter, 1968; Diamond and Evans, 1973.)

Lord challenged all of the previous studies because the instructions did not include directions that would inform examinees of the guessing strategy that would maximize their scores, while at the same time discouraging random guessing. He pointed out that although examinees might have been informed that formula scoring was the method being used, that they probably not been instructed how best to respond under that scoring method.

The Slakter (1968) study was of particular interest, in that it was the prototype of several studies that followed Lord's challenge. In the Slakter study the examinees had filled in responses to test items administered

under formula scoring instructions. When the test was completed, the examinees were asked to fill in responses to items that they had omitted during the formal testing period. Besides the aforementioned challenge by Lord, Slakter's study was also challenged because examinees had been allowed time enough to discuss the test items before filling in the omitted responses.

Cross and Frary (1977) designed a study to meet Lord's challenge and to test his premise. They provided instructions for a test that included the proper strategy for maximizing one's score under formula scoring. They had examinees in a college chemistry course fill in the responses to items they had previously omitted under formula scoring instructions, without allowing time for discussion of the test before the second responses were recorded. They found, contrary to Lord's premise, performance on previously omitted items was better than would be expected by chance. Bliss (1977) replicated their study with elementary school children and obtained the same result. In both studies, it was found that the scores of examinees, who reportedly complied with instructions to guess only if they could eliminate at least one choice, were most negatively affected by omissions under formula scoring instructions. These latter studies were challenged, however, by Angoff and Schrader

(1984) because the condition of responding to previously omitted questions was considered not to be realistic enough to generalize results to a true testing situation.

Angoff and Schrader (1984) approached the problem differently. They looked at differences between groups of examinees who took tests under different sets of instructions, number-right instructions for one group and formula instructions for the other. They compared formula (CFG) scores to corrected-number-right (CNR) scores. Corrected-number-right scores are those scores obtained under number-right instructions, to which the correction-for-guessing formula has been applied. They found no significant differences between the CNR and CFG scores of the different groups. They, therefore, interpreted their results as supporting Lord's premise and providing justification for formula scoring.

Their study also has been challenged on several bases. Albanese (1986) noted that volunteers were used for the study and challenged the validity of this in this particular context. While volunteers can be used for many studies, the use of volunteers in this case was questionable. This is because the examinee must perceive that his test performance will have a real effect on his life if the penalty for guessing is to deter his guess-
ing.

More important, a between groups design, in itself, does not provide a direct test of Lord's premise. The question is not whether formula instructions produce results equivalent to number-right instructions across all examinees on average; but rather whether there is a subset of examinees, however small that subset may be, who because of their test-taking behavior are differentially penalized by their response to the instructions.

Frery, Cross, and Sewell (1985) designed a study to meet the above challenges. As Angoff and Schrader, they compared formula (CFG) to corrected-number-right (CNR) scores. Frery, Cross, and Sewell, however, used a within subjects design, wherein CNR and CFG scores were obtained for each examinee for comparable test halves, administered under different sets of directions. Further, they used non-volunteer subjects, i.e. their course grades were based on the scores being studied. They classified their subjects according to the number of items omitted on the formula scoring directions test half. They found that students who omitted items had CFG scores that were significantly lower than their CNR scores, which could have been because they underestimate the degree of their knowledge or because they are less willing to guess when they are not certain of the correctness of their response. Frery, Cross, and Sewell also found that stu-

dents who reported using a more conservative strategy under formula scoring, as opposed to number-right scoring, had significantly lower CFG scores than CNR scores. It is particularly noteworthy that 49% of the examinees were in this category.

Their design was challenged, however, because they used a measure of omissiveness that was obtained on the same test for which the comparison of CFG scores to CNR scores was made. The lack of experimental independence may have confounded their results. To date, then, Lord's challenge remains unmet. To date, the attempts to meet the challenge have focused on a possible subset of examinees who may be adversely affected by formula scoring. In particular, the concern has been that this subset of examinees may omit items under formula scoring instructions, for which they have a better than chance probability of responding correctly.

Frary (1988) performed a simulation study, a major feature of which was distinguishing inappropriate omissions when informed from those made when misinformed. The results suggested that individuals who omit items inappropriately are likely to omit items under both conditions. In this case, a loss of points due to omission of items for which the examinee had partial knowledge would be offset by not losing points due to omission of

items for which the examinee was misinformed. An example of this effect was illustrated in the introduction (page 6). This result emphasizes two points: Losses due to inappropriate omissions will occur largely in the absence of misinformation; and an obstacle to identifying losses due to inappropriate omissions can be the presence of misinformation.

The present study focused on the possible subset of examinees who may omit items inappropriately on a formula scored test. As in the Angoff and Schrader (1984) study, the CNR and CFG scores were compared. As in the Frary, Cross, and Sewell (1985) study, a within subjects design was used, and the subjects were not volunteers. The measure of omissiveness, however, was obtained on a test taken previous to the experimental testing phase. Finally, the examinees were informed on four occasions of the guessing strategy that would maximize their scores, while discouraging random guesses.

Chapter 3

METHODS

This was a study of omissions on a multiple-choice test that had been administered under correction-for-guessing scoring instructions. The primary question was whether or not examinees omit only items for which they have no knowledge.

Data Collection

The subjects for this study were 660 students who enrolled in a junior level Civil War history sequence, not open to freshmen, at Virginia Tech. The three quarter sequence was offered during the academic year, 1987-88. During the two quarters of the sequence included in the study, four examinations were given: a fall midterm, a fall final, a winter midterm, and a winter final. The winter final examination was the experimental testing phase of the study. A score from an examination was included in the analyses only if the examination was taken in class during the regularly scheduled examination period.

Preliminary testing phase

The midterm examinations of the fall and winter quar-

ters each consisted of 54 four-choice multiple choice items. Oral instructions for these examinations included a brief statement indicating that formula scoring was the scoring method in effect. The formula scoring directions indicating the best strategy for responding to test items were printed on the front of the examinations (Appendix A). The final examination of the fall quarter, consisting of 67 multiple choice items, was also administered under formula scoring instructions. For this examination, however, the instructions were read to the examinees and were projected onto a screen for them to read (Appendix B). The purpose of administering these tests under formula scoring directions was to give students practice and feedback in taking tests under formula scoring instructions.

Measures obtained in the preliminary testing phase

Omissiveness.

Omissiveness (tendency to omit) is defined as the total number of responses omitted by an examinee on the winter midterm examination. In previous studies (Cross and Frary, 1977; Angoff and Schrader, 1984; Frary, Cross, and Sewell, 1987) this variable has been positively skewed. Approximately 25% of the examinees or more

omitted no items on even very difficult tests, and then the frequencies decreased to a very small percentage omitting as many as half of the responses to test items.

Risk-taking behavior.

Risk-taking behavior was assessed by including two types of "obscure" items on the fall final examination. These "obscure" items were constructed to measure knowledge of material so obscure that it was assumed that virtually no examinee would know the correct answer. The material for these items was obtained from Long (1971), Spencer (1986), and Wakelyn (1977). Four of the obscure items had one choice designed to be blatantly incorrect (Appendix D). Those obscure items constructed to have an obvious distractor are labeled Type I. An example of a Type I obscure item follows:

- On April 25, 1861, U. S. troops surrendered at:
1. Carthage, Missouri.
 2. Rockport, Arkansas.
 3. Saluria, Texas.
 4. Boise, Idaho.

For these items to function as intended, examinees should recognize the obviously incorrect choice, eliminate it, and provide a random guess among the remaining choices.

The remaining four obscure items were constructed to have no response choice that was deemed to be an obvious distractor (Appendix E). These are labeled Type II.

An example of a Type II item follows:

Lincoln's message to Governor Pickens, of South Carolina, that Fort Sumter would be resupplied was delivered by:

1. John L. Worden.
2. Leroy P. Walker.
3. Robert S. Chew.
4. James L. Chesnut.

For these items to function as intended, examinees should omit these items if they are to comply with the formula scoring directions which discourage guessing if a choice cannot be eliminated as a distractor.

The eight obscure items were distributed randomly among the regular test items. There were two forms of the Fall final examination. The response choices for each of the obscure items were ordered differently on the two forms of the test in order to prevent the possible confounding factor of preference for a particular choice position; i.e. 1, 2, 3, 4. It was intended that these items appear to be standard items.

Risk.

Risk is defined as the number of obscure items that an examinee did not omit. Formula scoring is designed to encourage examinees to omit items when no distractor can be eliminated as implausible. On the other hand, if one or more choices can be eliminated, it is to the examinee's advantage to guess among the remaining choices. Therefore, Risk was broken down into two variables:

Risk I for the Type I obscure items and Risk II for Type II obscure items. In order to examine the extent to which examinees, in general, respond to the two types of obscure items, the percent of examinees who responded to these items was also computed.

Appended to the end of the fall final examination were three questions designed to ascertain the examinees' strategies for guessing under formula scoring instructions (Appendix F). In light of the results reported in the next chapter, only one of these, question 69, was used in the analysis.

Experimental testing phase

The final examination of the winter quarter, the experimental testing phase of the study, consisted of three parts. One part consisted of 30 items preceded by number-right instructions, and another part consisted of 30 items preceded by formula scoring instructions. The final part consisted of items designed to ascertain the examinees' opinions about guessing on multiple-choice tests, past instruction or advice that they had received relative to guessing on multiple-choice tests, and the examinee's gender.

In order not to confound the effects of the instructions with the order in which the instructions were

given, four forms of the final examination were produced with separate written instructions on each part (Appendix G). The forms were ordered as follows:

- Form I: part A - number-right scoring instructions and question set 1
part B - formula scoring instructions and question set 2
- Form II: part A - formula scoring instructions and question set 2
part B - number-right scoring instructions and question set 1
- Form III: part A - formula scoring instructions and question set 1
part B - number-right scoring instructions and question set 2
- Form IV: part A - number-right scoring instructions and question set 2
part B - formula scoring instructions and question set 1.

The following steps were taken to create question sets of equal difficulty:

1. Items from final examinations administered in previous years were obtained and sorted according to difficulty level, as computed from item analyses.
2. New items were generated from lecture notes and from the textbook. Four people, a Civil War expert, two Civil War buffs, and a recent college graduate who had a major in history, reviewed the items for difficulty.
3. All items were sorted according to the topic covered, the source of the question, and perceived difficulty. Each set of questions contained

equal numbers of items by topic, source, and perceived difficulty.

4. The correct choices for each question set consisted of eight '1' responses, seven '2' responses, eight '3' responses, and seven '4' responses. The questions were randomly ordered so that the correct responses (1,2,3,4) were randomly ordered.

The final part (C) of the winter final examination consisted of ten questions (Appendix H) designed to ascertain the following:

1. the examinees' strategies for guessing under formula scoring instructions;
2. their beliefs about guessing and formula scoring;
3. information examinees had received outside of this study relative to guessing under formula scoring and the source of this information;
4. each examinee's gender.

After being collated, the four forms of the examinations were distributed alternately. While this method is admittedly not a random assignment, there was no reason to suspect bias. It was decided that this was the best way to distribute the tests efficiently.

All examinees were informed orally at the outset of the examination period that there were different instruc-

tions on each part of the examination (Appendix I). For grading purposes, the responses to parts A and B were scored in a manner consistent with the corresponding directions for that part. For research analyses, the correction-for-guessing formula was applied to scores for both parts of the examination, thus yielding the CNR and CFG scores.

Of the 660 students who enrolled in this sequence, 200 took all four examinations, 290 took just the two examinations of the winter quarter, 495 (the 200 and 290 just mentioned and an additional five who had not taken the winter midterm examination during the regularly scheduled period) took the winter final examination, and 360 (the 200 mentioned above and an additional 160 who did not enroll during the winter quarter of the sequence) took the fall final examination.

Analysis

Part I - Test of Lord's Premise

The question being addressed in this part of the study was: If examinees have a tendency to omit responses to items on formula-scored, multiple-choice tests, does the assumption hold that they will omit only those questions that would have necessitated a random guess among all choices under number-right scoring

instructions? The basis for comparison was responses recorded under number-right scoring instructions and correction-for-guessing scoring instructions.

Preliminary analysis.

The following preliminary questions were addressed:

- a. Are question sets 1 and 2 equivalent?
- b. Is there an effect relative to the order in which the instructions were given?

These two questions were addressed using two 2X2 factorial analyses of variance (ANOVA), one on CNR scores and the other on CFG scores. Independent variables were question set and order of instructions. There were two levels for each of these variables. The design is shown in Table 1. Since retaining the null hypothesis was the desired outcome of these tests, the conservative alpha level of .20 was used for these tests.

Primary analysis.

Paired scores (CNR,CFG) for each examinee were used for the primary analysis. The CNR and CFG scores were analyzed using a two-way, repeated measures, factorial ANOVA. The between subjects factor was the measure of omissiveness obtained on the winter midterm examination. Omissiveness was reduced to an ordered categorical vari-

able. The categories were as follows: 0 omissions, 1 - 3 omissions, and more than 3 omissions. The within subjects factor was the response instructions.

If Lord's assumption is tenable, the expected value of (CNR - CFG) should be zero for all examinees. If his premise is not tenable the alternative hypothesis is that for those examinees who omit responses on the winter examination the value, (CNR - CFG), will be significantly greater than for those examinees who do not omit responses. If it could be concluded that the question sets were of equal difficulty then the expected value of CNR - CFG must equal zero for examinees who omit no items on that examination. In this case the potential to observe differences between CNR and CFG scores will be among students who omit items under formula scoring instructions. If on the other hand, the question sets are not of equal difficulty, it would be reasonable to assume that differences between CNR and CFG for the non-omitters would merely reflect the differences in the question sets. In this case, the non-omitters essentially become the control group with whom to compare the omitters.

The design for this analysis is shown in Table 2. The hypothesis of particular interest was the interaction. An alpha level of .05 was used for these tests.

Table 1

Analysis of Variance Design for Testing the Effect of Question Set and Order for CNR and CFG Scores

	Order of Instructions	
	number right precedes formula	formula precedes number right
question set 1	form I part A CNR	form II part B CNR
question set 2	form IV part A CNR	form III part B CNR

	Order of Instructions	
	number right precedes formula	formula precedes number right
question set 1	form IV part B CFG	form III part A CFG
question set 2	form I part B CFG	form II part A CFG

Table 2

Repeated Measures Analysis of Variance Design for Testing the
Effect of Omissiveness on CNR and CFG Scores

	INSTRUCTIONS	
	number right instructions	formula instructions
(1) 0 omissions	<u>CNR scores</u>	<u>CFG scores</u>
(2) 1 - 3 omissions	<u>CNR scores</u>	<u>CFG scores</u>
(3) > 3 omissions	<u>CNR scores</u>	<u>CFG scores</u>

Part II - Risk-Taking Behavior on Obscure Items

The question being addressed is: What is the extent of guessing when examinees have a) no knowledge about an item and b) when they can eliminate exactly one choice with near certainty? There were 360 students who were included in this analysis.

Preliminary analysis.

The first question to be addressed was whether the selection of the correct response to a Type I obscure item could reasonably be attributed to chance and whether the blatantly incorrect response could be deemed obviously incorrect. If the Type I items are truly obscure, with one of the distractors being obviously incorrect, then it would be expected that by eliminating the obvious distractor that the probability of selecting the correct choice would be 1/3 and the probability of selecting an incorrect choice would be 2/3, thus yielding the following expected score for these four items:

$$E(S) = (1/3)(4) - (1/3)(2/3)(4) = 4/9.$$

The following hypothesis was tested:

$$H_0 : \mu_{\sigma_2} - 4/9 = 0$$

$$H_1 : \mu_{\sigma_2} - 4/9 \neq 0$$

Since the desired outcome was to accept the null hypothesis, alpha = .20 was the criterion for rejection.

For the Type II obscure questions, where no distractor can be eliminated, the probability of selecting the correct choice would be expected to be $1/4$ and the probability of selecting an incorrect choice would be expected to be $3/4$, thus yielding the following expected score:

$$E(S) = (1/4)(4) - (1/3)(3/4)(4) = 0.$$

So the following hypothesis was tested:

$$\begin{aligned} H_0 : \mu_{s_i} &= 0 \\ H_1 : \mu_{s_i} &\neq 0 \\ \alpha &= .20 \end{aligned}$$

Only those sets of items, if any, for which the null hypothesis was accepted were used for the following analysis.

Primary analysis.

In order for CFG scores to be psychometrically superior to number right scores, examinees should omit responses to Type II items and should respond to Type I items. The condition of responding to a Type II item, although not ideal, would be tolerable. As Lord has demonstrated, the most reliable score is obtained when completely random guesses are eliminated. However, the omission of Type I items creates the potential for bias. The potential for bias exists when items are omitted for which there is a better than chance probability of choosing the correct response. A 95% Confidence Interval was

computed for the percent of responses to the Type I obscure items and to the Type II obscure items, only if the sets of items met the criteria defined above.

Part III - Response Strategies

The following specific questions were addressed by means of hypotheses testing:

1. Is there a relationship between gender and omissiveness?
2. Is there a relationship between taking special classes on test taking skills and omissiveness?
3. Are attitudes towards formula scoring and omissiveness independent factors?

This information was obtained on the final examination of the winter quarter. All examinees who took the winter final examination composed a single sample of 495 for this part of the analysis.

The analysis of this data was primarily descriptive. In particular, cross-tabulations of responses to the questions in part C of the final examination and categories of omissiveness (obtained from the midterm examination) are provided.

Chapter 4

RESULTS

Table 3 contains the descriptive statistics for each form of the final examination. For the entire class (n = 495) the mean number-right score was 18.07, out of a possible score of 30; and the mean formula score was 13.98, out of a possible score of 30. The mean score of the whole test was 32.05, out of a possible score of 60. Application of Spearman-Brown formula to Cronbach's alpha reliability coefficients for the 30 item half-length tests suggests a total test reliability ranging from .67 to .79.

Part I - Test of Lord's Premise

Preliminary Analyses.

The results of the preliminary tests of question set and order effect are found in Table 4 through Table 7.

Under number-right instructions, it was concluded that there was neither a question set effect nor an order effect. Two tests of these hypotheses were done: one on number-right scores and one on CNR scores. Since there were very few (19) omissions, the conclusions were the same for both tests and only NR scores are reported.

Table 3

Descriptive Statistics and Reliability Estimates for
Experimental Test by Form N = 495

Form	Question Set	Scoring Method	N	Mean	SD	alpha
I	1	NR	125	17.84	4.06	.60
	2	CFG		14.36	4.53	.60
II	1	NR	125	18.02	4.00	.57
	2	CFG		14.92	4.47	.57
III	2	NR	123	17.89	3.84	.55
	1	CFG		13.27	5.58	.65
IV	2	NR	122	18.52	3.49	.50
	1	CFG		13.34	5.18	.60
All		NR	495	18.07	3.85	
		CFG		13.98	4.99	
		Total		32.05	7.84	

Table 4

Descriptive Statistics of CNR by Question Set and Order
of Instructions N = 495

Question Set	Order of Instr	Form	Part	N	M	SD
1	NR,CFG	I	A	125	13.80	5.37
1	CFG,NR	II	B	125	14.08	5.32
2	NR,CFG	IV	A	122	14.70	4.60
2	CFG,NR	III	B	123	13.91	5.14

Table 5

ANOVA Summary for Question Set and Order Effect on CNR
Scores N = 495

Source	df	SS	MS	F	F prob
Set	1	16.73	16.73	< 1	
Order	1	8.18	8.18	< 1	
Set by Order	1	35.70	35.70	1.36	.24
Error	491	12858.59	26.19		

Table 6

Descriptive Statistics of CFG by Question Set and Order
of Instructions N = 495

Question Set	Order of Instr	Form	Part	N	M	SD
1	NR, CFG	IV	B	122	13.34	5.18
1	CFG, NR	III	A	123	13.27	5.58
2	NR, CFG	I	B	125	14.36	4.53
2	CFG, NR	II	A	125	14.92	4.47

Table 7

ANOVA Summary for Question Set and Order Effect on CFG
Scores N = 495

Source	df	SS	MS	F	F prob
Set	1	220.10	220.10	8.95	.003
Order	1	7.25	7.25	< 1	
Set by Order	1	12.51	12.51	< 1	
Error	491	12076.69	24.59		

When the question sets were administered under formula scoring instructions, the mean score for question set 1 was significantly less than the mean score for question set 2 (set 1: $M = 13.31$; set 2: $M = 14.64$; $p\text{-value} = .003$). The interesting point about this result is that there was not a significant difference in the mean number of omissions between the two sets of questions (set 1: $M = 2.85$, $SD = 2.78$; set 2: $M = 2.75$, $SD = 2.80$). There was not evidence of an order effect.

Because of the above results, the forms of the final examination were separated into two samples for the primary analysis:

1. Forms I and II, on which question set 1 had been administered under number-right scoring instructions and question set 2 had been administered under formula scoring instructions;

2. Forms III and IV, on which question set 1 had been administered under formula scoring instructions and question set 2 had been administered under number-right scoring instructions.

Primary Analysis.

The results of the analysis of Forms I and II are shown in Table 8, Table 9, and Figure 1. The mean CFG scores were significantly greater ($p\text{-value} < .03$) than

Table 8

Descriptive Statistics of CNR and CFG Scores by Category
of Omissiveness for Forms I and II Combined N = 246

Category	N	CNR question set 1		CFG question set 2		total
		M	SD	M	SD	M
1						
0 omits	81	14.56	5.55	15.15	4.59	14.86
2						
1-3 omits	76	15.12	5.44	15.84	4.09	15.48
3						
>3 omits	89	12.47	4.73	13.12	4.48	12.80
combined	246	13.98		14.63		

Table 9

Repeated Measures ANOVA Summary Table of CNR and CFG by
 Category of Omissiveness for Forms I and II Combined
 N = 246

Source	df	SS	MS	F	F prob
<u>Between</u>					
Category	2	662.77	331.39	9.10	.0002
Error	243	8848.60	36.41		
<u>Within</u>					
Score	1	52.71	52.71	5.11	.0246
Score by Category	2	.34	.17	< 1	
Error	243	2505.48	10.31		

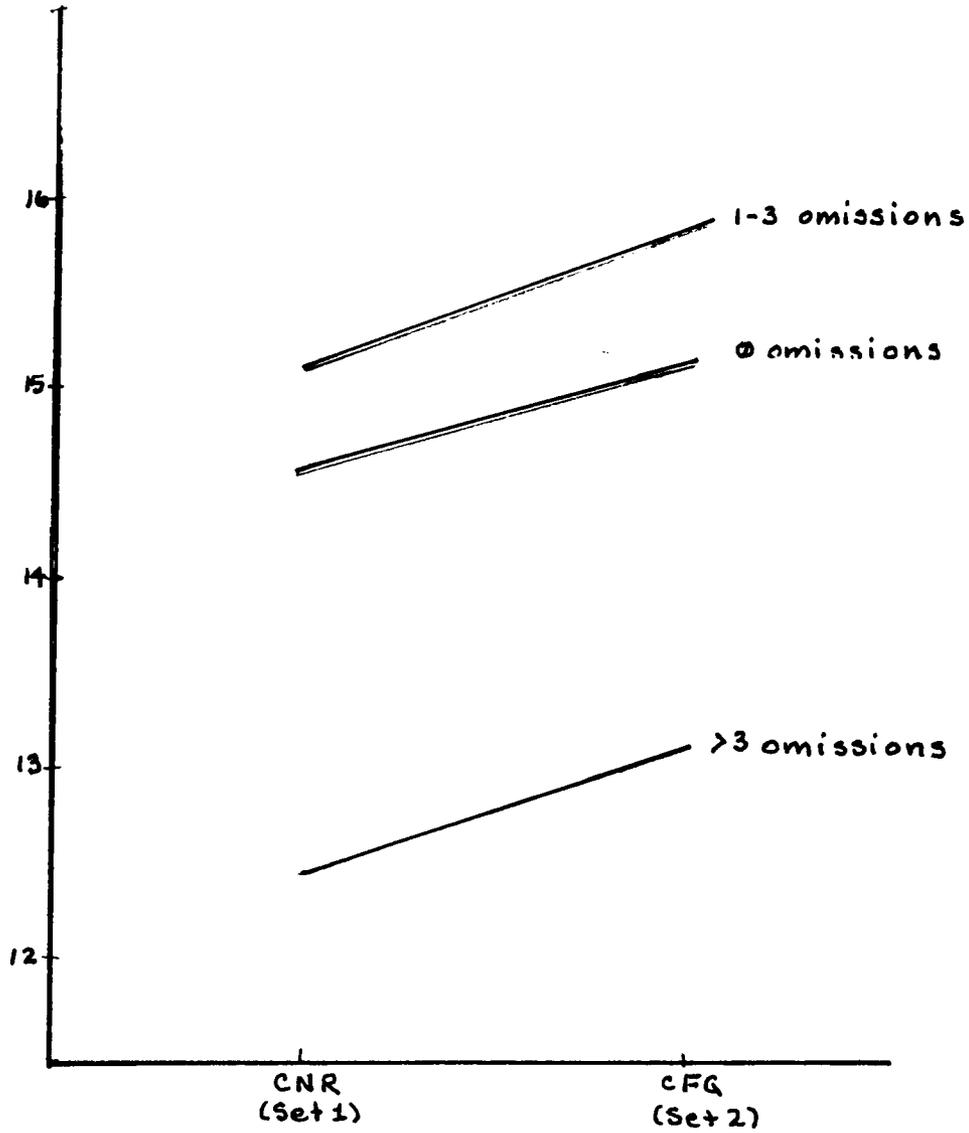


Figure 1. Plot of Mean CNR and CFG Scores by Category of Omissiveness for Forms I and II

the mean CNR scores. Although the category of high omitters had significantly lower scores ($\alpha = .05$) under both sets of directions than both the non-omitters and the moderate omitters, the difference between CNR and CFG for this group was parallel to that of the other two groups. There was clearly no interaction (p -value = .98) between the score differences and the categories of omissiveness. These data clearly do not support the research hypothesis which anticipates that CFG scores will be lower, not higher, than CNR scores, especially among those in the highest omission category.

The results of the analysis of Form III and Form IV are shown in Table 10, Table 11, and Figure 2. In this case the overall mean CFG scores were lower than the mean CNR scores. Overall there was not a significant difference between the mean scores of the three categories of omissiveness. Again there was not a significant interaction between category of omissiveness and difference between CNR and CFG scores; however, the p -value in this case was .06. The value (CNR - CFG) for the non-omitters was .18, whereas the value (CNR - CFG) for the high omitters was 1.83. This would tend to support the research hypothesis. Clearly, the contradictory results associated with the two question sets underscores the need to examine the difference in the two question sets.

Table 10

Descriptive Statistics of CNR and CFG Scores by Category
of Omissiveness for Forms III and IV Combined N = 244

Category	N	CNR question set 1		CFG question set 2		total
		M	SD	M	SD	M
1						
0 omits	81	14.60	5.11	14.42	5.57	14.51
2						
1-3 omits	65	14.52	5.01	13.75	5.18	14.14
3						
>3 omits	98	14.01	4.56	12.18	5.11	13.10
combined	244	14.34		13.34		

Table 11

Repeated Measures ANOVA Summary Table of CNR and CFG by
 Category of Omissiveness for Forms III and IV Combined
 N = 244

Source	df	SS	MS	F	F prob
<u>Between</u>					
Category	2	193.02	96.51	2.38	.0947
Error	241	9772.14	40.55		
<u>Within</u>					
Score	1	101.93	101.93	9.20	.0027
Score by Category	2	62.09	31.05	2.80	.0626
Error	241	2668.91	11.07		

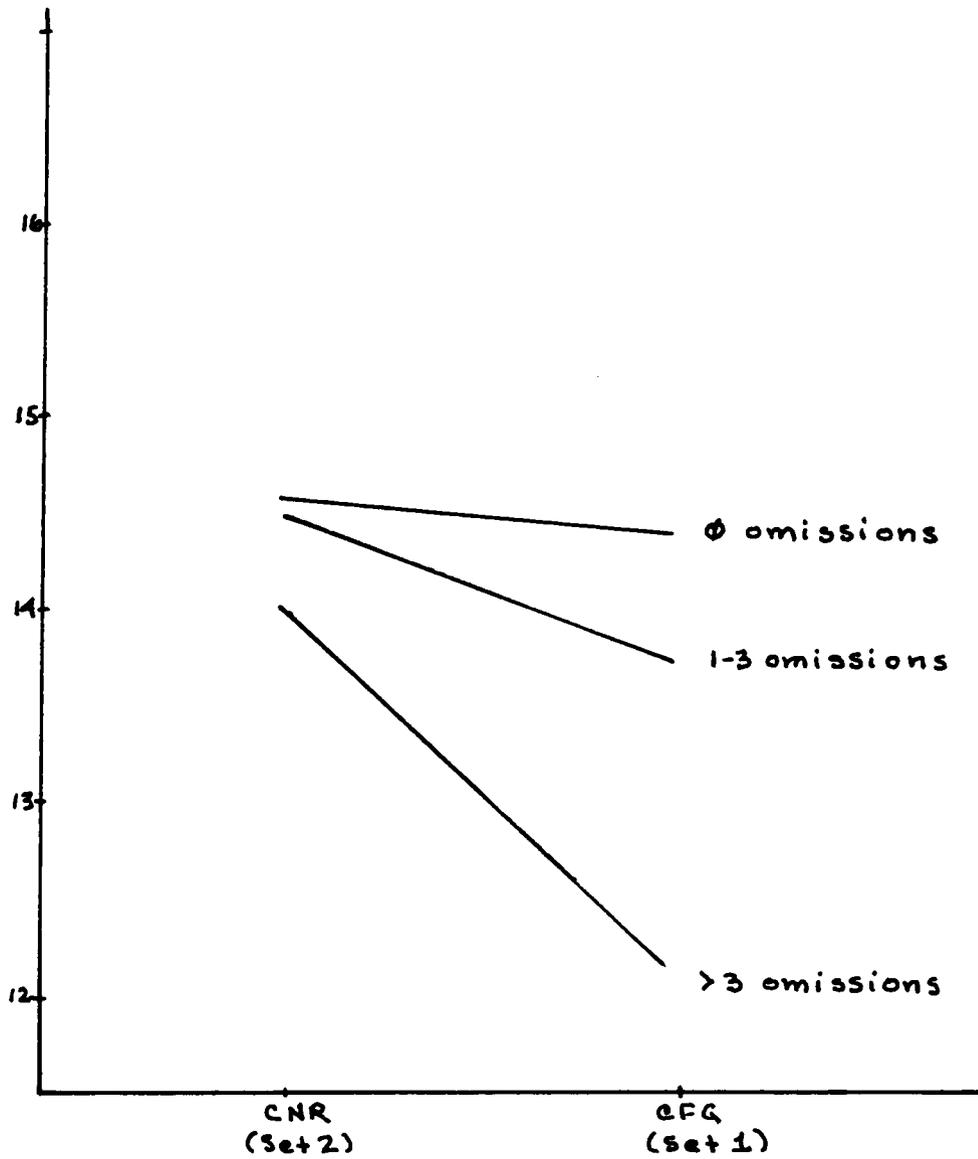


Figure 2. Plot of Mean CNR and CFG Scores by Category of Omissiveness for Forms III and IV

Part II - Risk-Taking Behavior on the Obscure Items

The frequencies of responses to the Type I and Type II questions are given in Table 12 and Table 13.

It was concluded that the Type I obscure questions did not function as hypothesized ($M = -.185$; $SD = .93$; $t = -12.83$; and $p\text{-value} < .001$). As explained in Chapter 3, the expected CFG score across these four items was 4/9. Rather than rule out the blatantly incorrect choices and guess randomly among the remaining three choices, it appears that the obviously wrong choices were also popular choices for two of these items.

However, it was concluded that the Type II obscure questions did meet the criteria set forth in Chapter 3 ($M = -.056$; $SD = .88$; $t = -1.195$; and $p\text{-value} = .23$). The 95% Confidence Interval for percent responding is (.601, .729). It is clear that examinees do not omit, generally, even the most difficult items. In fact the response rate exceeded 50% for each item.

The following points are noteworthy:

1. The most remarkable feature of these results is that in every case the most frequently chosen response on one form of the examination was also the most frequently chosen response on the second form. Given that the choices were not in the same order, that the items were not in the same position on the two forms, and that the

Table 12

Frequencies of Responses to Type I Obscure Items by Form
for 360 Examinees Who Took Fall Final Examination

Form	Item	Omits	Keyed Choice-Form A			
			1	2	3	4
A	3	40	34 ^b	33	17 ^a	56
B	29	57	19(4)	39(2)	14(1)	51(3)
A	10	39	42	43	53 ^a	3 ^b
B	4	43	54(1)	22(3)	55(4)	6(2)
A	16	39	33 ^a	5 ^b	35	68
B	6	45	21(3)	4(1)	42(2)	68(4)
A	59	36	69	30	14 ^a	31 ^b
B	67	46	60(1)	34(3)	16(4)	24(2)

Notes:

Numbers in parentheses are keyed choice on Form B.

^aKeyed correct choice

^bKeyed "obvious" distractor

Table 13

Frequencies of Responses to Type II Obscure Items by Form
for 360 Examinees Who Took Fall Final Examination

Form	Item	Omits	Keyed Choice-Form A			
			1	2	3	4
A	2	66	45 ^a	18	31	20
B	35	82	41(2)	10(1)	24(3)	23(4)
A	6	54	20	32	25 ^a	49
B	23	57	14(1)	35(2)	27(4)	47(3)
A	7	37	50 ^a	25	53	15
B	50	50	25(2)	30(1)	65(3)	10(4)
A	49	68	12	38	55	7 ^a
B	15	68	7(1)	37(2)	55(4)	9(3)

Notes:

Numbers in parentheses are keyed choice on Form B.

^aKeyed correct choice

correct response was chosen no more often than would have been predicted by chance, this suggests that hunches based on something common but unrelated to fact play a definite role in choice selection on multiple choice tests.

2. Even more remarkable is the fact that a substantial number of those examinees who indicated that they never guess or guess only when they can eliminate an incorrect choice or have a hunch chose to respond to these items, albeit incorrectly (see Tables 14 and 15). It is clear that an examinee's definition of hunch may have little to do with relevant information.

3. Further, it should be noted that at least one of the "obviously" incorrect choices of the Type I items were chosen as being correct by nearly one third of the examinees. It is somewhat comforting that these were generally low scorers.

The most important conclusion from this part of the study is that examinees were generally quite willing to guess in the absence of knowledge. Again, as stated earlier, this is tolerable, although not ideal, as long as there is not a bias introduced by the formula directions when one is reluctant to guess.

Table 14

Cross Tabulation of Risk I with Responses to Strategy
Question on Fall Final Examination N = 344

		Strategy			total
		1	2	3	
R I S K I	0	5	12	2	19
	1	5	24	4	33
	2	5	40	3	48
	3	3	46	5	54
	4	3	82	105	190
Total		21	204	119	344

Note:

Strategy Question

Quite apart from what the directions suggested, what did you usually do when you were not certain of the answer to a question?

1. I refrained from guessing.
2. I guessed only if I had a hunch as to the correct answer or could eliminate at least one choice as incorrect.
3. I provided an answer even if my answer was a sheer guess among all choices.

Table 15

Cross Tabulation of Risk II with Responses to Strategy
Question on Fall Final Examination N = 344

		Strategy			
		1	2	3	total
	0	6	26	3	35
R	1	5	35	7	47
I	2	4	55	5	64
S	3	4	40	7	51
K	4	2	48	97	147
II					
Total		21	204	119	344

Note:

Strategy Question

Quite apart from what the directions suggested, what did you usually do when you were not certain of the answer to a question?

1. I refrained from guessing.
2. I guessed only if I had a hunch as to the correct answer or could eliminate at least one choice as incorrect.
3. I provided an answer even if my answer was a sheer guess among all the choices.

Part III - Response Strategies

For the information questions appended to the final examination, omissions ranged from 10% to 20% of the 495 examinees. Therefore, a Chi-square test of independence was done to test whether the category of omissiveness and responding to an information item were independent factors. In every case the p-value was greater than .15

It was concluded that gender and omissiveness are independent factors (p-value > .20, see Table 16). It was also concluded that having taken a class in test taking skills and omissiveness are independent factors (p-value > .20, see Table 17).

Approximately half of the examinees reported that the most frequently given advice relative to guessing under formula scoring instructions was to guess if they could eliminate one distractor as being incorrect. Only 1.5% of the examinees reported that the advice that they had received was never to guess. Twenty-two percent reported that they had been advised never to omit responses. The rest indicated that they had not received advice relative to guessing.

As expected, responses to questions about the examinees' strategies relative to guessing under formula scoring instructions and omissiveness categories are dependent factors (p-values < .001, see Tables 18, 19, and

20). Again, however, as noted earlier in the discussion of consistency of omissiveness, there were a substantial number who indicated that they never omit who omitted items, and several who indicated that they never guess who were non-omitters.

It was concluded that the responses to the question eliciting the examinees' opinions of formula scoring and their omissiveness were dependent factors (p -value = .011, see Table 21). The following breakdown is of particular interest:

1. Of the non-omitters, 55 chose the "don't care" response, while the expected value was 43.76 (Chi-square contribution = 2.82);

2. Of the highest omitters, 27 chose the "approve and it doesn't bother me" response, while the expected value was 16.86 (Chi-square contribution = 6.19);

3. Of the moderate omitters, 16 chose the "approve, but it bothers me" response, while the expected value was 10.87 (Chi-square contribution = 2.42).

4. Finally it is noteworthy that 46% of the respondents indicated that they do not approve of formula scoring, and an additional 9% approve but are bothered by formula scoring.

Table 16

Cross Tabulation of Gender by Omissiveness Category
 N = 403

	Gender		
	male	female	total
1	101	35	136
2	92	23	115
3	115	37	152
total	308	95	403

Chi-square (2) = 1.218

Chi-square prob. = .544

Table 17

Cross Tabulation of Advice from Class in Test Taking
Skills with Category of Omissiveness N = 395

	Class		total
	yes	no	
1	15	117	132
2	12	101	113
3	13	137	150
total	40	355	395

Chi-square (2) = 0.603

Chi-square prob. = .740

Table 18

Cross Tabulation of Responses to Item 61(Strategy) with
Category of Omissiveness N = 443

	Strategy			total
	1	2	3	
1	7	86	59	152
2	6	102	17	125
3	9	133	24	166
total	22	321	100	443

Chi-square (4) = 35.26

Chi-square prob. < .001

Note:

Item 61

Quite apart from what the directions suggested, how did you usually respond when you were uncertain of the answer to a question on the part of the test that is scored with a correction for guessing?

1. I refrained from guessing.
2. I guessed only if I had a hunch as to the correct answer or could eliminate at least one choice as incorrect.
3. I provided an answer even if my answer was a sheer guess among all choices.

Table 19

Cross Tabulation of Responses to Item 62 (Guess Less)
with Category of Omissiveness N = 441

	Guess Less		
	yes	no	total
1	58	93	151
2	86	38	124
3	125	41	166
total	269	172	441

Chi-square (2) = 50.297

Chi-square prob. < .001

Note:

Item 62

On the formula scored part of the test, did you adopt a more conservative strategy (guess less) than for the part to be scored number right (no penalty for incorrect answer)?

1. Yes
2. No

Table 20

Cross Tabulation of Responses to Item 65 (Guess If You Run Out of Time) with Category of Omissiveness N = 435

	Guess If You Run Out of Time			total
	1	2	3	
1	29	88	32	149
2	14	92	15	121
3	17	131	17	165
total	60	311	64	435

Chi-square (4) = 17.629

Chi-square prob. = .001

Note:

Item 65

If you did not have time to read and consider some of the questions on a multiple-choice test, would you mark choices for these questions?

1. not at all
2. only if number-right scoring was to be used
3. even if formula scoring was to be used

Table 21

Cross Tabulation of Responses to Item 69 (Opinion of CFG)
with Category of Omissiveness N = 396

	Response					total
	1	2	3	4	5	
1	55	11	8	29	31	134
2	31	7	16	28	31	113
3	43	27	14	31	34	149
total	129	45	38	88	96	396

Chi-square (8) = 19.765

Chi-square prob. = .011

Note:

Item 69

Which of the following best describes your reaction to formula scoring of multiple-choice tests? (choose only one)

1. I don't care how the test is scored.
2. I think it is a good idea because it keeps people from guessing, and it does not bother me.
3. I think it is a good idea because it keeps people from guessing, but it bothers me.
4. I think it is a bad idea because you should get credit for every answer that you get right.
5. I think it is a bad idea because I don't do as well when I have to be concerned about losing points for a wrong answer.

Recap of the results of the tests of Lord's premise

The mean scores of the question sets were not significantly different when administered under number-right scoring instructions but were significantly different when administered under formula scoring instructions. Under formula scoring instructions the mean score for question set 1 was significantly less than the mean score for question set 2. This outcome was completely unexpected.

Lord's premise was tested twice: once for forms on which question set 1 had been administered under number-right instructions (Forms I and II) and once for forms on which question set 1 had been administered under formula scoring instructions (Forms III and IV). The results reflected the difference in the question sets reported above. For the first analysis, the p-value for the interaction was .98; for the second it was .06, with the high-omitters showing the greatest CNR - CFG difference.

Although two other questions had been posed, and the results of those analyses are reported, the difference in the question set means and the failure to support the research hypothesis became the driving force to do several post hoc analyses. These now are reported in chronological sequence, with only two of the analyses being reported in detail.

Post hoc analyses

Despite the difference in question set mean scores, the first post hoc analysis was to replicate the Frary, Cross and Sewell (1985) study in which the measure of omissiveness had been obtained on the portion of the final examination that had been administered under formula scoring instructions. The results were that the interaction was significant in the hypothesized direction at alpha levels less than .04 and .05 for Forms I and II and for Forms III and IV, respectively. The analysis of the scores of all forms combined into a single sample of paired CNR-CFG scores yielded a p-value of .0032 for the interaction. This may well reflect the credibility of Angoff's challenge to the validity of that design. On the other hand, this may emphasize that a measure of omissiveness which disregards the stimulus of the test being taken may be inappropriate. This result led to an analysis of omissiveness and to several analyses based on different measures of omissiveness.

Omissiveness

The measure of omissiveness for the primary analysis was the total number of omissions on the winter midterm examination. For Forms I and II, table 22 shows that 56% of those examinees who were in the "0 omissions" category

on the mid-term examination, omitted responses on the final examination. For Forms III and IV, table 23 shows that 44% of those examinees, who were in the "0 omissions" category on the mid-term examination, omitted responses on the final examination. Even though one can conclude that omitting items on one examination and omitting items on another are dependent factors (Chi-square probabilities were less than .001 for both analyses), more noteworthy is the fact that almost half (47%) of the examinees switched categories from one examination to the other.

It was thought that a better measure of a tendency to omit might be a measure of consistency of omissiveness. For the 200 examinees who enrolled in both quarters of the sequence, a measure of consistency of omissive behavior was computed. A two-category omissiveness index was used:

1. no omissions on any of the four examinations administered;
2. at least one omission on each of the four examinations.

With only these two categories of omissiveness, the repeated measures ANOVA was done (tables 25 and 27), again for Forms I and II separately from Forms III and IV. As can be seen in Tables 24 and 26, there were very

Table 22

Cross Tabulation of Omissions on Final Examination with
 Category of Omissiveness, Obtained from the Midterm
 Examination for Forms I and II N = 246

	Category on Final Examination			
	1 (0 omits)	2 (1-2 omits)	3 (>2 omits)	total
1	36	28	17	81
2	24	33	19	76
3	10	26	53	89
total	70	87	89	246

Table 23

Cross Tabulation of Omissions on Final Examination with
Category of Omissiveness, Obtained from the Midterm
Examination for Forms III and IV N = 244

	Category on Final Examination			
	1	2	3	total
1	45	21	15	81
2	14	39	12	65
3	13	31	54	89
total	72	91	81	244

few examinees in either category (consistent non-omitters, $n = 34$, and consistent omitters, $n = 34$). As with the previously described analyses, there was not a significant interaction between score differences (CNR-CFG) and the categories of omissiveness. Figures 3 and 4 illustrate once again the difference in the results for the two question sets. The consistent omitters actually did slightly better under CFG instructions ($M = 13.36$) than under NR instructions ($M = 12.77$). This was not the case for either category of omitters based on the midterm examination. As with the previous analyses, the mean score for consistent omitters was significantly less than for the non-omitters.

Several other measures of omissiveness were investigated, in addition to those previously reported. It was thought that perhaps an examinee's statements about his omissiveness or his attitude towards formula scoring might be more appropriate indicators of the potential for a differential effect on score differences. Each of the strategy questions appended to the winter final examination was used as a means of classifying examinees, and the repeated measures analysis of variance was done. None provided support for the research hypothesis. Consistently the number of examinees misclassified according to the final examination classification was quite large.

Table 24

Descriptive Statistics of CNR and CFG Scores for Consistent Omitters and Non-omitters for Forms I and II Combined N = 30

Category	N	CNR		CFG	
		M	SD	M	SD
Non-omitters	13	16.77	5.00	16.39	4.48
Omitters	17	12.53	3.99	14.53	4.39

Table 25

Repeated Measures ANOVA Summary Table of CNR and CFG by
Category of Omissiveness for Forms I and II Combined

N = 30

Source	df	SS	MS	F	F prob
<u>Between</u>					
Category	1	136.83	136.83	4.76	.038
Error	28	805.32	28.76		
<u>Within</u>					
Score	1	9.61	9.61	< 1	
Score by Category	1	20.94	20.94	1.96	.172
Error	28	298.54	10.66		

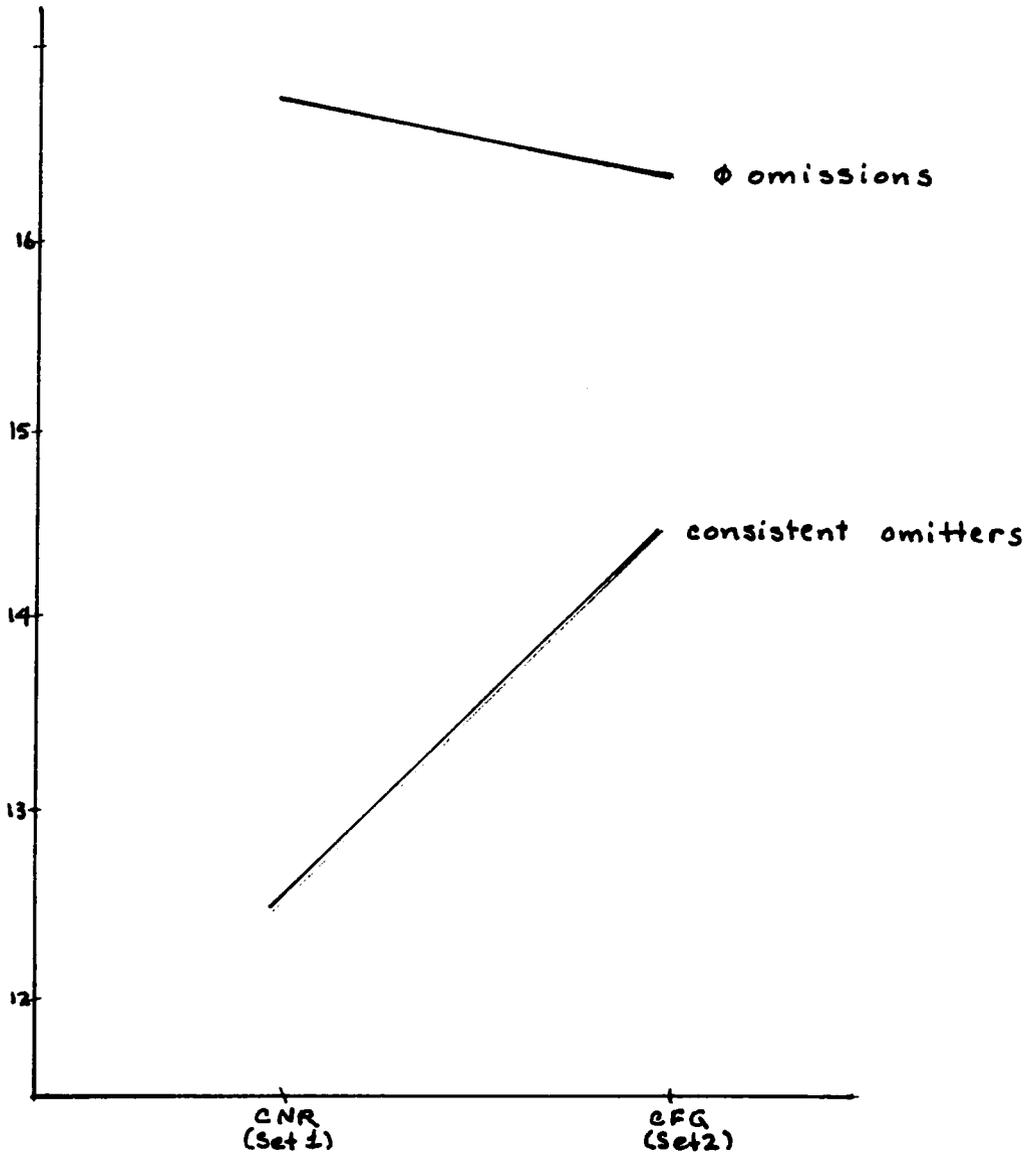


Figure 3. Plot of Mean CNR and CFG Scores by Category of Omissiveness Based on Consistency for Forms I and II

Table 26

Descriptive Statistics of CNR and CFG Scores for Consistent Omitters and Non-omitters for Forms III and IV Combined N = 38

Category	N	CNR		CFG	
		M	SD	M	SD
Non-omitters	21	14.52	5.66	14.76	5.04
Omitters	17	13.00	4.85	12.18	5.20

Table 27

Repeated Measures ANOVA Summary Table of CNR and CFG by
 Category of Omissiveness for Forms III and IV Combined
 N = 38

Source	df	SS	MS	F	F prob
<u>Between</u>					
Category	1	76.32	76.32	1.86	.181
Error	36	1533.38	42.59		
<u>Within</u>					
Score	1	1.61	1.61	< 1	
Score by Category	1	5.29	5.29	< 1	
Error	36	400.14	11.12		

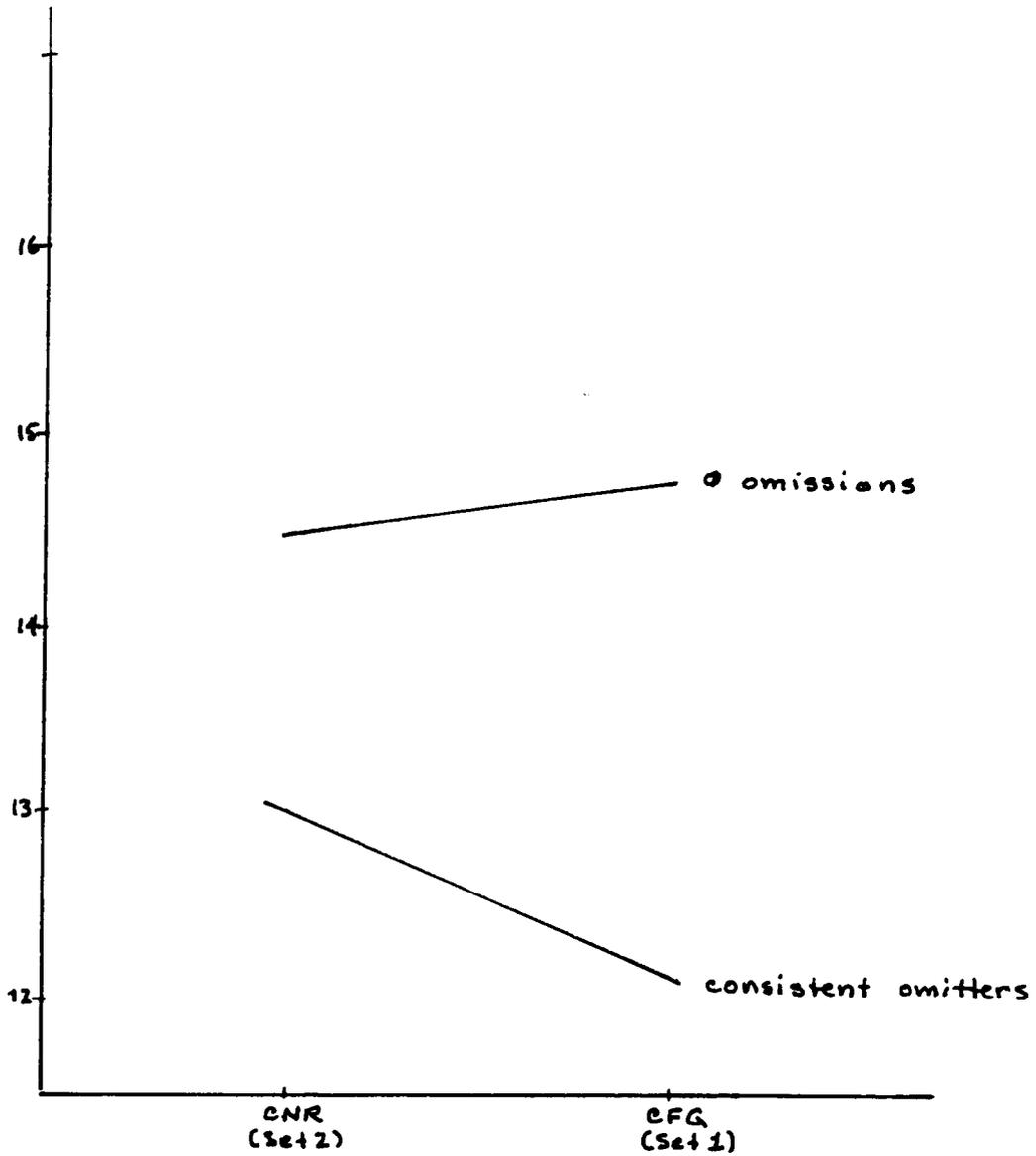


Figure 4. Plot of Mean CNR and CFG Scores by Category of Omissiveness Based on Consistency for Forms III and IV

Analyses based on subsets of the final examination items

Analyses were undertaken in which the repeated measures analysis of variance was done using scores obtained from subsets of the final examination items. For one of these analyses those items that could have been classified as faulty, e.g. items for which examinees chose a distractor more frequently than the correct response, were removed. This particular condition suggested misinformation, which would counteract losses due to inappropriate omissions, as described earlier. The research hypothesis was still not supported. The final examination was then limited to include only difficult items, the thought being that these items elicit the greatest number of omissions and thus may be more sensitive. Again the hypothesis was not supported. In every analysis, except the replication of the Frary, Cross, and Sewell (1985) design, the pattern seen in the primary analysis was repeated. The fact that mean scores between the questions sets were significantly different under formula scoring instructions was the dominant factor of the analysis.

It seemed imperative to investigate the difference between the two question sets, therefore the following portion of the study ensued.

Investigation of Lord's premise by items

For this analysis the scores of the entire group (n = 495) were used. Each item had been administered to half of the examinees under number-right instructions and to half of the examinees under formula scoring instructions. The following value, which for convenience will be called L, was computed for each item.

$$L = \frac{NR_1 + (1/4)S_1}{n_i} - \frac{NR_2 + (1/4)S_2}{n_i}$$

where;

NR₁ = the number of correct responses to an item when the item is administered under number right instruction;

S₁ = the number of omissions of an item when administered under number right instructions;

n_i = the number of subjects per group

NR₂ = the number of correct responses to an item when the item is administered under formula scoring instructions;

S₂ = the number of omissions of an item when administered under formula scoring instructions.

The value L is merely the difference in percent correct under the two sets of instructions, if one were to add 1/4 of the number of omissions to the number of correct responses. If Lord's premise, that an item omitted under formula scoring instructions would only be replaced by a random guess under number-right scoring instructions, is correct then the value of L would be expected to equal zero. A positive value of L would indicate that examinees had performed better under number-right scoring instructions, vis a' vis formula scoring instructions. A negative value would indicate the reverse.

The mean value of L for the first question set is .015, whereas the mean value of L for the second question set is -.007. Overall the variable L has a slightly positively skewed, unimodal distribution, with an overall mean value of .004 and SD = .045 (see Figure 5). This would tend to support Lord's premise.

The value of L was also computed for each item in order to investigate the relation of item difficulty and percent of responses omitted to Lord's premise (see Table 28).

A breakdown of the items by difficulty and omissions, yielded the following results:

1. Moderately easy to easy items (percent correct under number-right scoring instructions greater than 70%) had a mean value of L equal to .003 and SD = .039. Eight of these items were in question set 1 and ten were in question set 2. Even for those items for which there were many omissions, the values of L were near zero.

2. Difficult items (percent correct less than 50%) also had small values of L. There were 17 items that were in this category, seven in set 1 and ten in set 2. The mean value of L for these items was -.010 and SD = .034. Seven of these items elicited a high rate of omissions (percent of omissions greater than 10%), but apparently appropriately.

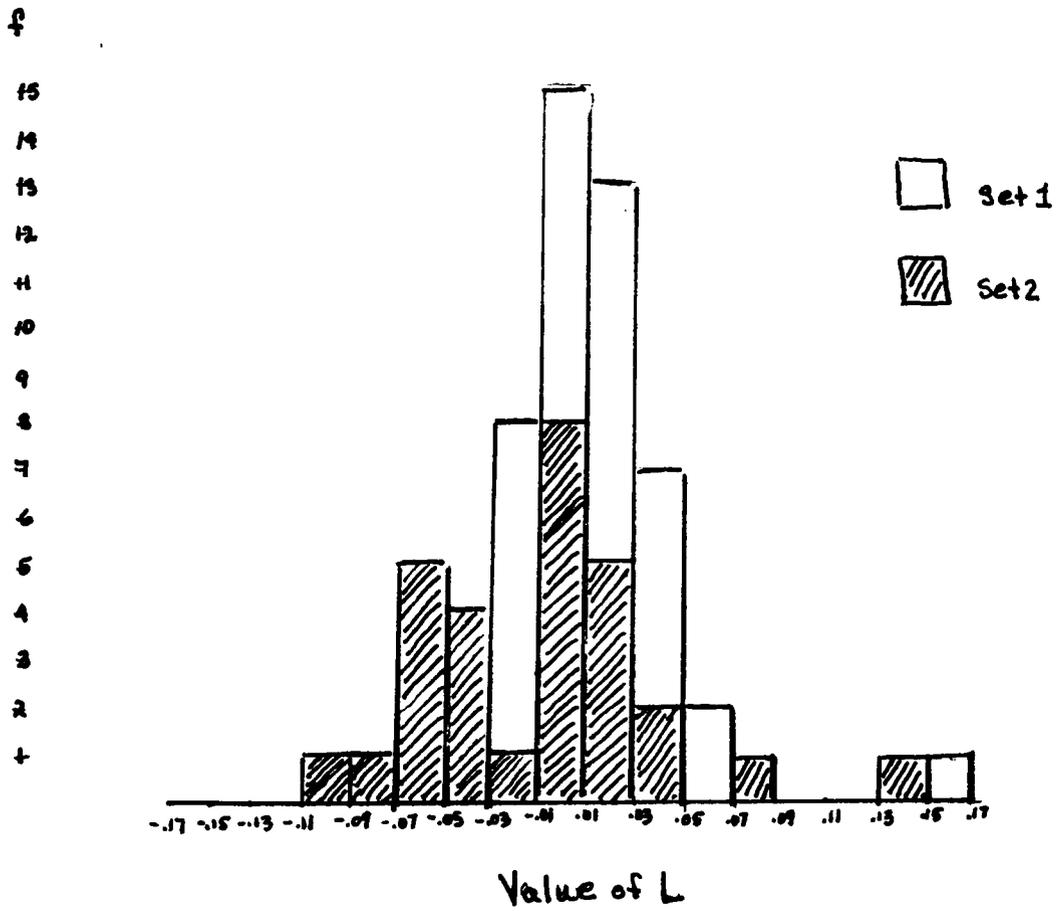


Figure 5. Histogram of the Variable L

Table 28

Item	% Correct	% Omit	L
1	.83	.04	.018
2	.59	.17	.030
3	.90	.02	-.023
4	.55	.07	-.006
5	.65	.02	-.003
6	.65	.17	.153
7	.70	.09	-.005
8	.76	.03	.005
9	.76	.03	.047
10	.80	.04	-.027
11	.58	.12	.021
12	.55	.00	.022
13	.94	.02	.011
14	.39	.47	.021
15	.52	.02	-.021
16	.59	.04	.042
17	.52	.06	.027
18	.93	.01	-.009
19	.37	.16	-.022
20	.39	.06	.001
21	.61	.08	.046
22	.86	.03	.061
23	.38	.27	-.029
24	.66	.09	.000
25	.39	.07	.021
26	.50	.23	.052
27	.50	.25	-.014
28	.12	.06	-.026
29	.40	.04	.028
30	.56	.02	.037

Table 28 (continued)

Item Information

Item	% Correct	% Omit	L
31	.64	.05	-.051
32	.69	.02	-.032
33	.48	.12	.012
34	.67	.19	.139
35	.18	.04	.046
36	.65	.10	-.058
37	.91	.01	-.004
38	.73	.02	-.007
39	.40	.08	-.054
40	.79	.04	-.042
41	.73	.07	-.088
42	.37	.17	-.055
43	.61	.03	-.092
44	.55	.02	.027
45	.92	.02	.022
46	.71	.07	-.016
47	.47	.04	-.068
48	.28	.25	-.050
49	.67	.06	.021
50	.76	.01	-.007
51	.47	.06	.045
52	.30	.27	.005
53	.82	.18	.008
54	.74	.20	.027
55	.49	.10	.006
56	.69	.03	.011
57	.64	.08	.003
58	.65	.17	-.004
59	.28	.30	-.049
60	.86	.07	.082

3. Moderately difficult items (item difficulty between 50 and 70%), having a rate of omissions greater than 10%, yielded the largest values of L. There were seven items that were in this category. More important, five of these items were in question set 1 and only two were in question set 2. The mean value of L for these items was .054 and SD = .067. These items were dramatically different from items of the same difficulty level having fewer omissions and from items of greater and lesser difficulty.

All but one of these seven items were taken from the textbook and dealt with topics that were generally not discussed in class lectures. It is easy to see how these could have been omitted by a number of examinees under the formula scoring directions; but it is important to note that under number-right instructions, when the examinees nearly always guessed when uncertain, that they guessed at better than chance probability. It is also important to note that items of moderate difficulty are the desirable items for discrimination (See Lord, 1968, or Ebel, 1979).

4. It appears that an important difference between the question sets was in the difference in the distribution of the difficulty levels of the items. For question set 1 there were seven difficult items, 15 moderately

difficult items, and eight easy items. For question set 2, there were ten items in each category. These distributions yielded mean item difficulties that were not significantly different for the two question sets. Likewise, the mean number of omissions were not significantly different for the two sets. The preceding analysis, however, suggests that items that are moderately difficult which elicit large numbers of omissions, may present the greatest challenge to Lord's premise.

CONCLUSIONS

The concern that generated this study was the concern that correction-for-guessing might be unfair, not in the scoring itself, but in the eliciting of inappropriate omissions by a subset of examinees. Frederic Lord (1975) advanced the premise that responses omitted under formula scoring instructions would only be replaced by random guesses under number-right scoring instructions. The research hypothesis of this study was that the difference between corrected-number right (CNR) scores and corrected-for-guessing (CFG) scores would be greater for examinees who demonstrate a tendency to omit items under correction-for-guessing scoring instructions than for examinees who demonstrate a tendency not to omit responses under correction-for-guessing scoring instructions. The research hypothesis was not supported, but it would be premature to conclude that there is not a differential effect on test score associated with omissiveness.

Two major problems were identified which cast doubt on the results of this study but which may provide suggestions for future research. These problems concerned the test itself and the measure of omissiveness.

It is clear that the difference between question set 1 and question set 2 affected the outcome of this study.

In particular, had both question sets been like set 1, the research hypothesis probably would have been supported. On the other hand, had both sets been like set 2, it would not. The important point is that these sets were carefully prepared and actually met the criteria of having mean item difficulties that were not significantly different, and of having mean numbers of omissions that were not significantly different. Likewise, the question sets contained equal numbers of items for each topic and from each source. The reliability of the final examination, although not high, exceeded the reliability of tests used in previous studies.

An examination of test items revealed that moderately difficult items that elicit a large number of omissions yield a large difference between the number-right percents and the formula percents. This was not true for easy or difficult items. Further this investigation revealed that five of these moderately difficult items were in question set 1 and only two were in the second question set.

These results show that subtle differences in tests may yield dramatic differences relative to this research hypothesis. Additionally tests need to be such that generalizations can be made to the situations of concern, the aptitude and achievement testing situation. It is

not at all clear which of the question sets would more closely resemble those used by the Educational Testing Service, for example.

The second problem concerned the measure of omissiveness. In order to avoid the criticism of experimental dependence, omissiveness in the current study was defined on the basis of responses to tests administered prior to the test used to compare CNR and CFG mean differences. Unfortunately, the tendency to omit was not very consistent from test to test. Indeed inspection of table 22 on page 61 shows that 47% of the examinees, classified by the midterm examination, were misclassified on the measure of omissiveness relative to the final examination.

Those studies which have supported the differential effects hypothesis, but which have been challenged either on the basis of internal or external validity, have all included some consideration of omissiveness on the test under study, i.e. the tendency to omit items relative to the perception of the difficulty of the items. The attempt to meet the challenges to validity in effect removed the important consideration of the effect of the test under study on omissiveness.

There did not appear to be a problem with too few omissions. Even though examinees knew that the grades would be based on ranking, the evidence is that there

were as many or more omissions for the number of items on the test as reported in the studies by Cross and Frary (1977); Angoff and Schrader (1984); and Frary, Cross, and Sewell (1985). There were a substantial number of omissions on both the midterm examination and the final examination.

In the attempt to make sure that the examinees understood the best strategy for responding under formula scoring instructions, another confounding factor may have been introduced. It has been noted that 47% of the examinees either increased or decreased their number of omissions on the final examination relative to the midterm examination. After the midterm examination, the examinees had the opportunity to see their scores. It could well be that examinees who scored relatively low on the midterm examination decided that they had omitted too many or too few responses.

Clearly, the issue is not resolved, but it needs to be. Both Frary (1988) and Albanese (1987) have demonstrated that scores of some students may be substantially, adversely affected by inappropriate omissiveness. Frary (1988) has demonstrated that identifying the adverse effect may be particularly difficult because of the offsetting positive effect of omissiveness in the face of wrong partial information. This study has iden-

tified a particular category of items which seem to elicit inappropriate omissions. These items were balanced by other items with high omission rates, in that the mean value of the L statistic overall for high omission items was negative. This study has also shown that a substantial number of examinees do not omit responses. Further it was shown that examinees' intended strategies and beliefs relative to formula scoring are clearly related to their omissiveness. Finally more than half of these examinees indicated that they were either bothered by formula scoring or felt that it was a "bad idea."

From a practical point of view, it appears that the resolution of the problem will have to come from the producers of the standardized tests that are used to measure aptitude and achievement. Even though the Cross and Frary (1977) study was challenged because the testing situation that they used might not be generalizable to a regular testing situation, their design nevertheless provides the most direct test of Lord's premise. Their design also prevents the problem with the measure of omissiveness, also. Given the results of this study, it seems that an even greater challenge to validity would be in generalizing conclusions based on an in-class test to a standardized test. Considering the results of the Frary (1988) study, an improvement to the Cross and Frary

design would be to include examinee ability as a between groups factor.

Secondly, Angoff and Schrader (1984) have data from the Scholastic Aptitude Test in which items were administered under the two conditions, formula scoring instructions and number-right scoring instructions. It would be informative to know if the same values of L relative to high omissiveness for moderately difficult items found in this study were also obtained on those items. If so, then an examination of the omission rates on the moderately difficult items of standardized tests would be valuable information.

Given the importance of the tests for which formula scoring is used, the topic should be pursued. Further, in pursuing this topic valuable insights may be gained about test-taking behavior in general.

REFERENCES

Albanese, M.A. (1987) A study of the practical effects of the correction for guessing. An unpublished paper.

Albanese, M.A. (1986) The correction for guessing: a further analysis of Angoff and Schrader. Journal of Educational Measurement, 23, 225-235.

Angoff, W.H., and Schrader, W.B. (1986) A rejoinder to Albanese, "The correction for guessing: a further analysis of Angoff and Schrader. Journal of Educational Measurement, 23, 237-243.

Angoff, W.H., and Schrader, W.B. (1984) A study of hypotheses basic to the use of rights and formula scores. Journal of Educational Measurement, 21, 1-17.

Bliss, L.B.A. (1980) A test of Lord's assumption regarding examinee guessing behavior on multiple-choice tests using elementary school students. Journal of Educational Measurement, 17, 147-153.

Cross, L.H., and Frary, R.B. (1977) An empirical test of Lord's theoretical results regarding formula scoring of multiple-choice test. Journal of Educational Measurement, 14, 313-321.

Davis, F.B. Educational measurements and their interpretation. Belmont, California: Wadsworth Publishing Company, Inc. 1964.

Diamond, J. and Evans, W. (1973). The correction for guessing. Review of Educational Research, 43, 181-191.

Ebel, Robert E. Essentials of educational measurement. Englewood Cliffs, New Jersey: Prentice-Hall, Inc. 1979.

Frary, R.B. (1987) Why formula scoring works as well as it does: A simulation study. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, Louisiana, 1988.

Frary, R.B., Cross, L.H., and Sewell, E.H. (1985) Partial information and the "correction" for guessing. Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago, Illinois, 1985.

Lindquist, E.F., ed. Educational measurement. Washington, D.C.: American Council on Education, 1966. pp.270-370.

Long, E.B. The Civil War day by day, an almanack 1861-1865. Garden City, New York: Doubleday and Company, Inc. 1971.

Lord, F.M. (1975) Formula-scoring and number-right scoring. Journal of Educational Measurement, 12, 7-12.

Lord, F.M. and Novick, M. R. Statistical theories of mental test scores. Reading, Massachusetts: Addison-Wesley Publishing Company. 1968.

Randall, J. G. and Donald, D. H. The Civil War and reconstruction. Lexington, Massachusetts: D. C. Heath and Company. 1969.

Slakter, M.J. (1968b). The effect of guessing strategy on objective test scores. Journal of Educational Measurement, 5, 217-221.

Spencer, James. Civil War generals, categorical listings and a biographical directory. Westport, Connecticut: Greenwood Press. 1986.

Wakelyn, Jon L. Biographical dictionary of the Confederacy. Westport, Connecticut: Greenwood Press. 1977.

Appendix A

Instructions for Midterm Examinations

INSTRUCTIONS FOR MID-TERM EXAMINATION

Mark your answers on the answer sheet provided. Use a No. 2 pencil. Each question has only one right answer. Erase thoroughly if you decide to change an answer. Be sure to mark your FORM LETTER on your answer sheet.

Your test scores will be based on the number of questions you answer correctly minus a fraction of the number you answer incorrectly. If you have some knowledge of a question, you may be able to eliminate one or more of the answer choices as wrong. In this case, it is generally to your advantage to guess which of the remaining choices is correct. Furthermore, it is improbable that random or haphazard guessing will change your scores significantly.

Appendix B

Instructions for Fall Final Examination

YOUR TEST SCORES WILL BE BASED ON THE NUMBER OF QUESTIONS YOU ANSWER CORRECTLY MINUS A FRACTION OF THE NUMBER YOU ANSWER INCORRECTLY. IF YOU HAVE SOME KNOWLEDGE OF A QUESTION, YOU MAY BE ABLE TO ELIMINATE ONE OR MORE OF THE ANSWER CHOICES AS WRONG. IN THIS CASE, IT IS GENERALLY TO YOUR ADVANTAGE TO GUESS WHICH OF THE REMAINING CHOICES IS CORRECT. FURTHERMORE, IT IS IMPROBABLE THAT RANDOM OR HAPHAZARD GUESSING WILL CHANGE YOUR SCORE SIGNIFICANTLY.

Appendix C

Fall Final Examination

HISTORY 3051
FALL 1987
FINAL EXAM - FORM A

Write your answers on the answer sheet provided. Use a No. 2 pencil. Each question has only one right answer. Be sure to erase thoroughly, if you change an answer. MARK FORM A ON YOUR ANSWER SHEET!

1. At the same time General Grant was campaigning against Henry-Donelson-Shiloh, General John Pope was winning an important campaign capturing 5000 Confederates at:
 1. Island Number Ten.
 2. Port Hudson.
 3. New Madrid.
 4. Fort Pillow.
2. Which of the following occurred first?
 1. skirmish near Adamsville, Tennessee
 2. USS Adirondack wrecked
 3. Apalachiola, Florida, blockaded
 4. Sioux uprising at Acton, Minnesota.
3. Which of the following blockade runners was the first to be captured?
 1. Monitor
 2. Albion
 3. Alert
 4. Columbia
4. The battle plan which Lee adopted for relieving Richmond in June, 1862, was:
 1. to blast the army into retreat with heavy artillery.
 2. to make frontal assaults against McClellan's center.
 3. to isolate and attack the Union army south of the Chickahominy.
 4. to attack the Union right flank so as to isolate it from supplies from Washington.
5. The figure who gained the greatest fame at Manassas/Bull Run was:
 1. Beauregard.
 2. Stuart.
 3. Johnston.
 4. Jackson.
6. Lincoln's message to Governor Pickens, of South Carolina, that Fort Sumter would be resupplied, was delivered by:
 1. John L. Worden.
 2. Leroy P. Walker.
 3. Robert S. Chew.
 4. James L. Chesnut.
7. Brig. General Joseph Reid Anderson's occupation before the war was:
 1. iron works superintendent.
 2. lawyer.
 3. Baptist minister.
 4. police chief.

- 2 -

8. In population base (and hence manpower resources) the North outnumbered the South by:
 1. 2:1.
 2. 4:1.
 3. 8:1.
 4. 10:1.

9. Richmond became the capital of the Confederacy because:
 1. of its closeness to Washington.
 2. it was the largest city in the South.
 3. Virginia demanded the capital as a condition for joining the Confederacy.
 4. it was the agricultural center of the South.

10. On April 25, 1861, U. S. troops surrendered at:
 1. Carthage, Missouri.
 2. Rockport, Arkansas.
 3. Saluria, Texas.
 4. Boise, Idaho.

11. The fall of Fort Henry came primarily because of:
 1. the Federal gunboats.
 2. incompetence on the part of the Confederate commander.
 3. the Tennessee River.
 4. Sidney Johnston's lack of concern for the area.

12. The official name for the 1861 - 1865 struggle is the:
 1. Civil War.
 2. War between the States.
 3. War of the Rebellion.
 4. War of Secession.

13. By fighting defensively, the South:
 1. fought fewer battles.
 2. needed fewer men.
 3. could choose the time and place for battle.
 4. isolated itself from necessary supply lines.

14. The first capital of the Confederacy was:
 1. Montgomery, Alabama.
 2. Atlanta, Georgia.
 3. Charleston, South Carolina.
 4. New Orleans, Louisiana.

15. The relief expedition to Fort Sumter did NOT include which of the following ships?
 1. Baltic
 2. Pawnee
 3. Powhatan
 4. Harriet Lane

- 3 -

16. At the Battle of Wilson's Creek, the Union withdrawal was led by:
 1. Major Samuel Sturgis.
 2. General "Stonewall" Jackson.
 3. Major Sterling Price.
 4. Brig. General Nathaniel Lyon.
17. The greatest technological revolution in naval warfare during the Civil War was:
 1. the use of mines.
 2. the use of submarines as tactical weapons.
 3. ironclad ships.
 4. the perfection of a blockade.
18. According to the text, the knottiest problem of the Confederacy was:
 1. drafting a constitution.
 2. securing men for the armies.
 3. raising enough food.
 4. securing enough munitions to wage war.
19. The army unit where esprit de corps, pride, and "home" feeling prevailed was the:
 1. company.
 2. regiment.
 3. brigade.
 4. division.
20. The leading cause of death in the Union army was:
 1. camp fever.
 2. pneumonia.
 3. consumption.
 4. diarrhea and dysentery.
21. Which of the following criticisms of Jefferson Davis is inaccurate?
 1. He was a poor administrator.
 2. He was insensitive to public opinion.
 3. He insisted on combining civilian and military leadership of the Confederacy in one person, himself.
 4. He lacked an overall strategic plan for winning the war.
22. Which of the following is NOT true of the South's geographical advantages in the war?
 1. The Appalachians divided the war into major theaters.
 2. Moving southward up the Shenandoah actually took Union troops farther away from Richmond.
 3. Defending Richmond was difficult because of the wide area between the mountains and the coast.
 4. The rivers of Virginia ran in an advantageous direction for defense.

- 4 -

23. Confederate conscription was a disaster primarily because of:
 1. sweeping exemption laws enacted in rapid fashion.
 2. inability of the government to enforce the law after a few months.
 3. dissension in the Congress over the wording of the law.
 4. President Davis' opposition to conscription.
24. The postmaster general under Abraham Lincoln was:
 1. Edward Bates, of Missouri.
 2. Caleb Smith, of Indiana.
 3. Simon Cameron, of Pennsylvania.
 4. Montgomery Blair, of Maryland.
25. By far the most revolutionary innovation in this conflict was the development of:
 1. exploding cannon shells.
 2. the canister.
 3. the rifle.
 4. land mines.
26. The effective range of the smooth barreled musket was:
 1. 80 yards.
 2. 160 yards.
 3. 250 yards.
 4. 400 yards.
27. Southern advantages as the war began did NOT include:
 1. a more experienced president.
 2. a strong aptitude for military life.
 3. a well-organized militia.
 4. the ability to strike quickly with a small but mobile navy.
28. Nineteen Federal army posts were delivered to Texas authorities by:
 1. General David E. Twiggs.
 2. Major Robert J. Anderson.
 3. General Henry W. Halleck.
 4. General S. B. Buckner.
29. The climax of the first battle along Bull Run came when:
 1. Confederates counter-attacked against the Union center.
 2. a Shenandoah Valley brigade held its position atop Henry House Hill.
 3. Federals slammed into Evans' weak South Carolina lines.
 4. Johnson's troops linked with those of Beauregard.
30. The most preferred close range cannon was:
 1. developed by John Parrot.
 2. the Napoleon.
 3. the double barrel cannon.
 4. developed by James Burton.

- 5 -

31. Which of the following is NOT a result of First Manassas?
 1. The South became energized, ready for further battle.
 2. The North began rapidly mobilizing for a long war.
 3. General "Stonewall" Jackson became an overnight Southern hero.
 4. General George McClellan assumed command of Union troops in the field.
32. Josiah Gorgas is most remembered as:
 1. the man in charge of all Confederate munitions.
 2. the man who succeeded Charles Sumner in the U. S. Senate.
 3. McClellan's chief of staff.
 4. Jefferson Davis' principal adviser.
33. The "Anaconda Plan" is associated with the strategy proposed by:
 1. Thomas Jackson.
 2. Winfield Scott.
 3. Ulysses Grant.
 4. George McClellan.
34. With the invention of the rifle, the advantage in battle:
 1. switched to the offense.
 2. switched to the defense.
 3. switched to the cavalry.
 4. switched to the infantry.
35. Which of the following is not an innovation of this war?
 1. use of land mines
 2. war photography
 3. massive staff planning
 4. mass and maneuver tactics
36. McClellan's Peninsular Campaign exemplified the military philosophy of:
 1. mass and maneuver (not to fight at all).
 2. destroying the enemy's army.
 3. developing strong defensive works and waiting for the enemy to attack.
 4. bombarding the enemy's capital into submission.
37. Of the following, which Federal fort was NOT seized before the Civil War?
 1. Fort Pulaski
 2. Fort Barrancos
 3. Fort Jefferson
 4. Fort Morgan
38. The weather condition most characteristic of the 1862 Peninsular Campaign was:
 1. rain.
 2. heat.
 3. cold.
 4. fog.

- 6 -

39. The basic structure of an army formation was (from smallest to largest):
1. company/regiment/brigade/division/corps.
 2. regiment/brigade/company/division/corps.
 3. company/brigade/division/regiment/corps.
 4. brigade/division/company/regiment/corps.
40. General George B. McClellan was a field commander:
1. widely disliked by his men.
 2. lacking in administrative talents.
 3. beset by strong inner doubts.
 4. who enjoyed Lincoln's confidence throughout the first year of the war.
41. The developer of the grooved barrel was:
1. James Burton.
 2. Claude Minie.
 3. John Parrott.
 4. Charles Winfield.
42. The Seven Days Battles ended with the Confederate defeat at:
1. Mechanicsville.
 2. Savage Station.
 3. Malvern Hill.
 4. White Oak Swamp.
43. Treatise on Tactics, the guiding textbook for military tactics at the beginning of the Civil War, was written by:
1. Scott.
 2. Forest.
 3. Jomini.
 4. von Clausewitz.
44. In which of the following resources did the South exceed the North at war's outset?
1. capital
 2. military leadership
 3. manufactured goods
 4. transportational facilities
45. Which of the following did NOT appear for the first time in the American war?
1. telegraph
 2. large staffs and staff planning
 3. concept of total war
 4. submarines
46. Which of these statements is FALSE?
1. Regiments were designated by number and state (e.g., 18th Virginia).
 2. A division was the basic unit sent into battle.
 3. A brigade generally had its own artillery and cavalry support.
 4. A regiment was a local product and consisted of 1,000 men.

- 7 -

47. The Civil War food which tended to "petrify" by the time it reached the front lines was:
1. beef jerky.
 2. cornmeal.
 3. hardtack.
 4. dessicated vegetables.
48. Treatise on Tactics DID NOT teach that:
1. war is an extension of politics.
 2. the good general amasses a large army.
 3. mass and maneuver are the imperatives in waging war.
 4. the object of war is to attack relentlessly.
49. The "Florence Nightingale of the Southern Army" was:
1. Jane Conwell Shroud.
 2. Sarah Smith McCallum.
 3. Elizabeth Cage Hayes.
 4. Ella King Newsom.
50. Which of these statements is NOT true of naval warfare in the 1860s?
1. Blockades were found to be useless in a war of this scope.
 2. Mines were used extensively in rivers and harbors.
 3. Submarines were first used as tactical weapons.
 4. Developments in American naval warfare rendered all other navies obsolete.
51. Which of these was NOT a problem for the Confederate government?
1. tariffs
 2. the seizure of U.S. forts in the South
 3. control of Indian affairs
 4. harmonious selection of men for President and Vice President
52. The last state to get the approval of its citizens to secede was:
1. Arkansas.
 2. Virginia.
 3. Tennessee.
 4. North Carolina.
53. The majority of Prentiss' men in the Hornet's Nest at Shiloh were from:
1. Michigan.
 2. Illinois.
 3. Iowa.
 4. Wisconsin.
54. General McDowell:
1. never won a Civil War battle.
 2. won battles mostly on "luck," not ability.
 3. wore a huge hat in battle so his staff officers could easily locate him.
 4. was generally liked by his soldiers.

- 8 -

55. The Confederate defeat that caused Albert Sidney Johnston to consolidate his forces at Forts Henry and Donelson was the battle of:
1. Shiloh.
 2. Pea Ridge.
 3. Wilson's Creek.
 4. Mill Springs.
56. At the outset of the war:
1. the South believed it would be over quickly.
 2. the North believed it would be long and drawn out.
 3. neither of the above
 4. both 1 & 2
57. In the battle for Fort Donelson:
1. the Confederate commander was General Lloyd Tilghman.
 2. the flooding Cumberland River aided the Union victory.
 3. General Buell reinforced Grant at the last moment to insure victory.
 4. none of these
58. The North named battles after:
1. towns.
 2. mountains.
 3. streams.
 4. military officers.
59. The oldest Civil War general on either side was:
1. Ethan Allen Hitchcock.
 2. John Henry Winder.
 3. John Ellis Wool.
 4. J.E.B. Stuart.
60. "Johnny Reb" would name the first major battle of the war as :
1. The Battle for Northern Virginia.
 2. The Battle of Bull Run.
 3. The Battle for the Orange and Alexandria Railroad.
 4. The Battle of Manassas.
61. When Lincoln assumed the presidency in 1861, he:
1. entered office with a clear mandate from 60% of the voters.
 2. found a healthy U. S. treasury.
 3. organized a coalition government of Democrats and Republicans.
 4. found many government officials leaving to join the confederacy.
62. In which campaign did Confederate troops break through Union lines, only to turn around and fight their way back into their original works?
1. Fort Magruder
 2. Fort Henry
 3. Fort Donelson
 4. Fort Monroe

- 9 -

63. Which one of the following was NOT a result of the Henry-Donelson campaign?
1. It demonstrated that the war in the West was going to a long and painful process.
 2. It opened up all of middle and western Tennessee to Federal invasion.
 3. It brought about the capture of the largest number of Americans in the nation's history.
 4. It led immediately to the downfall of the first capital seized in the war.
64. One of the chief borrowing measures of the Confederacy was called:
1. the "railroad loan."
 2. the "foreign loan."
 3. the "produce loan."
 4. the "cotton loan."
65. Major Robert Anderson, who surrendered Fort Sumter, had immense pride in one West Point cadet who performed outstandingly in the West Point classroom and outstandingly at Fort Sumter. That cadet was:
1. Lee.
 2. Beauregard.
 3. Jackson.
 4. Stuart.
66. The incident that led directly to Virginia's secession was:
1. secession of South Carolina.
 2. firing on Fort Sumter.
 3. Lincoln's call for troops.
 4. Federal refusal to abandon Harper's Ferry.
67. The "Bull Run Races" referred to the rapidity of the:
1. Federal advance.
 2. Federal retreat.
 3. Confederate advance.
 4. Confederate retreat.

Appendix D

Type I Obscure Questions

TYPE I OBSCURE QUESTIONS
FORM A

3. Which of the following blockade runners was the first to be captured?
 1. Monitor
 2. Albion
 3. Alert
 4. Columbia

10. On April 25, 1861, U. S. troops surrendered at:
 1. Carthage, Missouri.
 2. Rockport, Arkansas.
 3. Saluria, Texas.
 4. Boise, Idaho.

16. At the Battle of Wilson's Creek, the Union withdrawal was led by:
 1. Major Samuel Sturgis.
 2. General "Stonewall" Jackson.
 3. Major Sterling Price.
 4. Brig. General Nathaniel Lyon.

59. The oldest Civil War general on either side was:
 1. Ethan Allen Hitchcock.
 2. John Henry Winder.
 3. John Ellis Wool.
 4. J. E. B. Stuart.

Appendix E

Type II Obscure Questions

TYPE II OBSCURE QUESTIONS
FORM A

2. Which of the following occurred first?
 1. skirmish near Adamsville, Tennessee
 2. USS Adirondack wrecked
 3. Apalachiola, Florida, blockaded
 4. Sioux uprising at Acton, Minnesota.

6. Lincoln's message to governor pickens, of South Carolina, that Fort Sumter would be resupplied, was delivered by:
 1. John L. Worden.
 2. Leroy P. Walker.
 3. Robert S. Chew.
 4. James L. Chesnut.

7. Brig. General Joseph Reid Anderson's occupation before the war was:
 1. iron works superintendent.
 2. lawyer.
 3. Baptist minister.
 4. police chief.

49. The "Florence Nightingale of the Southern Army" was:
 1. Jane Conwell Shroud.
 2. Sarah Smith McCallum.
 3. Elizabeth Cage Hayes.
 4. Ella King Newsom.

Appendix F

Addendum to Fall Final Examination

- 10 -

We are interested in your understanding of what the test directions advised you, relative to guessing, and whether you followed this advice or adopted a guessing strategy of your own. Please record your answer to the next three questions on your answer sheet, beginning at number 68. Your responses to these questions will in no way affect your grade, so please be candid.

68. According to the directions given at the start of the test, you were advised to:
1. refrain from answering a question if you were not certain of the answer.
 2. refrain from answering a question if you could eliminate only one or two of the choices.
 3. refrain from answering a question if you had to guess among all choices.
 4. answer all questions even if your answer would represent a sheer guess.
69. Quite apart from what the directions suggested, what did you usually do when you were not certain of the answer to a question?
1. I refrained from guessing.
 2. I guessed only if I had a hunch as to the correct answer or could eliminate at least one choice as incorrect.
 3. I provided an answer even if my answer was a sheer guess among all the choices.
70. Did you adopt a more conservative strategy when taking this test than you would have if your score were to be simply the number that you answered correctly?
1. Yes
 2. No

Appendix G

Winter Final Examination

HISTORY 3052
 Winter 1988
 FINAL EXAM - FORM C

Write your answers on the OPSCAN sheet provided. Use a No. pencil. Each question has only one correct answer. Be sure to erase thoroughly, IF YOU CHANGE AN ANSWER. MARK FORM C ON YOUR OPSCAN SHEET.

DIRECTIONS FOR PART I: FORMULA SCORING (CORRECTION-FOR-GUESSING)

Your score on this part of the test will be the number of questions you answer correctly minus a fraction of the number you answer incorrectly. If you have some knowledge of a question, you may be able to eliminate one or more of the answer choices as wrong. In this case, it is to your advantage to guess which of the remaining choices is correct. However, it is improbable that random or haphazard guessing will change your score significantly.

1. The Confederate brigade that sacrificed itself in securing the Confederate line at the Battle of the Wilderness, but not until Lee had gone behind the lines, was from:
 1. Texas.
 2. Alabama.
 3. Tennessee.
 4. North Carolina.
2. In the 1862 election, Lincoln's Republican Party:
 1. lost control of the House of Representatives.
 2. won control of the House of Representatives with a larger majority than in 1860.
 3. won majorities in the populous states of New York, Pennsylvania, and Ohio.
 4. lost the majority vote in Lincoln's home state of Illinois.
3. The only time that Grant lost his patience in 1864-1865 and ordered a senseless frontal assault was at:
 1. the Crater.
 2. Spotsylvania.
 3. Cold Harbor.
 4. the Wilderness.
4. Sherman's famous march began on November 10, 1864 at Atlanta and ended on March 19, 1865 at:
 1. Savannah, Georgia.
 2. Charleston, South Carolina.
 3. Columbia, South Carolina.
 4. Goldsboro, North Carolina.
5. The Confederate general with whom President Davis quarreled the most violently was:
 1. P. G. T. Beauregard.
 2. A. Sidney Johnston.
 3. Joseph E. Johnston.
 4. Braxton Bragg.

- 2 -

6. The outcomes of the congressional campaigns of 1862 were PRIMARILY influenced by:
 1. conscription.
 2. the "confiscation" act.
 3. military outcomes.
 4. the Emancipation Proclamation.
7. The Union general infamous for his administration of New Orleans, after the Union occupation was:
 1. Banks.
 2. Geary.
 3. Beauregard.
 4. Butler.
8. The most accurate statement about Civil War naval operations is that:
 1. the Union blockade of the South became less important as the war progressed.
 2. blockade-running was not successful because the ships were not fast enough.
 3. the Union navy failed to occupy strategic Southern ports.
 4. the advent of ironclads rendered the world's navies obsolete.
9. Lee's most reliable corps commander in the last year of the war died in battle early on the morning that Grant's army broke through the Confederate lines. That general was:
 1. "Stonewall" Jackson.
 2. A. P. Hill.
 3. J. E. B. Stuart.
 4. George E. Pickett.
10. This general won his greatest victories through siege operations:
 1. Grant.
 2. Sherman.
 3. Lee.
 4. Pickett.
11. The most significant factor leading to Lincoln's re-election in 1864 was:
 1. Fremont's running as a third party candidate, causing a split in the Democratic Party's vote.
 2. the Union occupation of Atlanta.
 3. Horace Greeley's support for Lincoln.
 4. Salmon P. Chase's loss of support in the Republican Party.
12. Which of the following is the LEAST likely to have occurred in the Civil War?
 1. If a man had a broken leg, it would have been amputated.
 2. If a man had a ruptured appendix, it would have been removed.
 3. If a man had an amputation, he would have been sedated with ether.
 4. If a man had a fever, it would have been measured with a thermometer.

- 3 -

13. From class discussion, it is clear that Professor Robertson:
 1. favors re-enactments of Civil War battles because they encourage the study of American history.
 2. favors re-enactments of Civil War battles because they tend to demonstrate the superiority of Confederate generals.
 3. does not favor re-enactments of Civil War battles because they are far too costly.
 4. does not favor re-enactments of Civil War battles because they do not accurately portray the grimness of war.
14. Anti-Lincoln radicals were appeased by the retirement from the cabinet of:
 1. Montgomery Blair.
 2. Andrew Johnson.
 3. Edward Bates.
 4. William Dennison.
15. The glory which the Confederate Navy realized came through the use of:
 1. individual privateers.
 2. the tactical submarine.
 3. ironclad ships.
 4. blockade-runners.
16. The most serious misjudgment of the South, stemming from its isolation was the misjudgment:
 1. of the North's industrial might.
 2. of the deep love for the Union by Northerners.
 3. of England's and France's dependence upon cotton.
 4. over the nature of immigrant participation in the war.
17. When Grant launched his battle plan of 1864, his leader of the Army of the James was:
 1. Sherman.
 2. Meade.
 3. Butler.
 4. Sigel.
18. The 1864 campaign began in Virginia when Lee took advantage of terrain and attacked Grant at:
 1. Manassas.
 2. the Wilderness.
 3. Spotsylvania.
 4. Orange.
19. The last seaport of the Confederacy to fall into Union hands was:
 1. Mobile.
 2. Wilmington.
 3. Charleston.
 4. Galveston.

- 4 -

20. Which of the following was the least significant factor contributing to the food shortage in the South during the war?
 1. poor transportation facilities
 2. the Union blockade of Southern ports
 3. Union raids
 4. Federal occupation of productive lands
21. From the Confederate point of view, the chief obstacle to peace was:
 1. the Northern insistence on emancipation of all slaves.
 2. the prospect of Northern confiscation of Southern property in the event of peace.
 3. the Northern insistence on re-union.
 4. the impending support of the Confederacy by England.
22. In the campaign from the Wilderness to Cold Harbor, Union losses were:
 1. nearly equal to Lee's whole army.
 2. less than Confederate losses.
 3. equal to Confederate losses.
 4. proportionately greater than Confederate losses.
23. The first engagement involving Negro soldiers occurred in:
 1. Pennsylvania.
 2. Missouri.
 3. Tennessee.
 4. Louisiana.
24. The Confederate privateer noted for capturing two dozen Union whalers and not surrendering until seven months after Appomattox was the:
 1. Shenandoah.
 2. Tennessee.
 3. Florida.
 4. Sumter.
25. It can be concluded from class discussion, that, during the Civil War:
 1. prostitutes were found almost solely in the company of officers, rather than enlisted men.
 2. venereal diseases were more prevalent in the North than in the South.
 3. fidelity was more common than infidelity.
 4. letter writing was more important for the soldiers, than for their wives or girlfriends.
26. Approximately what percent of negro soldiers died in the Civil War?
 1. 25
 2. 33
 3. 50
 4. 60

- 5 -

27. He was the leader of the "Copperheads" or Peace Democrats in the North. His name was:
1. C. L. Vallandigham.
 2. Horace Greeley.
 3. G. H. Pendleton.
 4. S. P. Chase.
28. Which of these is NOT true concerning the Northern naval blockade?
1. The South was able to import much of its firearms, artillery, and ammunition from Europe, despite the blockade.
 2. Few of the blockade-runners were actually seized by the Union fleet.
 3. The full effect of the Union blockade is best measured by the number of large commerce ships that could not run the blockade.
 4. The blockade-runners, as well as being fast, could carry a great deal of cargo.
29. Prior to the end of the war, blacks officially:
1. served only in the Union army.
 2. served only in the Confederate army.
 3. served in both armies.
 4. served in neither army.
30. This organization may be regarded as the forerunner of the Red Cross:
1. the army ambulance corps.
 2. the Christian Commission.
 3. the United States Sanitary Commission.
 4. the Women's Medical Care Auxiliary.

- 6 -

DIRECTIONS FOR PART II: NUMBER-RIGHT SCORING

Your score on this part of the test will be the number of questions answered correctly. It is to your advantage, therefore, to answer every question in this part of the test, even if your answer to some questions is a sheer guess.

31. When Lee surrendered at Appomattox, he had approximately how many men left in his immediate army?
 1. 6,000
 2. 16,000
 3. 26,000
 4. 46,000
32. When Grant took command of the Union armies, his idea of warfare was:
 1. to push for dramatic and climactic engagements with smaller forces.
 2. to conduct war as chess instead of checkers.
 3. to drive after enemy armies rather than aim for enemy cities.
 4. to assume the defensive and force Lee to come to him.
33. Desertion was a significant factor during the Civil War. According to the text, the most accurate statement regarding this factor is:
 1. deserters usually just found their way home and sat out the remainder of the war.
 2. deserters on both sides were primarily conscripts.
 3. desertion in the Union armies was a factor primarily limited to the Army of the Potomac.
 4. desertion in the Confederate armies was more extensive than in the Union armies.
34. The people in the U. S. government who stood to gain the most from Lincoln's death were the:
 1. conservative Democrats.
 2. radical Republicans.
 3. moderate Republicans.
 4. the presidential Cabinet members.
35. The last important battle of the war for Lee's army was at:
 1. Wilmington, North Carolina.
 2. Five Forks, Virginia.
 3. Saylor's Creek, Virginia.
 4. Bentonville, North Carolina.
36. The general most responsible for whatever success there was in the integration of blacks in federal armies was:
 1. Grant.
 2. Thomas.
 3. Sheridan.
 4. Sigel.

- 7 -

37. Of the following, who played the most important role in trying to unseat Jefferson Davis as president of the Confederacy?
1. Robert E. Lee
 2. P. G. T. Beauregard
 3. William L. Yancey
 4. Alexander H. Stephens
38. When soldiers fell wounded in battle, their chances of survival were:
1. 1 of 2.
 2. 1 of 4.
 3. 1 of 10.
 4. 1 of 43.
39. After the fall of Vicksburg, the key to Union strategy in the West was the city of:
1. Memphis.
 2. New Orleans.
 3. Chattanooga.
 4. Jackson.
40. The "Battle of the Bloody Angle" was fought at:
1. Cold Harbor.
 2. Spotsylvania.
 3. Clark's Mountain.
 4. New Market.
41. The 1864 Democratic candidate for President was:
1. Salmon P. Chase.
 2. Horace Greeley.
 3. George B. McClellan.
 4. Clement L. Vallandigham.
42. "King Cotton" was:
1. comparable in importance to the state's rights issue to Southerners.
 2. an argument used by abolitionists to oppose slavery.
 3. such an important issue that the Confederate Congress passed a Cotton Embargo Act.
 4. such an important issue that the Federal Congress passed a Cotton Embargo Act.
43. Flaming alcohol was often used to treat:
1. pneumonia.
 2. wounds caused by the "Minie" ball.
 3. amputation infections.
 4. venereal disease.
44. From class discussion, it seems clear that the most significant event of the Civil War was:
1. the battle of Gettysburg.
 2. the Confederate surrender at Vicksburg.
 3. the appointment of Grant as supreme commander of Union forces.
 4. the way in which Lee surrendered the Army of Northern Virginia.

- 8 -

45. At Appomattox, Lee chose not to resort to guerrilla warfare because:
 1. guerrilla warfare would not bring the South its independence.
 2. his army was trapped.
 3. he did not have a force sufficient to fight guerrilla tactics.
 4. he knew that Grant would offer "can't refuse" peace terms.
46. A fundamental motive in Southern diplomacy with the nations in Europe was reliance upon:
 1. continued military victories in the field.
 2. the ability to keep Southern ports open.
 3. the need for cotton in Europe.
 4. the persuasiveness of Southern statesmen.
47. For the year 1864, the best description of the Army of the Potomac was that:
 1. it was successful because of the inclusion of the added manpower provided by blacks and conscripts.
 2. it was unsuccessful because of the dispute over leadership between Grant and Meade.
 3. it was successful because of the new plans of battle provided by Grant.
 4. it was unsuccessful because of the loss of experienced soldiers.
48. The Confederate general noted for success at the battle of the Crater was:
 1. William Mahone.
 2. George Pickett.
 3. J. B. Hood.
 4. Wade Hampton.
49. During the summer of 1864 while Union forces beseiged Petersburg, a Confederate corps under the command of Gen. Jubal Early:
 1. surrendered en masse to Grant's army because Early had stoutly opposed secession in the first place.
 2. successfully engaged Union forces at New Market, with the help of VMI cadets.
 3. conducted a raid upon Washington and came within sight of the Capitol.
 4. conducted the "Trevilian Raid" north and west of Richmond.
50. Men and women of the 1860's loved each other more intensely than today primarily because:
 1. mass media was not as dominating then, as it is now.
 2. Victorian customs fostered deep romance.
 3. life was much simpler then.
 4. the "feminist" movement had not begun.

- 9 -

51. Which of the following is the most accurate description of Lincoln's religious beliefs?
1. He was essentially indifferent to religion, holding no particular views.
 2. He was an atheist.
 3. He was a pantheist.
 4. He was a regular adherent to Christian worship.
52. Of the lower southern states, the one from which the largest number of men enlisted in the Union Army was:
1. Alabama.
 2. Florida.
 3. Texas.
 4. Louisiana.
53. George Root is most remembered as a:
1. songwriter.
 2. Confederate statesman.
 3. leading surgeon.
 4. negro soldier.
54. In addition to the Congressional elections of 1862, Lincoln faced another serious political crisis which was:
1. the Supreme Court decision in the Dred Scott case.
 2. a threat by members of his own party to make significant changes in his cabinet.
 3. the declaration by Vice President Hamblin that he no longer supported Lincoln's policies.
 4. the resignation of McClellan.
55. The USS Kearsage trapped the USS Alabama and sunk it in the port of:
1. Cherbourg.
 2. Mobile.
 3. New Orleans.
 4. Lisbon.
56. The most distinguished fighter in C. S. Navy was:
1. Stephen Mallory.
 2. Raphael Semmes.
 3. David Farragut.
 4. Horace Hunley.
57. According to the text, the Confederacy in time came to look on Europe primarily for:
1. warships.
 2. money.
 3. medicines.
 4. additional manpower.

- 10 -

58. The least effective trench pattern used in the Civil War was called:
1. salient.
 2. serial.
 3. straight.
 4. semi-inverted.
59. The last of the Southern forces to surrender were:
1. Joseph E. Johnston's at Durham, North Carolina.
 2. Nathan Bedford Forrest's at Montgomery, Alabama.
 3. Richard Taylor's at Citronelle, Alabama.
 4. Kirby Smith's at New Orleans, Louisiana.
60. Lee's overstretched defensive line around Petersburg was finally breached at the Battle of:
1. Fort Steadman.
 2. Bermuda Hundred.
 3. Dinwiddie Court House.
 4. Five Forks.

Appendix H

Addendum to Winter Final Examination (Part C)

- 11 -

PART III:

We are very interested in your responses to the following questions. Please answer the questions to the best of your ability with the knowledge that your answers to these questions will not affect your grade in any way.

61. Quite apart from what the directions suggested, how did you usually respond when you were uncertain of the answer to a question on the part of the test that is scored with a correction for guessing?
 1. I refrained from guessing.
 2. I guessed only if I had a hunch as to the correct answer or could eliminate at least one choice as incorrect.
 3. I provided an answer even if my answer was a sheer guess among all the choices.
62. On the formula scored part of the test, did you adopt a more conservative strategy (guess less) than for the part to be scored number right (no penalty for incorrect answer)?
 1. Yes
 2. No
63. Is it your understanding that formula scoring (subtracting points for incorrect answers) designed to penalize those who disregard the directions and make sheer guesses?
 1. Yes
 2. No
64. Do you believe that there is an element of dishonesty in guessing when unsure of the answers to multiple-choice questions, regardless of how the test is scored?
 1. Yes
 2. No
65. If you did not have time to read and consider some of the questions on a multiple-choice test, would you mark choices for these questions?
 1. not at all
 2. only if number-right scoring was to be used
 3. even if formula scoring was to be used
66. Outside of this class, which of the following best describes the instruction or advice that you have received relative to guessing on formula-scored multiple-choice tests?
 1. Don't leave any questions blank - answer every question, whether it is a sheer guess or not.
 2. If you have no idea as to the answer, leave it blank; but otherwise, answer the question.
 3. Never guess, answer only the questions that you're sure of.
 4. I don't recall ever receiving advice relative to guessing on formula-scored multiple-choice tests.

- 12 -

67. What or who was the source of this instruction or advice?
1. I have not received such advice or instruction.
 2. special class on test-taking skills
 3. advice of a teacher, other than in a special class
 4. advice of a parent
 5. advice of a friend
68. You attended:
1. public high school
 2. parochial, or denominational, high school
 3. private high school, other than denominational
69. Which of the following best describes your reaction to formula scoring of multiple-choice tests? (choose only one)
1. I don't care how the test is scored.
 2. I think it is a good idea because it keeps people from guessing, and it does not bother me.
 3. I think it is a good idea because it keeps people from guessing, but it bothers me.
 4. I think it is a bad idea because you should get credit for every answer that you get right.
 5. I think it is a bad idea because I don't do as well when I have to be concerned about losing points for a wrong answer.
70. You are a:
1. male.
 2. female.

Appendix I

Oral Instructions for Winter Final Examination

INSTRUCTIONS FOR WINTER FINAL EXAMINATION

PRELIMINARY ORAL INSTRUCTIONS

The test you are about to take was made up by Professor Robertson and consists of 60 multiple-choice questions. Four forms of the test have been prepared: Forms A, B, C, and D. The same 60 questions appear on each form. The only difference is the order of the questions.

Each form has two sets of directions, one for the first 30 questions and a different set for the last 30 questions. Not only are the directions different for the first and second halves, but each half will be scored differently. Therefore it is very important that you read both sets of directions carefully when you get your test questions.

Your score on this test will be the sum of your scores on the two halves of the test. Grades will be assigned to the scores as is the usual practice for this course.

Appended to the end of the final examination are 10 additional questions designed to assess your beliefs about multiple-choice tests. Please record your answers to these questions on the same machine-readable answer sheet you used for the test beginning with number 61.

**The two page vita has been
removed from the scanned
document. Page 1 of 2**

**The two page vita has been
removed from the scanned
document. Page 2 of 2**