To the Graduate Council:

I am submitting herewith a thesis written by James A. Eison entitled "An Investigation of the Alleged Position Response Set in Multiple-Choice Tests with a Built-in Control for the Variable of Response Attractiveness." I recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Arts, with a major in Psychology.

Raymond R. Shrader
Raymond R. Shrader, Major Professor

We have read this thesis
and recommend its acceptance:

Fumiko Samejima

Accepted for the Council:

Hilton A. Smith
Vice Chancellor
Graduate Studies and Research

AN INVESTIGATION OF THE ALLEGED POSITION RESPONSE SET IN MULTIPLE-CHOICE TESTS WITH A BUILT-IN CONTROL FOR THE VARIABLE OF RESPONSE ATTRACTIVENESS

A Thesis

Presented for the

Master of Arts

Degree

The University of Tennessee

James A. Eison

December 1975

1266397

ACKNOWLEDGMENTS

I would sincerely like to thank Drs. Raymond R. Shrader, Schuyler W. Huck, and Fumiko Samejima for their scholarly help and valuable assistance on this project.

And to my parents, whose unfaltering support and encouragement has been with me throughout my graduate career, I can only say I love you both.

ABSTRACT

Several research efforts investigating the possible existence of a position response set in multiple-choice examinations have previously produced equivocal results. The present investigation represents an improvement over previous studies in that (a) it employed a research design which controls the possible contaminating effects of differential response attractiveness, and (b) the data were subjected to a proper statistical analysis. The results of two separate experiments (one using traditional vocabulary test items, the other using a regular class examination) failed to substantiate the claim that a position response set affects an examinee's test-taking behavior.

TABLE OF CONTENTS

CHAPTER		PAGE
I.	INTRODUCTION	1
II.	EXPERIMENT 1	10
	Method	10
	Subjects	10
	Materials	10
	Procedure	11
	Results	12
III.	EXPERIMENT 2	14
	Method	14
	Subjects	14
	Materials	14
	Procedure	15
	Results	16
IV.	DISCUSSION	20
REFERENCES		27
APPENDIXES		30
	Appendix A	31
	Appendix B	33
	Appendix C	45
	Appendix D	47
	Appendix E	48
	Appendix F	52
VITA		55

LIST OF TABLES

TABLE		PAGE
1.	Chi-Square Values for the Nine "Pure Guess" Items in Experiment 1	13
2.	Chi-Square Values for the Five "Pure Guess" Items in Experiment 2	17
3.	Chi-Square Values for the Five "Pure Guess" Items in Experiment 2 Performed on the Lower 25 Percent of the Class	18
4.	Chi-Square Values for the Five "Pure Guess" Items in Experiment 2 Performed on the Upper 25 Percent of the Class	19
5.	Chi-Square Values for the Nine "Pure Guess" Items in Experiment 1--A Comparison of Alternative Attractiveness.	45
6.	Chi-Square Values for the Five "Pure Guess" Items in Experiment 2--A Comparison of Alternative Attractiveness.	47
7.	Chi-Square Values for the Nine "Pure Guess" Items in Experiment 1--An Alternative by Alternative Analysis.	48
8.	Chi-Square Values for the Five "Pure Guess" Items in Experiment 2--An Alternative by Alternative Analysis.	52

CHAPTER I

INTRODUCTION

Response sets have been defined as "any tendency causing a person consistently to make different responses to test items than he would have made had the same content been presented in a different form" (Cronbach, 1946). Probably the most frequently investigated form of response set has been the position response set, which operationally has been defined as a "deviation from chance at or beyond the 5 percent level of confidence for the number of answers selected at each position" (Rapaport and Berg, 1955). For example, in a forty item multiple-choice examination, chance would maintain that, on the average, each individual would select ten options at each of the four positions, while a significant deviation from this distribution would indicate the presence of a position response set.

Previous experimental research that has attempted to investigate this phenomenon has consistently yielded contradictory results. A basic methodological difference in these studies can neatly divide them into two groups: those that have used highly unstructured tests, such as "questionless questionnaires" or "imaginary inventories," and those that have used "real" objective tests in natural settings.

Looking first at studies that have used unstructured tests, Rapaport and Berg (1955) report that "in an imaginary questionnaire in which subjects were to select one of four options when no actual question was stated, over 60 percent of those responding chose the third position" (p. 59). In an unpublished doctoral dissertation, Metfessal (1955) also found evidence of a position response set when nonsense syllables were used as responses to test questions (cited in Wilbur, 1970). Other studies involving highly unstructured tests that support the existence of a non-random preference for certain positions include Goodfellow (1940) and Whitfield (1950).

Turning to those studies which have used genuine objective style exams and which have found evidence of a position response set, McNamara and Weitzman (1945) examined the results of tests given at Navy Flight Preparatory Schools and found that the placement of choices in the various response positions has a small but statistically significant effect upon the difficulty level of the questions. Using a 50 item questionnaire that examined student attitudes towards various study habits, Mathews (1927) showed that "the position of a printed response on a questionnaire with five possible responses influenced individuals' answers to items on the blank" (p. 134). Gustav (1963) found

similar results with objective classroom tests as well.

However, contrary to these findings, several investigators have found no evidence of a position response set in other studies that have similarly used multiple-choice examinations of a meaningful nature. For example, Rapaport and Berg (1955) administered a seventy item test of general knowledge to 284 college students. Forty of the four option multiple-choice questions dealt with material so dificcult that only by chance would a subject choose the correct answer. As noted previously, the authors suggest that chance would maintain that an individual is expected to select ten options at each of the four response positions. To identify a response bias, chi-square values were computed by comparing the expected with the observed frequencies of option choices for every subject. Rapaport and Berg found "that there was no generalized tendency for such biased responses to appear" (p. 61). Hopkins and Hopkins (1964) and Wilbur (1970) also found no evidence that might suggest the existence of a position response set.

Following a somewhat different approach, Marcus (1963) administered four different forms of a previous years' final examination to some 400 students enrolled in an introductory psychology class. On each form the arrangement of correct choices was randomized throughout

the test by a scheme that allowed each position an equal number of correct choices, and which placed the correct choice for each item in a different position on each of the four forms. Though Marcus did not find a single consistent position response set favoring one position over the four forms, he did observe a significant difference over all forms between observed and expected frequencies for total number of responses to each position. He attributes this finding to the possibility that the most plausible distractors within each test form were not evenly distributed among the four possible positions. Marcus suggests that one promising line of research would be to "attempt to get at a method for equalizing the attractiveness of the distractors" (p. 51).

Before a clarifying or definitive statement concerning the position response set can be made, however, there are several aspects of these many studies that need to be examined and questioned more closely. I will address myself now to three issues of particular concern.

At least two studies (Rapaport and Berg, 1955; Clark, 1956) which failed to find a position response set, and at least one study which did reveal a position response set (Gustav, 1963), used statistical analyses that were based on the assumption that if a position response set were not present, the distribution of

incorrect responses that the group would make would not differ significantly from chance. It is this writer's contention, however, that this type of examination of the distribution of incorrect responses cannot provide a clear and convincing demonstration of the alleged position response set that would be free from alternative explanations. In fact, any number of possible extraneous factors, other than a position response set, might account for a student's non-random selection of incorrect answers. For example, as is often the case, a distractor may bear a close resemblance to something other than the correct answer that the instructor may have said or that the student vaguely remembers reading. Or perhaps one or more of the choices may contain some specific determiner that would clue the student to either choose or avoid the alternative. In short then, some of the incorrect answers will appear to be more attractive than others, for reasons entirely other than their position, and will consequently be chosen more frequently.

Furthermore, in that few classroom test constructors, and few researchers of the position response set, have taken the time and effort to insure that each alternative has been randomly assigned to each of the positions, it might be hypothesized that if there were a tendency for the test author to assign, either consciously or

unconsciously, the more attractive distractors to any one of the response positions, one would then expect that position to be chosen non-randomly over the other positions.

Thus, to statistically test the notion of a position response set against the assumption that the total number of incorrect answers should be equally divided among the alternative positions seems unjustifiable. This researcher suggests that the true question to which an experimental investigation of the position response set must directly address itself to is "When guessing, do students select a given response solely on the basis of its particular position or location within the cluster of alternatives, irrespective of the level of attractiveness it may have."

The second issue concerns the use, or possible misuse, of the chi-square statistic by several researchers who have investigated position response sets. For instance, some researchers (Clark, 1956; Marcus, 1963) have added together the frequencies with which correct answers, and/or distractors, have been chosen from each of the individual questions contained on the test, and then used the chi-square statistic to test if any one position were chosen significantly more often than the others. I suggest that this is incorrect for two reasons.

First, a fundamental assumption of the chi-square statistic requires that there be independence between each of the events being observed. If one acknowledges that a student's answer to any individual test question might have been influenced in any way by his answers to the previous questions, then to sum the total number of responses as described above would be improper. Sometimes information contained in one question may lead a student to answer a subsequent question in a particular manner; sometimes students look for symmetrical patterns on their answer sheets which might clue them in on how to guess on a particular item; and sometimes, if you acknowledge that a position response set does indeed exist, students may consistently choose a particular response position. And so, to use the chi-square statistic when possible linkages or connections between test answers may exist would be to violate this most basic assumption.

The second problem with summing across items is that the value of N (the total number of responses) in the chi-square formula becomes inflated, thus increasing power unjustifiably. Douglas (1969) has demonstrated that doubling the total number of responses used in a chi-square analysis, while holding constant the specific proportion of responses found in each cell, can lead a researcher to a conclusion different from that

which would be found if only half of the total number of cases were used. And so, in a hypothetical example, summing the frequencies in the manner previously described would mean that if a fifty item multiple-choice test were given to 100 students, the value of N in the resulting chi-square formula would become 5000; considerably higher than the N=100 that would be obtained if the same chi-square statistic were to be applied to any single test question. Thus, the effect of having the same individuals reply fifty times might well be to make very small differences appear statistically larger, and of greater significance, than they actually are.

A third, and final, concern with previous studies involves the fact that two investigations that have failed to show the position response set (Hopkins and Hopkins, 1964; Wilbur, 1970) were conducted with subjects drawn from either elementary school or high school populations. One must question if their results can be generalized to a more practiced, more sophisticated, and test-wise population of college freshmen and sophomores.

The purpose of this study is to investigate the existence of an alleged position response set in college-level multiple-choice tests using a research design which controls for the possible contaminating

influence of differential response attractiveness. Furthermore, the present investigation employed an approach to statistical analysis that avoided problems inherent in some previous research studies.

To provide for greater generalizability of the results, the present study was conducted in two separate phases. Experiment 1 used a traditional vocabulary test similar to that used by Jones (1972). Experiment 2 used a regular course exam in a natural setting to assure a high level of student motivation to achieve success.

CHAPTER II

EXPERIMENT 1

Method

Subjects. The subjects (Ss) used in Experiment 1 were 88 undergraduate students (39 males and 49 females) enrolled in an introductory psychology course at The University of Tennessee during the spring quarter, 1974. Participation in this study was voluntary, with all Ss receiving extra-credit in their psychology course for their help.

Materials. Four modified versions of the 45-item multiple-choice vocabulary test used by Jones (1972) were prepared. Of these 45 items, nine were "pure guess" items in that the stem words were chosen from a list of words that appear less than four times in 18,000,000 (Thorndike and Lorge, 1944). To insure that the Ss did not, by chance, know these words, the one-word response alternatives to these nine critical items were neither synonyms nor antonyms of the stem word.

The four versions of the vocabulary test were constructed such that each response alternative to the nine "pure guess" items appeared once in each of the

four possible positions. The following balanced Latin square design was used to accomplish this goal:

	Choice Order			
	1st	2nd	3rd	4th
Form W	a	b	c	d
Form X	d	c	b	a
Form Y	c	a	d	b
Form Z	b	d	a	c

wherein the letters a through d stand for the response alternatives. By using the four form format and systematically varying the position the four response alternatives would occupy across each different form, it would now be possible to statistically compare the total number of times each response position was chosen while the possible confounding effects of differential response attractiveness would be controlled by the counterbalancing scheme.

Procedure. All subjects were randomly assigned one of the four forms of the examination which they completed during a single thirty-minute time period. The test was described to the Ss as being "similar to those used on graduate school admissions tests." Subjects were further instructed that although there were a few very hard items which almost no one would answer correctly, it was essential that a response be

made to every item on the test no matter how unfamiliar the word seemed.

Results

For each of the nine critical items, the responses from all Ss were collapsed across the four forms. This resulted in a total frequency for each of the four response positions, these data being uncontaminated by response attractiveness due to the counterbalancing technique employed in developing the four forms.

A one-item chi-square test was applied to the data generated by each of the nine "pure guess" items (see Table 1). The resulting values of the test statistic ranged from 0.7 to 5.5, with none exceeding the .05 critical value (df = 3) of 7.817. These results indicate that the frequency of responses to the four choice positions did not differ beyond what could be expected by chance. Even for the item having the largest calculated value (5.5), the four frequencies did not seem overly disparate: 30 (a), 24 (b), 18 (c), and 16 (d).

TABLE 1

CHI-SQUARE VALUES FOR THE NINE "PURE GUESS" ITEMS IN EXPERIMENT 1

Items	Response Position 1	2	3	4	χ^2	df	sig
#1							
Obtained Frequency*	28	18	26	16	4.72	3	NS
Expected Frequency	22	22	22	22			
#2							
Obtained Frequency	19	29	22	18	2.81	3	NS
Expected Frequency	22	22	22	22			
#3							
Obtained Frequency	15	22	30	21	5.13	3	NS
Expected Frequency	22	22	22	22			
#4							
Obtained Frequency	17	25	21	25	2.00	3	NS
Expected Frequency	22	22	22	22			
#5							
Obtained Frequency	30	24	18	16	5.45	3	NS
Expected Frequency	22	22	22	22			
#6							
Obtained Frequency	23	26	15	24	3.18	3	NS
Expected Frequency	22	22	22	22			
#7							
Obtained Frequency	22	20	29	17	3.54	3	NS
Expected Frequency	22	22	22	22			
#8							
Obtained Frequency	24	23	19	22	0.63	3	NS
Expected Frequency	22	22	22	22			
#9							
Obtained Frequency	26	26	23	13	5.17	3	NS
Expected Frequency	22	22	22	22			

*The Obtained Frequency values correspond to the frequencies for each Response Position summed across the four forms.

CHAPTER III

EXPERIMENT 2

Method

Subjects. The Ss used in Experiment 2 were 136 undergraduate students (63 males and 73 females) enrolled in a single section of the same psychology course during the summer quarter, 1974. Unlike the first experiment, the Ss here were unaware of their actual participation in a psychological experiment.

Materials. Four versions of a 45-item multiple-choice test covering the course reading assignments were prepared. Of these items, five were "pure guess" items. Although the stem and accompanying response alternatives for these five items had apparent face validity with respect to the other forty items, they were in all other ways imaginary and meaningless, with no scorable correct answer. For example, embedded within ten genuine items dealing with prejudice was the following "pure guess" question:

Which of the following is most closely related to simple prejudice?

(a) interpersonal verticality

(b) analitic identification

(c) interpersonal contrients

(d) pseudo-maxi identification

The same procedure as used in the first experiment was employed in constructing the four forms of the test for Experiment 2. The principle of the balance Latin square caused each of the nonsense choices of the five "pure guess" items to appear once in each of the four response positions. This arrangement controlled for the possibility of "response attractiveness" becoming confounded with the alleged "position" response set.

Procedure. The Ss were randomly assigned to the four forms of the test. Though the class was aware of the fact that multiple forms of the same exam had been used in the past to safeguard against possible cheating, none of the Ss were aware of the true differences between the four forms. Likewise, the Ss were completely naive to the experimental question under investigation. To further maintain this naivete and to minimize any possible anxiety that might have been aroused in the minds of the brighter students upon finding five unfamiliar questions, the class was instructed that the exam contained five items which were very difficult. Correct answers to these questions would earn extra credit; however, incorrect answers would definitely not be

penalized nor would they in any way affect or lower the distribution of grades.

Results

As in Experiment 1, the responses from the total group of Ss were collapsed across forms, with a one-item chi-square being applied to the resulting data of each "pure guess" item. Four out of these five analyses yielded non-significant results, with the remaining chi-square value being significant at the .05 level (see Table 2).

To determine whether a position response set might be associated only with students who are low in ability or who don't study, the total group of 136 Ss was subdivided into quartiles on the basis of their scores on the forty real test items. For the top and bottom groups (n = 34 in each case), the responses to the five "pure guess" items were examined after collapsing across forms. Of the ten chi-square tests conducted, none of the calculated values exceeded the .05 critical value (see Tables 3 and 4).

TABLE 2

CHI-SQUARE VALUES FOR THE FIVE "PURE GUESS" ITEMS IN EXPERIMENT 2

Items	Response Position 1	2	3	4	χ^2	df	sig
#1							
Obtained Frquency*	43	27	33	32	3.97	3	NS
Expected Frequency	34	34	34	34			
#2							
Obtained Frequency	34	36	37	28	1.44	3	NS
Expected Frequency	34	34	34	34			
#3							
Obtained Frequency	40	28	42	20	9.76	3	$p<.05$
Expected Frequency	34	34	34	34			
#4							
Obtained Frequency	38	34	28	34	2.35	3	NS
Expected Frequency	34	34	34	34			
#5							
Obtained Frequency	36	42	30	28	3.52	3	NS
Expected Frequency	34	34	34	34			

*The Obtained Frequency values correspond to the frequencies for each Response Position summed across the four forms.

TABLE 3

CHI-SQUARE VALUES FOR THE FIVE "PURE GUESS" ITEMS IN EXPERIMENT 2 PERFORMED ON THE LOWER 25 PERCENT OF THE CLASS

Items	Response Position 1	2	3	4	χ^2	df	sig
#1							
Obtained Frequency*	11	4	10	11	3.77	3	NS
Expected Frequency	9	9	9	9			
#2							
Obtained Frequency	10	11	7	8	1.11	3	NS
Expected Frequency	9	9	9	9			
#3							
Obtained Frequency	13	7	10	6	3.33	3	NS
Expected Frequency	9	9	9	9			
#4							
Obtained Frequency	8	15	5	8	6.00	3	NS
Expected Frequency	9	9	9	9			
#5							
Obtained Frequency	11	13	8	4	5.11	3	NS
Expected Frequency	9	9	9	9			

*The Obtained Frequency values correspond to the frequencies for each Response Position summed across the four forms.

TABLE 4

CHI-SQUARE VALUES FOR THE FIVE "PURE GUESS" ITEMS IN EXPERIMENT 2 PERFORMED ON THE UPPER 25 PERCENT OF THE CLASS

Items	Response Position 1	2	3	4	χ^2	df	sig
#1							
Obtained Frequency*	11	7	11	7	1.77	3	NS
Expected Frequency	9	9	9	9			
#2							
Obtained Frequency	7	6	16	7	7.33	3	NS
Expected Frequency	9	9	9	9			
#3							
Obtained Frequency	10	10	10	6	1.33	3	NS
Expected Frequency	9	9	9	9			
#4							
Obtained Frequency	9	4	11	12	4.22	3	NS
Expected Frequency	9	9	9	9			
#5							
Obtained Frequency	8	14	4	10	5.77	3	NS
Expected Frequency	9	9	9	9			

*The Obtained Frequency values correspond to the frequencies for each Response Position summed across the four forms.

CHAPTER IV

DISCUSSION

Unlike the findings of McNamara and Weitzman (1945) and Gustav (1963), the present investigation offers no significant evidence that a position response set influences student "guessing" behavior on genuine, college-level multiple-choice examinations.

To help clarify the controversy surrounding the position response set, Experiment 1 followed a procedure similar to that used by previous researchers whose findings have oftentimes been both confusing and contradictory. A further refinement in Experiment 2 was the fact that the true nature of the five "key items" was concealed from the students, thus avoiding many of the reactive consequences (such as the Hawthorne effect, experimental demand characteristics, etc.) that often result when subjects are aware of the fact that they are in a psychological experiment or that they are being observed. Furthermore, by incorporating the experiment into a regularly scheduled class examination, a high level of student motivation and involvement was presumably insured.

Rather than employ a randomization scheme to assign alternatives to response positions, as did Marcus

(1963), this writer used a Latin square sequence which provided greater control against the possibility that effects attributable to differential response attractiveness might influence, and consequently confound, the effects caused by specific position preferences. The type of counterbalancing technique used in this study, which presented each alternative once in each response position, across the four forms, insured that there would be an equal distribution of the most plausible, or attractive alternatives among each of the four response positions. Thus, when the responses from all the subjects were summed across the four test forms, equality of response attractiveness was achieved, and the total frequency obtained for each of the four response positions could then be examined to assess the effects of a position response set.

It can be mentioned here that Tables 5 and 6 (see Appendixes C and D) provide an examination of the total frequency with which each of the alternatives to the fourteen "pure guess" items was chosen, with responses being summed across the four different forms. Of the nine "pure guess" items used in Experiment 1, two items revealed significant differences in response attractiveness, and of the five "pure guess" items used in Experiment 2, four showed similar significant differences. This suggests that even when the possible influence of

response position is accounted for, the attractiveness of response alternatives was not found to be equal. This was observed more so on the psychology test items used in Experiment 2 than on the vocabulary test items used in Experiment 1. This observation thus suggests the importance of using a counterbalancing technique to assign alternatives to positions in investigations of the position response set.

If the subjects that participated in this study did indeed have a specific position preference, it would appear as a pattern of significant differences in each of the analyses performed on the fourteen "pure guess" items. However, as shown in Tables 1 and 2 (pages 13 and 17, respectively), this clearly was not the case.

It might be noted that on Item #3 in Experiment 2 a significant preference favoring positions 1 and 3 over positions 2 and 4 was observed (see Table 2). This finding is close to what would be expected, with the .05 level of significance predicting a ratio of 1 out of 20 results that will appear to be significant. Further, a post-experimental examination of this item has failed to suggest any alternative explanation that might account for this finding. It should also be noted that when the responses of the upper and lower quartile of the class that participated in Experiment 2 were examined separately,

the results were uniformly negative (see Tables 3 and 4, pages 18 and 19).

Rapaport and Berg (1955) have proposed that "when a student responds to an item he may respond on one of three levels":

> 1. Direct information, in which the subject makes his choice on the basis of sure knowledge, or what he believes to be sure knowledge;
>
> 2. Formal characteristics of the item, in which the subject is not certain of the correct answer and attempts to settle upon one option by eliminating others, using anything which occurs to him, i.e., number of syllables or words, vague familiarity, absolutes like never, always, etc.;
>
> 3. Response sets, in which neither sure knowledge nor formal item characteristics avails; hence, the only remaining basis for choice is the subjects "feeling" about how he might guess in choosing an option (p. 60).

In the present two investigations the use of unanswerable "pure guess" items insured that all subjects would have no direct information concerning the correct answers. Further, the four test forms with their balanced Latin square design allowed this researcher to control against the possibility that the influence of formal characteristics of the item (which can be thought of as being equivalent to the sum of all of the factors which comprise response attractiveness) might become confounded with the effects of a position response set. It appears that when the influence of response attractiveness is experimentally controlled for, there is little

support for the notion that students will respond to a single alternative position in a consistent manner.

Has the question of the position response set been at last put to rest? Anecdotally, I might briefly mention a rather startling conversation that occurred after I explained the experimental question that was being investigated to the subjects who had participated in Experiment 2. A female student meekly raised her hand and informed the class that Glamour Magazine's annual "How to do it better" issue, which was on sale at the University bookstore, had an article entitled "Pass multiple-choice tests when you don't know the answers." Citing the advice of a researcher in testing and education, a Mr. S. Hoffman, the first rule of thumb that the article offered was, "The correct answer tends to be the third choice if there are five alternatives, and the second or third choice if there are four alternatives" (p. 81). The rationale given for this suggestion was that a teacher will try to "bury the correct answer among the alternate response distractors." At that point I could only conclude that perhaps if more of the class had read Glamour Magazine prior to their participation in these studies, the results of my investigations might have been markedly different.

Two additional considerations that might be of interest to future investigations of the position

response set might also be noted. The first concern is the fact that this study examined the total group's replies in search of a position preference without examining within, or intra-individual, choices. The rather limited number of "pure guess" items used in these two studies prevented a statistically meaningful analysis of this type of data. A closer look at intra-individual preferences, on an exam which contains a far greater number of "pure guess" items might be interesting.

The second consideration involves the possibility that the particular sequential arrangement of alternatives may influence a student's guessing behavior. Though the original intent and design of this study did not directly address itself to this question, a post-hoc analysis of an exploratory nature can be provided. When each of the response alternatives was examined separately (see Appendixes E and F) and the frequency with which each alternative was selected (as its position varied across the four forms) was statistically compared, the following was found. Out of fifty-six chi-square tests that resulted when the four alternatives to each of the fourteen "pure guess" items were examined separately, twelve were found to be significantly different, at the .05 level, from the distribution of responses predicted by random guessing.

Unfortunately, in that many of these fifty-six chi-square tests involved very low expected frequencies, the meaningfulness of such tests must be seriously questioned. A further examination of this data can be offered, however, by directing attention to only those chi-square tests that do have a large enough expected value, which arbitrarily might be set at seven. When this is done, it appears that of the nineteen chi-square tests that can be examined, only three reach the significance level of .05. This finding appears to be more in line with the major findings and conclusions of this investigation discussed previously.

REFERENCES

REFERENCES

Berg, I. A. and Rapaport, G. M. Response bias in an unstructured questionnaire. The Journal of Psychology, 1954, 38, 475-481.

Clark, E. L. General response pattern to five-choice items. Journal of Educational Psychology, 1956, 47, 110-117.

Cronbach, L. J. Response sets and test validity. Educational and Psychological Measurement, 1946, 6, 475-494.

Gustav, A. Response set in objective achievement tests. The Journal of Psychology, 1963, 56, 421-427.

Hopkins, K. D., and Hopkins, B. R. Intraindividual and interindividual positional preference response styles in ability tests. Educational and Psychological Measurement, 1964, 24, 801-805.

Jones, P. D. The differential formation of response sets by specific determiners. (Doctoral dissertation, The University of Tennessee, Knoxville.) Dissertation Abstracts International, 1972, 33, 05A, 2238. (University Microfilms No. 72-27477.)

McNamara, Lt. W. J., and Weitzman, Lt. E. The effect of choice placement on the difficulty of multiple-choice questions. Journal of Educational Psychology, 1945, 36, 103-113.

Marcus, A. The effect of correct response location on the difficulty level of multiple-choice questions. Journal of Applied Psychology, 1963, 47, 48-51.

Mathews, C. O. The effect of the order of printed response words on an interest questionnaire. Journal of Educational Psychology, 1927, 18, 445-457.

Metfessel, N. S., and Sax, G. Systematic biases in the keying of correct responses on certain standardized tests. Educational and Psychological Measurement, 1958, 18, 787-790.

Rapaport, G. M., and Berg, I. A. Response sets in multiple-choice tests. Educational and Psychological Measurement, 1955, 15, 58-62.

Wevrick, L. Response set in a multiple-choice test. Educational and Psychological Measurement, 1962, 22, 533-538.

Wilbur, P. H. Positional response set among high school students on multiple-choice tests. Journal of Educational Measurement, 1970, 7, 161-163.

APPENDIXES

APPENDIX A

SAMPLE TEST FORM USED IN EXPERIMENT 1

FORM 1

The following is a vocabulary test similar to those used on Graduate School Admissions Tests. There are a few very easy items which nearly everyone will get right and a few very hard items which almost no one will answer correctly. For our research purposes we need to have an answer for every item on the test, so no matter how unfamiliar a word may be, please record the best response you can, even if it is a pure guess. It is also critical to our research that you work straight through the test and not skip around in answering the questions. Please place all your answers on the answer sheet. Do not write on the test itself.

ON YOUR ANSWER SHEET PLACE THE NUMBER OF THE WORD WHICH HAS THE MEANING CLOSEST TO THE FIRST (CAPITALIZED) WORD FOR EACH ITEM.

1.	INJURY	1) damage 2) decision 3) investigation 4) question
2.	MATURE	1) balanced 2) huge 3) adult 4) sincere
3.	SOLID	1) trusted 2) heavy 3) square 4) firm
4.	CONCEAL	1) imagine 2) hide 3) invasion 4) vice
*5.	APOTHEGM	1) constellation 2) overture 3) journey 4) danger
*6.	CORIANDER	1) chemist 2) stolidity 3) predicament 4) pillar
7.	LUDICROUS	1) beggarly 2) ridiculous 3) doleful 4) drowsy
8.	IMPERCEPTIBLE	1) immune 2) imperious 3) intangible 4) impertinent
9.	ROBUST	1) healthy 2) automatic 3) bankrupt 4(blithe
10.	TEDIOUS	1) complex 2) magnetic 3) orderly 4) tiresome
*11.	FUSTY	1) particular 2) comical 3) pastoral 4) gaudy
12.	FABLE	1) myth 2) invention 3) monster 4) oracle
13.	FRAGMENT	1) portion 2) conflict 3) crystal 4) structure
14.	DISPERSE	1) incline 2) scatter 3) regulate 4) design
15.	ACCORD	1) conclusion 2) agreement 3) contentment 4) report
16.	CONTEMPLATE	1) consider 2) regulate 3) confound 4) design

*Denotes "pure guess" items.

*17.	DILURIAL	1) gaily 2) distant 3) caustic 4) irreverent
18.	FLACCID	1) flaky 2) flashy 3) flabby 4) flawless
19.	OSCILLATE	1) castigate 2) coagulate 3) luxuriate 4) fluctuate
20.	THWART	1) beguile 2) induce 3) foil 4) disguise
21.	ACCESS	1) emphasis 2) restraint 3) entry 4) submission
22.	IMITATE	1) disclose 2) introduce 3) copy 4) reduce
*23.	GLAIVE	1) coward 2) shine 3) valley 4) trick
24.	CAPACITY	1) melody 2) slope 3) corruption 4) ability
25.	INSINUATE	1) affront 2) beguile 3) imply 4) fluctuate
26.	DIMINISH	1) penetrate 2) extend 3) discourage 4) decrease
27.	ADVERSARY	1) opponent 2) defender 3) lunatic 4) radical
28.	SIMULATE	1) liberate 2) feign 3) retaliate 4) bequeath
*29.	ESCRITOIRE	1) salad 2) fortitude 3) journey 4) clearness
30.	REFRIGERATE	1) solidify 2) pay 3) chill 4) settle
31.	MEAGER	1) acute 2) scanty 3) ample 4) needy
32.	SCARE	1) examine 2) tremble 3) bump 4) alarm
33.	INDUCE	1) appropriate 2) emerge 3) fascinate 4) persuade
34.	FORLORN	1) elastic 2) desolate 3) exposed 4) alien
*35.	UMBRA	1) lagoon 2) mushroom 3) pain 4) report
36.	HEX	1) bewitch 2) unlock 3) agree 4) perpetuate
37.	IMMORTAL	1) enormous 2) mysterious 3) deathless 4) evil
38.	EXILE	1) isolate 2) decay 3) punish 4) banish
39.	INCESSANT	1) perpetual 2) hysterical 3) orthodox 4) intricate
40.	SELECT	1) appear 2) choose 3) persuade 4) suggest
*41.	VIRAGO	1) connoisseur 2) poison 3) earthquake 4) alcoholic
42.	DISCLOSE	1) conceal 2) reveal 3) conclude 4) open
43.	GHASTLY	1) grim 2) furious 3) gigantic 4) haunted
44.	MENACE	1) disagree 2) promote 3) rustle 4) threaten
*45.	IMMERSE	1) hoist 2) redeem 3) submerge 4) refresh
46.	SIFFLE	1) burn 2) blow 3) strike 4) launder

*Denotes "pure guess" items.

APPENDIX B

SAMPLE TEST FORM USED IN EXPERIMENT 2

QUIZ 2 FORM A

PROFESSOR: Ray Shrader INSTRUCTOR: Jim Eison

COORDINATOR: Steve Kennedy

PSYCHOLOGY 2130

SPRING 1974

DIRECTIONS: Put your name on BOTH TEST BOOKLET AND ANSWER SHEET. On the reverse side of the answer sheet, fill in the boxes with your name and student ID number (Social Security #). Use pencils only, and make complete erasures; the tests are computer graded, and any stray marks may count as incorrect answers.

Choose the best answer, and fill in the corresponding space on the answer sheet.

1. After Muzafer Sherif and his associates had arbitrarily divided boys in a summer camp into two groups, the bulldogs and the red devils, they found that intergroup hostility could be created by providing _____ and reduced by providing _____.

   1. inequitable rewards, equitable rewards
   2. unilateral rewards, bilateral rewards
   3. competitive activities, cooperative activities
   4. negative stereotypes, positive stereotypes

2. According to our text (Aronson) which of the following methods is most effective in reducing aggressive behavior in children?

   1. pure reason
   2. punishment of an aggressive model
   3. punishment of the aggressor
   4. rewarding alternative behavior patterns

3. In an encounter group, participants typically learn by

   1. experiencing and doing things for themselves.
   2. listening to the advice of the leader.
   3. watching how other group members characteristically behave.
   4. doing prerequisite reading before their actual participation.

4. According to Aronson, very often a person who we are very close to may have _____ reward value for us than would a stranger.

   1. greater
   2. less
   3. far greater
   4. the same

5. In Goldberg's study in which subjects read an article supposedly written by either John T. McKay (a man) or Joan T. McKay (a woman) it was found that

   1. women rated the article as better written when the author was a woman than when he was a man.
   2. the quality of the article was rated equal in both conditions but the women were more persuaded by it when the author was a woman.
   3. changing one letter in the author's first name from an "h" to an "a" had no effect on the subjects rating of the article.
   4. women rated the article as better written when the author was a man than when she was a woman.

6. An individual with an authoritarian personality structure would be most likely to endorse which one of the following statements:

   1. things are rarely either black or white, good or bad.
   2. a leader must regard the opinions of his followers.
   3. children should learn proper respect for their teachers.
   4. people vary as to attitudes and opinions that they hold.

7. The general reward theory of attraction is based upon which of the following principles?

   1. we tend to like a person better whose behavior

provides us with maximum reward at minimum cost.
2. the more a person rewards us, the more we will like him.
3. we tend to like a person better who rewards us only when we deserve the reward.
4. the more often we see another person (repeated exposure), the more we will like him.

8. The "Scapegoat" Theory of prejudice relates to the phenomenon of:

1. displaced aggression
2. socioeconomic problems in a deprived culture
3. authoritarian personality structure
4. conformity and compliance

9. As Edward Jones says, "flattery will get you somewhere," except when

1. you are too generous with your praise
2. you have someone else convey the praise for you
3. the person you are flattering likes you very much
4. the person you are flattering thinks you are trying to manipulate him.

10. According to a study by Dunnette, compared to members of a discussion group, members of T-groups showed _____ after the group experience.

1. a greater increase in rejection of traditional sex roles
2. a greater increase in empathy
3. a decrease in compliance to the instructions of leaders
4. a greater increase in conformity

*11. Which of the following is most closely related to simple prejudice?

1. interpersonal veriticality
2. anaclitic identification
3. interpersonal contrients
4. pseudo-maxi identification

12. Albert Bandura and his associates demonstrated that seeing another person behave aggressively can:

---

*Denotes "pure guess" items.

1. reduce aggressive behavior in children
2. increase aggressive behavior in young children
3. create frustration and aggression
4. reduce cognitive dissonance

13. If college profs were sneaky and manipulative, how might they get you to like their courses more?

    1. give you low grades at first and then give you increasingly higher grades as the semester progresses.
    2. demand that you undergo a lot of effortful bureaucratic hassle in order to register for the course.
    3. disagree with many of your ideas about the course material, but express a lot of liking for you as a person.
    4. all of the above are true.

14. Dion and Bersheid have found that one is aware of the physical attractiveness of those around them as early as:

    1. the beginning of junior high
    2. elementary school years
    3. nursery school
    4. two years of age

15. In the experiment by Barker, Dembo and Lewin in which children were delayed or not delayed from playing with attractive toys, it was found that the greater aggression shown by children in the delayed condition probably occurred because they

    1. were angry at other children who played with the toys
    2. were suffering from a deprivation of toys at home
    3. developed cognitive consonance by deciding that they didn't want to play with the toys after all
    4. didn't get what they were led to expect

16. Research indicates that when people do favors for us,

    1. we only like them better if we don't feel that uncomfortable obligations have been placed upon us
    2. we only like them better if the favor was anticipated
    3. we only like them better if our self-esteem is low
    4. we always like them better than when they don't

17. Which of the following is not a valid goal of the T-group as defined in the book?

    1. to get rid of your personal hang-ups (therapy)
    2. to get insight into your own behavior
    3. to get insight into the behavior of other people
    4. to get valid feedback from others about your effect on them.

18. The parent who uses severe punishment as a means of modifying undesirable behavior tends to produce children

    1. who develop great hatreds for authority figures in later life
    2. who become shy and apathetic
    3. who are extremely aggressive when not under parental supervision
    4. who learn the proper way to behave but who generally suffer from high levels of anxiety as well

19. In Aronson's discussion of encounter groups, he states that feedback expressed in terms of _____ is a lot easier for the recipient to listen and deal with than feedback in the form of _____.

    1. facts; opinions
    2. feelings; judgements, evaluations
    3. opinions; facts
    4. judgements and evaluations; feelings

20. When discussing the relative importance of (a) physical attractiveness and (b) attitude similarity it appears that

    1. physical attractiveness is generally the more important of the two
    2. attitude similarity is generally the more important of the two
    3. women were more influenced by physical attractiveness than were men
    4. the relative importance of each depends upon the type of relationship (friend, date, or marriage partner) that is envisioned

21. Which of the following is probably <u>not</u> a cause of prejudice?

    1. the displacement of aggression (scapegoating)

2. the genetic inferiority of certain racial groups
3. economic or political competition
4. conformity to social norms and conventions

*22. Studies of encounter groups have shown that people are generally most "tectonistic" during which period of their lives

1. 15-24
2. 25-34
3. 35-44
4. 45-54

23. In general, getting someone to do a favor is _____ effective in increasing his liking for you than your doing a favor for him, due to _____.

1. less; your dissonance reduction
2. less; his negative reinforcement
3. more; his dissonance reduction
4. more; your positive reinforcement

24. Freud, and the psychoanalytic school of psychology, relate the presence of aggression in all humans to

1. instinct
2. abnormal brain wave patterns
3. frustration
4. learned behavior patterns

25. The greater the empathy a person has for others

1. there is no relationship between empathy and aggression (2, 3, and 4 are false)
2. the less he resorts to aggressive actions
3. the more frequently he reports experiencing aggressive dreams
4. the more he resorts to aggressive actions

26. When a near-perfect person commits a blunder, his attractiveness generally _____. When a mediocre person commits the same blunder, his attractiveness generally _____.

1. stays the same, decreases
2. decreases, increases
3. increases, decreases
4. decreases, decreases

---

*Denotes "pure guess" items.

27. A study by Diamond and Shapiro showed that following a T-group experience subjects were more hypnotizable than matched controls. They interpreted this result as evidence of:

    1. increased trust
    2. greater fear of rejection
    3. increased gullibility
    4. lower self-esteem

28. Typically we say that in the T-group we learn from people who are not experts. Yet people are experts in two important areas crucial to the group process. These are:

    1. the interpretation and evaluation of other people's behavior
    2. group dynamics and deindividuation
    3. interpersonal communication and empathy
    4. their own feelings about themselves and about the behavior of others

29. Which of the following statements is true?

    1. hurting another person subsequently decreased the aggressor's negative feelings towards the victim
    2. hurting another person subsequently increases the aggressor's negative feelings towards the victim
    3. there is clear evidence to suggest that competitive games provide a really good outlet for the instinctive aggressive drive
    4. social psychologists have shown that man's innate patterns of behavior are rigidly fixed and determined (inflexible)

30. Caucasians, Black and Mexican-American gangs are having "wars" in a local park. Based on your knowledge of social psychological research, which of the following would you recommend as a way to reduce these group conflicts?

    1. increase the number of police and plain clothesmen patrolling the park
    2. begin a fund-raising project in which all must participate if the park is to be kept open
    3. get the leaders together to talk over their disputes
    4. assemble all the gang members in a group and give a dynamic, multi-media, emotion producing presentation on brotherhood

31. One of the reasons for the tremendous popularity of the phenomenon of grouping (encounter groups, T-groups, sensitivity groups, etc.) in recent years is that:

    1. the groups provide a possible way out from the anonymity and lack of personal identity so common in modern life
    2. it is inexpensive, simple, very safe form of psychotherapy
    3. the incidence of neuroses and psychoses is increasing in this country and one to one psychotherapy doesn't work well
    4. people are by and large frustrated sexually and grouping provides a socially approved outlet

32. Regarding the effects of observing violence on TV, there appears to be

    1. strong evidence that watching violent behavior on TV will have a therapeutic and cathartic effect on children and actually decrease their aggressive behavior
    2. little evidence that watching violent behavior on TV will increase the aggressive behavior of children
    3. the evidence is far too contradictory and tentative for any conclusions to be drawn
    4. there is some evidence that watching violent behavior on TV will increase the aggressive behavior of children

*33. Using principles derived from behavior modification research, C. Klipspringer has proposed a detailed plan to end international confrontation which has become known as

    1. IHTAR (International Harmony Through Arms Reduction)
    2. CRTID (Conflict Reduction Through International Dialogue)
    3. WRIT (Worldwide Reduction in International Tension)
    4. PTIT (Peace Through International Interdependence)

34. The single most powerful determinant of whether Jack, who does not hold himself in very high esteem, will like Jill would be

---

*Denotes "pure guess" items

1. whether Jack was in immediate need of a date for Friday night.
2. whether Jill held attitudes and values that were similar to Jack's.
3. whether Jill has complimentary personality needs from Jack.
4. whether Jack thought that Jill liked him.

35. The view that aggression is "an essential part of the life-preserving organization of "instincts" would express best the opinion of

1. Lorenz (ethological approach)
2. Dollard and Miller (frustration hypothesis)
3. Rogers (humanistic approach)
4. Bandura (social-learning approach)

36. The one determinant of whether a couple, on a blind date liked each other and repeated their date, was shown by Elaine Waister to be

1. their attitude similarity
2. their physical attractiveness
3. their intelligence
4. their complimentary personality needs

37. Most T-groups place a major emphasis on

1. sex role flexibility
2. sexual permissiveness
3. the here and now
4. interpersonal intellectualization

38. According to Aronson "opposites do attract - sometimes," which of the following pairs of opposite personality traits would be <u>most unlikely</u> to be attracted towards each other

1. nuturance - dependency
2. assertiveness - passivity
3. introversion - extroversion
4. masculinity - feminity

39. Which of the following is <u>not</u> a true statement regarding encounter groups?

1. scientific research into the effectiveness of groups has been successful in its attempt to control and manipulate the variables that produce the outcomes

2. no one should ever be forced to do something he doesn't want to do in a group meeting
3. in encounter groups, the participants are oriented toward making themselves vulnerable, because the norm is opened
4. maximal learning occurs when the group has learned to do its own work and leader intervention can thus be minimized

40. A good group leader will

1. serve in the capacity of "group therapist"
2. be a Ph.D. in clinical psychology with at least one year's practical experience in groups both as a member and as a leader
3. offer detail interpretations of, and explanations for, the groups behavior
4. be first and foremost a member of the group

*41. The guebracho phenomenon was first observed by researchers investigating

1. encounter groups
2. interpersonal attraction
3. prejudice
4. aggression

42. People that are rather rigid in their beliefs, who possess conventional values, are intolerant of weakness, and tend to be highly punitive, would be classified by spcyhologists as examples of

1. the empathetic personality
2. the neurotic personality
3. the authoritarian personality
4. the autocratic aggressive personality

43. The set of beliefs that we accept implicitly but of which we are unaware, because we cannot even conceive of alternative conceptions of the world, is known as

1. unitary thinking
2. interpersonal cognitive fixedness
3. unconscious personal judgements
4. non-conscious ideology

44. If two people work equally hard on the same task, and

---

*Denotes "pure guess" items.

by a flip of a coin, one receives a sizeable reward and the other receives nothing, observers show a strong tendency to rate the unlucky person

1. as having worked harder (law of unjust recompense)
2. as having worked less hard (just world phenomenon)
3. as having worked equally hard as the rewarded person
4. as just being plain unlucky (cruel world phenomenon)

45. Studies have indicated that by age _____ children have already developed clearly defined notions of what constitutes appropriate behavior for women and men.

1. 7
2. 6
3. 5
4. 4

46. A recent (1971) poll conducted by the National Opinion Research Center indicated that most anti-black prejudice was found

1. in groups just above the blacks socioeconomically
2. in the uppermost strata of the rich aristocracy
3. in other minorities with even lower socio-economic status than the blacks
4. in the upper middle class segment of Americans

47. Which of the following is <u>not</u> a true statement?

1. one can determine causality from correlational evidence
2. stateways can change folkways
3. frustration leads to a specific type of aggression--aggression against someone you hate
4. immediate desegregation was far more effective than gradual desegregation in the border states during the 1950's

48. If you were a public-housing administrator interested in decreasing racial antagonism, which of the following methods would you consider most effective in assigning apartments to black and white tenants?

1. Assign adjacent but separate buildings to black and white families

2. Assign apartments without regard to race so that black and white families live side by side.
3. Keep white and black families in separate clusters of buildings
4. Assign apartments so that each floor of the building is alternated with black and white families

49. Discriminatory hiring practices, restrictive zoning codes, and segregated schools are all examples of

1. personal stereotypes
2. out-group prejudice
3. institutional prejudice
4. simple prejudice

*50. According to Aronson's text, what percent of the people who have been through 1 or more encounter group sessions later report having ordurious experiences outside of the group?

1. 40 percent
2. 25 percent
3. 15 percent
4. 5 percent

---

*Denotes "pure guess" items.

APPENDIX C

TABLE 5

CHI-SQUARE VALUES FOR THE NINE "PURE GUESS" ITEMS IN EXPERIMENT 1--A COMPARISON OF ALTERNATIVE ATTRACTIVENESS

Item	Alternative A	B	C	D	χ^2	df	sig
#1							
Obtained Frequency*	17	20	33	18	7.54	3	NS
Expected Frequency	22	22	22	22			
#2							
Obtained Frequency	25	11	28	24	7.72	3	NS
Expected Frequency	22	22	22	22			
#3							
Obtained Frequency	33	15	11	29	15.45	3	P<.05
Expected Frequdncy	22	22	22	22			
#4							
Obtained Frequency	23	24	20	21	0.45	3	NS
Expected Frequency	22	22	22	22			
#5							
Obtained Frequency	23	16	24	25	2.27	3	NS
Expected Frequency	22	22	22	22			
#6							
Obtained Frequency	22	21	24	21	0.27	3	NS
Expected Frequency	22	22	22	22			
#7							
Obtained Frequency	11	32	27	18	11.90	3	P<.05
Expected Frequency	22	22	22	22			
#8							
Obtained Frequency	24	20	16	28	3.63	3	NS
Expected Frequency	22	22	22	22			

TABLE 5 (continued)

Item	Alternative A	B	C	D	χ^2	df	sig
#9							
Obtained Frequency	16	23	32	17	7.36	3	NS
Expected Frequency	22	22	22	22			

*The Obtained Frequency values correspond to the total frequencies for each Alternative summed across the four forms.

APPENDIX D

TABLE 6

CHI-SQUARE VALUES FOR THE FIVE "PURE GUESS" ITEMS IN EXPERIMENT 2--A COMPARISON OF ALTERNATIVE ATTRACTIVENESS

Item	Alternative A	B	C	D	χ^2	df	sig
#1							
Obtained Frequency*	33	47	33	22	9.26	3	P<.05
Expected Frequency	34	34	34	34			
#2							
Obtained Frequency	47	46	26	16	19.47	3	P<.05
Expected Frequency	34	34	34	34			
#3							
Obtained Frequency	14	15	40	67	55.47	3	P<.05
Expected Frequency	34	34	34	34			
#4							
Obtained Frequency	40	36	36	24	4.23	3	NS
Expected Frequency	34	34	34	34			
#5							
Obtained Frquency	44	27	41	24	8.47	3	P<.05
Expected Frequency	34	34	34	34			

*The Obtained Frequency values correspond to the frequencies for each Alternative summed across the four forms.

APPENDIX E

TABLE 7

CHI-SQUARE VALUES FOR THE NINE "PURE GUESS" ITEMS IN EXPERIMENT 1--AN ALTERNATIVE BY ALTERNATIVE ANALYSIS

Item	Response Position 1	2	3	4	χ^2	df	sig
#1							
Alternative A							
Obtained Frequency	5	5	3	4	0.6	3	NS
Expected Frequency	4.25	4.25	4.25	4.25			
Alternative B							
Obtained Frequency	7	5	6	2	2.8	3	NS
Expected Frequency	5	5	5	5			
Alternative C							
Obtained Frequency	10	5	10	8	2.0	3	NS
Expected Frequency	8.25	8.25	8.25	8.25			
Alternative D							
Obtained Frequency	6	3	7	2	3.8	3	NS
Expected Frequency	4.50	4.50	4.50	4.50			
#2							
Alternative A							
Obtained Frequency	3	11	9	2	9.4	3	P<.05
Expected Frequency	6.25	6.25	6.25	6.25			
Alternative B							
Obtained Frequency	5	2	1	3	3.2	3	NS
Expected Frequency	2.75	2.75	2.75	2.75			
Alternative C							
Obtained Frequency	4	12	8	4	6.3	3	NS
Expected Frequency	7	7	7	7			
Alternative D							
Obtained Frequency	7	4	4	9	3.0	3	NS
Expected Frequency	6	6	6	6			

TABLE 7 (continued)

Item	Response Position 1	2	3	4	χ^2	df	sig
#3							
Alternative A							
Obtained Frequency	8	11	8	6	1.5	3	NS
Expected Frequency	8.25	8.25	8.25	8.25			
Alternative B							
Obtained Frequency	1	3	10	1	14.6	3	P<.05
Expected Frequency	3.75	3.75	3.75	3.75			
Alternative C							
Obtained Frequency	1	1	3	6	6.1	3	NS
Expected Frequency	2.75	2.75	2.75	2.75			
Alternative D							
Obtained Frequency	5	7	9	8	1.2	3	NS
Expected Frequency	7.25	7.25	7.25	7.25			
#4							
Alternative A							
Obtained Frequency	2	11	9	1	13.0	3	P<.05
Expected Frequency	5.75	5.75	5.75	5.75			
Alternative B							
Obtained Frequency	4	6	6	8	1.3	3	NS
Expected Frequency	6	6	6	6			
Alternative C							
Obtained Frequency	2	6	5	7	2.8	3	NS
Expected Frequency	5	5	5	5			
Alternative D							
Obtained Frequency	9	2	1	9	10.8	3	P<.05
Expected Frequency	5.25	5.25	5.25	5.25			
#5							
Alternative A							
Obtained Frequency	7	7	3	6	3.1	3	NS
Expected Frequency	5.75	5.75	5.75	5.75			
Alternative B							
Obtained Frequency	6	6	3	1	4.5	3	NS
Expected Frequency	4	4	4	4			

TABLE 7 (continued)

Item	Response Position 1	2	3	4	χ^2	df	sig
Alternative C							
Obtained Frequency	8	4	6	6	1.3	3	NS
Expected Frequency	6	6	6	6			
Alternative D							
Obtained Frequency	9	7	6	3	3.0	3	NS
Expected Frequency	6.25	6.25	6.25	6.25			
#6							
Alternative A							
Obtained Frequency	5	10	5	2	6.0	3	NS
Expected Frequency	5.50	5.50	5.50	5.50			
Alternative B							
Obtained Frequency	4	6	7	4	1.3	3	NS
Expected Frequency	5.25	5.25	5.25	5.25			
Alternative C							
Obtained Frequency	7	6	2	9	4.3	3	NS
Expected Frequency	6	6	6	6			
Alternative D							
Obtained Frequency	7	4	1	9	7.0	3	NS
Expected Frequency	5.25	5.25	5.25	5.25			
#7							
Alternative A							
Obtained Frequency	6	0	2	3	6.8	3	NS
Expected Frequency	2.75	2.75	2.75	2.75			
Alternative B							
Obtained Frequency	3	11	15	3	13.5	3	P<.05
Expected Frequency	8	8	8	8			
Alternative C							
Obtained Frequency	11	2	4	10	8.7	3	P<.05
Expected Frequency	6.75	6.75	6.75	6.75			
Alternative D							
Obtained Frequency	2	7	8	1	8.2	3	P<.05
Expected Frequency	4.50	4.50	4.50	4.50			

TABLE 7 (continued)

Item	Response Position 1	2	3	4	χ^2	df	sig
#8							
Alternative A							
Obtained Frequency	10	1	0	13	21.0	3	P<.05
Expected Frequency	6	6	6	6			
Alternative B							
Obtained Frequency	5	6	7	2	2.8	3	NS
Expected Frequency							
Alternative C							
Obtained Frequency	9	2	2	3	8.5	3	P<.05
Expected Frequency	4	4	4	4			
Alternative D							
Obtained Frequency	0	14	10	4	16.6	3	P<.05
Expected Frequency	7	7	7	7			
#9							
Alternative A							
Obtained Frequency	6	1	7	2	6.5	3	NS
Expected Frequency	4	4	4	4			
Alternative B							
Obtained Frequency	5	9	6	3	3.3	3	NS
Expected Frequency	5.75	5.75	5.75	5.75			
Alternative C							
Obtained Frequency	12	11	4	5	6.3	3	NS
Expected Frequency	8	8	8	8			
Alternative D							
Obtained Frequency	3	5	6	3	1.6	3	NS
Expected Frequency	4.25	4.25	4.25	4.25			

APPENDIX F

TABLE 8

CHI-SQUARE VALUES FOR THE FIVE "PURE GUESS" ITEMS IN EXPERIMENT 2--AN ALTERNATIVE BY ALTERNATIVE ANALYSIS

Item	Response Position 1	2	3	4	χ^2	df	sig
#1							
Alternative A							
Obtained Frequency	9	7	8	9	0.3	3	NS
Expected Frequency	8.25	8.25	8.25	8.25			
Alternative B							
Obtained Frequency	15	10	10	12	1.5	3	NS
Expected Frequency	11.75	11.75	11.75	11.75			
Alternative C							
Obtained Frequency	9	7	11	6	1.8	3	NS
Expected Frequency	8.25	8.25	8.25	8.25			
Alternative D							
Obtained Frequency	10	3	4	5	5.3	3	NS
Expected Frequency	5.50	5.50	5.50	5.50			
#2							
Alternative A							
Obtained Frequency	13	12	13	9	0.9	3	NS
Expected Frequency	11.75	11.75	11.75	11.75			
Alternative B							
Obtained Frequency	12	11	12	11	0.1	3	NS
Expected Frequency	11.50	11.50	11.50	11.50			
Alternative C							
Obtained Frequency	7	8	6	5	0.8	3	NS
Expected Frequency	6.50	6.50	6.50	6.50			
Alternative D							
Obtained Frequency	2	5	6	3	2.5	3	NS
Expected Frequency	4	4	4	4			

TABLE 8 (continued)

Item	Response Position 1	2	3	4	χ^2	df	sig
#3							
Alternative A							
Obtained Frequency	6	2	6	0	7.8	3	NS
Expected Frequency	3.50	3.50	3.50	3.50			
Alternative B							
Obtained Frequency	5	5	5	0	5.0	3	NS
Expected Frequency	3.75	3.75	3.75	3.75			
Alternative C							
Obtained Frequency	11	6	10	8	1.7	3	NS
Expected Frequency	8.75	8.75	8.75	8.75			
Alternative D							
Obtained Frequency	18	15	21	12	2.2	3	NS
Expected Frequency	16.50	16.50	16.50	16.50			
#4							
Alternative A							
Obtained Frequency	11	10	10	9	0.2	3	NS
Expected Frequency	10	10	10	10			
Alternative B							
Obtained Frequency	9	12	2	13	8.2	3	P<.05
Expected Frequency	9	9	9	9			
Alternative C							
Obtained Frequency	7	9	11	9	0.9	3	NS
Expected Frequency							
Alternative D							
Obtained Frequency	11	3	5	3	7.8	3	P<.05
Expected Frequency	5.50	5.50	5.50	5.50			
#5							
Alternative A							
Obtained Frequency	11	13	11	9	0.7	3	NS
Expected Frequency	11	11	11	11			
Alternative B							
Obtained Frequency	7	9	4	7	1.9	3	NS
Expected Frequency	6.75	6.75	6.75	6.75			

TABLE 8 (continued)

Item	Response Position 1	2	3	4	χ^2	df	sig
Alternative C							
Obtained Frequency	9	13	10	9	0.8	3	NS
Expected Frequency	10.25	10.25	10.25	10.25			
Alternative D							
Obtained Frequency	9	7	5	3	3.3	3	NS
Expected Frequency	6	6	6	6			

VITA

James A. Eison was born in Brooklyn, New York on September 13, 1950. He attended elementary school in Flushing, New York and was graduated from John Bowne High School in 1968. The following year he entered State University of New York, College at New Paltz and in June 1972 he received the Bachelor of Arts degree, cum laude, in Psychology.

He entered The Graduate School at The University of Tennessee in September 1972 and received the Master of Arts degree with a major in Psychology in December 1975.

In 1974 he co-authored the book General Psychology: A Course Manual and Workbook (Collegiate Press) with Dr. Raymond Shrader, David Murray, and Robert Fisher.