THE APPLICATION OF THE TOOPS-ADKINS FORMULA IN THE CONSTRUCTION AND VALIDATION OF A TEST IN MATHEMATICS

by

AUBREY THORNTON EDWARDS

B. S., Kansas State College
of Agriculture and Applied Science, 1941

A THESIS

submitted in partial fulfillment of the

requirements for the degree of

MASTER OF SCIENCE

Department of Education and Psychology

KANSAS STATE COLLEGE
OF AGRICULTURE AND APPLIED SCIENCE

1946

TABLE OF CONTENTS

INTRODUCTION

The need for a carefully constructed diagnostic test in mathematics which required an understanding of basic principles for later study of physical science was apparent in Kansas high schools and colleges. The construction, with item analysis, of such a test was the problem for this thesis.

Various methods of item analysis were studied, and the one which Herbert A. Toops and Dorothy C. Adkins developed as a modification of the Pearson product moment correlation formula was chosen for use in constructing the test. This method determines the correlation of each item with the total test by finding the indices of validity of the various test items on the basis of internal consistency.

The test was so constructed that the answers could be indicated on a separate answer sheet such as the Perfo-Score Answer Sheet so that the test might be used again and again by purchasing only answer sheets.

The refined test itself is presented as the conclusion of this problem.

REVIEW OF LITERATURE

For the past fifteen years many methods for determining the validity of test items have been used and compared. Validity has been defined as the closeness of agreement of the scores of a test and some other objective measure of that which the test is used to measure. Bliss (2) defined the validity of an item in a test as "the extent to which the item differentiates between individuals of superior and inferior ability."

The method most commonly found in studies of item analysis is the Biserial r. It is recognized, in general, as being one of the best.

$$\text{Biserial } r = \frac{(M_1 - M_2)\ pq}{\sigma\ z}$$

M_1 = mean score of the pupils having the item right.

M_2 = mean score of the pupils having the item wrong.

p = percentage of pupils having the item right.

q = percentage of pupils having the item wrong.

σ = S. D. of all criterion scores.

z = ordinate of the normal probability curve cutting off p cases.

Clark's formula (3) for evaluating test items is as follows:

$$V = \frac{P - D}{1 - D}$$

V = validity of the item.

P = percentage of pupils who fail to answer the item correctly, it is the difficulty of the item.

D = percentage of the criterion group who fail to answer the items correctly. The criterion group is the D percentage of the class having the lowest scores.

The Vincent validity method (5), sometimes called the overlapping method, evaluates an item by measuring the extent to which the poor pupils overlap upon the good pupils on that item. Overlapping is measured by calculating the percentage of pupils failing the item who have criterion scores higher than the median score of the pupils passing the item. The following steps are followed in calculating the Vincent Overlapping score:

1. Arrange the pupils in the order of scores, beginning with the highest.

2. Find the total number of pupils having the item right.

3. Locate the median passing score, that is, the score of the middle pupil of those passing the item.

4. Count the number of failures above the median passing score.

5. Express the failures above the median passing score as a percentage of the total failures.

6. In ranking items by this method, that item ranks highest which shows the smallest percentage of overlapping.

In the Upper vs. Lower Third Method (5), the scores are ranked in the order of excellence and divided into upper, middle, and lower thirds. The evaluation of each item is the difference between the number of passes for the upper third, and the number of passes for the lower third. Those items with the greatest number of passes in the upper group in excess of the lower are considered the best items since they discriminate most between the two groups.

McCall's method (5), which resembles Biserial r, may be expressed as follows:

$$C = \frac{(M_1 - M_2)\ N_1 \times N_2}{N}$$

C = the coefficient of value for the item.

M_1 = the mean total score of those persons making a positive (correct) reaction to the item.

M_2 = the mean total of those persons making a negative (incorrect) reaction to the item.

N_1 = the number of persons reacting positively to the item.

N_2 = the number of persons reacting negatively to the item.

N = the total number of responses to the item.

The Summation of Agreements method (5) proposed by Lentz may be expressed:

$$C = \frac{N_{a_1} + N_{a_2} + N_{a_3} + \ldots N_{a_n}}{N}$$

C = the coefficient.

N_{a_1} = the total number of responses by the first subject which agree with that subjects response to the item being considered.

N_{a_2} = is the total number of responses by the second subject which agree with that subjects response to the item being considered.

N = the total number of persons responding to the item.

Holzinger's Method (6) makes use of this formula:

$$V = \frac{(Ru + Wl) - (Wu + Rl)}{1/2\ N}$$

R_u = number of "rights" in upper twenty-five percent of total group.

W_u = number of "wrongs" in upper twenty-five percent of total group.

R_l = number of "rights" in lower twenty-five percent of total group.

W_l = number of "wrongs" in lower twenty-five percent of total group.

N = total number of persons.

A very simple formula, Difference between Means, suggested by Swineford (6) is $V = M_R - M_W$

M_R = mean score of those persons making a positive answer to the item.

M_W = mean score of those persons making a negative answer to the item.

Other methods mentioned in the literature (5, 6) but generally disregarded as being impractical are Zubins index, Critical Ratio, Boos Method, McCall, Long, Bliss Method, and Long's Weighted Overlapping.

Swineford (6) recommended the Holzinger method and the Difference between Means as the best, because of simplicity of calculation, but warned that Holzinger's method does not consider the middle 50 percent of the subjects, and, further, that it penalizes items whose difficulty lies outside the range of 25 to 75 percent.

Lentz, Hershtein, and Finch (5) favored the Upper vs.

Lower Third Method, but stated that it also disregarded a middle one-third of the subjects.

Adkins and Toops (1) developed a simplified formula based on the Pearson product moment formula for correlation which overcomes many of the limitations mentioned in criticism of the other methods. They took the formula for the Pearson correlation coefficient of a dichotomous variable with a multiple categoried variable and simplified it for computational purposes by effecting in the multiple categoried variable an arbitrary rectangular distribution.

Toops and Adkins make the following statement:

> By using arbitrary forms of distribution of the criterion variable--a process which has but slight relative effect on the criterion coefficient of competing items--we have derived several very simple formulas for the correlation of a dichotomously-scored item and a multiple-categoried criterion. In our opinion, corrections for coarse grouping will be unnecessary. The size of N may be set arbitrarily by two considerations: (a) N must be an integral multiple of K (the number of categories) and (b) it must be large enough to render sampling errors negligible.

When the rectangular criterion distribution is divided into five equal categories, the simplified formula is:

$$r_{xy} = \frac{\Sigma Y_R}{\sqrt{2RW}}$$

x = an item scored dichotomously and administered with no time limit.

y = the criterion score.

R = the number of persons answering an item correctly.

W = the number of persons answering an item incorrectly.

ΣY_R = the sum of the criterion scores of the persons answering an item correctly.

The five equal categories are obtained by dividing the ranked scores of the individuals tested into five equal groups in the order of excellence.

If the five equal categories are given a coded value of 2, 1, 0, -1, -2, in the descending order, and a, b, c, d, and e represent the number of right answers in each respective group, then ΣY_R becomes 2a + b - d - 2e. The formula then, in a usable form, becomes:

$$r_{xy} = \frac{2a + b - d - 2e}{\sqrt{2RW}}$$

This formula was chosen as superior to the others and was used in the selection of the items for the construction of the test presented at the conclusion of this thesis.

DESCRIPTION OF PROCEDURE

Construction

A test of 130 items embodying the basic mathematical concepts of arithmetic, algebra, and concrete measures needed for an understanding of beginning courses in chemistry and physics was developed. Each question was stated and five possible answers were presented. The answers were lettered a, b, c, d, and e respectively. In most cases one answer was correct. To vary the presentation and to help prevent guessing, occasional items had no correct answer given, and the phrase "no correct answer" was one of the alternatives. This phrase was sometimes included in an item where the correct answer also appeared. It was hoped that this would tend to make the testee actually solve the example and not guess.

The test was administered to approximately 100 college freshmen, 100 Salina High School seniors, and 100 Junction City High School seniors. No time limit was established. The test was so long that those who took it were fatigued and many did not finish the 130 items. A total of 250 students finished the first 75 items, so from these, the 50 items with the highest validity coefficients were chosen for the refined test.

The answers were marked on the Perfo-Score answer sheets and were scored and the subjects ranked in the order of excellence. In this ranking each student's response to each of the 75 items was recorded. A response to any of the choices except the right one was scored as wrong, thus fulfilling the requirement of the formula for a dichotomously scored item. The ranks were then divided to five equal categories of 50 students in each, thus fulfilling the requirement of the formula for a multiple categoried criterion. The number of right answers to each item was recorded for every group. The total number of right answers and the total number of wrong answers for each item were determined from the data and the Toops-Adkins formula was applied.

Validity

The coefficient of validity for each item was found. In using the whole test as a criterion with which to correlate each item, it may be assumed that the test is valid. Great care was taken in the choice of items. Many texts, both college and high school, were studied, and items commonly basic in all were used in the test. The opinions of college and high school mathematics and science teachers were sought and found to agree that it was a test of basic mathematical concepts. Therefore, it seemed sound to use the Toops and Adkins method of item selection which determines the degree of consistency with which each test item measures whatever the whole test measures.

It was impossible to determine a coefficient of validity for the test by correlating the scores with grades in mathematics, because the students, especially the boys, who would probably have taken mathematics courses, were not in school but were in the army. The students tested, who were not drafted, were so widely scattered in different schools that later grades in mathematics could not be obtained. It would have been highly desirable to know the validity of the test found by comparing it with an outside criterion, but circumstances beyond the author's control prevented such a calculation.

Toops and Adkins recommended as ideal that the number of cases be 1000. This ideal is seldom attained in actual practice. In this study 250 were used. Since the formula uses the square root of the number of cases in the denominator, the $\sqrt{1000}$ is only about twice as large as the $\sqrt{250}$.

Table 1 gives the (r) coefficient of validity for each of the 75 items. The starred items are those chosen for the refined test. Four items showed coefficients of validity of .60 or better; sixteen, .50 to .60; twenty-three, .40 to .50; and seven, .35 to .40. The twenty-five items with validity coefficients below .35 were cast out. Some of the rejected items gave indices of validity lower than -.01.

These indices of validity may be too low by a small amount. Kelly (4) states that when a correlation is between ranks on one hand and variates on the other, the $r = 1.0233\rho$. The values for r in the tables are not corrected. Attention is called to this fact because an arbitrary rectangular distribution was set up which makes the r a little lower than it would be if the distribution were left in a form approaching normality.

Order of Difficulty

The 50 items chosen for the refined test were arranged in the order of difficulty according the number of right responses on each item. Table 2 gives the item number on the refined test, the item number on the original test, the number of right responses, and the coefficients of validity. It can be observed from the table that the easiest item was answered correctly by 216 of the 250 students taking the test, while the most difficult item was answered correctly by only 55. Question number 28 was answered correctly by 125, exactly one-half of the students. Henry (3) states "Apart from extreme items--those on which nearly all the pupils pass, or those on which nearly all fail--the difficulty of an item has little to do with its validity." His statement is upheld in this study, for in the 25 least difficult items, 13 are above .45, and in the 25 most difficult items, 15 are above .45.

Reliability

The reliability of a measure is its self consistency. A reliability coefficient was calculated by correlating the odd numbered items with the even numbered items on the original 75 item test as administered to 100 high school students. The Pearson product moment formula was used. The reliability coefficient was .875. This .875 was the reliability coefficient for one-half of the test. When the Spearman-Brown formula was applied, the reliability coefficient for the whole original test of 75 items was found to be .933.

The same formula was used to determine the reliability coefficient for a test of 50 items. This index was .904. This index might vary if computed directly from test and retest on the 50 item refinement for the order of the items was changed. It was impossible to obtain a figure on the 50 item test because the refined test was not administered to any large number of students. These figures are given merely as an indication that the reliability of the test is high.

Comparison of Item Validities

Coefficients of validity were calculated for 130 test items on one hundred high school students. In a separate calculation the indices of validity for 75 items administered to 90 college students were found. This was done to observe the effect on the coefficients of validity in using numbers of individuals much smaller than 100, and in using a greater number of items than 75. Table 3 shows a comparison of the validities of the 50 chosen items under these other circumstances. It was interesting to observe that in nearly every case the validity of the items remained high under the different conditions of small numbers of students and greater number of items.

The 100 students in the calculations using 130 items were not identical with those used in the 250 student and 75 item calculation. The 90 college students were not all the same individuals used in the college part of the 250 student study. Since it was impossible to obtain data for establishing a coefficient of validity with an outside criterion, the fact that the internal validities remained high under these varying circumstances was taken as an indication that the whole might be valid.

Key For Table 1

a - the number of right answers in the first group (highest fifth).

b - the number of right answers in the second group.

c - the number of right answers in the third group.

d - the number or right answers in the fourth group.

e - the number of right answers in the fifth group.

R - the total number of right answers.

W - the total number of wrong answers.

(2a + b - d - 2e) - the numerator of the Toops-Adkins formula.

r - coefficient of validity.

Table 1. Coefficients of validity of original 75 items.

Item	a	b	c	d	e	R	W	2a + b - d - 2e	2RW	$\sqrt{2RW}$	r
1	48	48	50	48	45	239	11	6	5258	72.5	.08
2	49	47	49	48	45	236	14	11	6608	81.3	.14
3	45	44	43	39	34	205	45	27	18450	135.0	.33
4	49	48	50	50	45	242	8	6	3872	62.2	.10
5	46	44	41	32	17	180	70	70	25200	159.0	.44*
6	49	47	48	47	44	235	15	10	7050	84.0	.12
7	48	50	47	46	30	221	29	40	12818	113.0	.35
8	50	50	48	38	30	216	34	52	14688	121.0	.43*
9	48	45	49	49	41	232	18	10	8352	91.4	.11
10	50	47	45	45	32	219	31	38	13577	116.5	.33
11	42	35	34	19	14	144	106	72	30528	174.8	.41*
12	48	45	43	34	11	181	69	85	24977	159.0	.54*
13	49	48	48	45	34	224	26	33	11648	108.0	.31
14	47	41	40	35	26	189	61	48	23058	151.9	.32
15	48	48	48	42	29	215	35	44	15054	122.9	.36*
16	47	47	46	37	25	202	48	54	19392	139.0	.39*
17	49	48	45	44	43	229	21	16	9618	98.1	.16
18	49	45	45	36	27	202	48	53	19392	139.0	.38*
19	48	46	38	33	24	189	61	85	23058	151.9	.56*
20	46	46	35	25	9	161	89	95	28658	169.2	.56*
21	40	35	29	28	18	150	100	51	30000	173.0	.29
22	41	33	22	17	10	123	127	78	31242	176.8	.44*
23	41	40	35	21	15	152	98	71	29792	172.5	.41*
24	46	45	39	40	30	200	50	37	20000	141.7	.26
25	28	17	17	7	5	74	176	56	26048	161.5	.35
26	46	43	42	38	16	185	65	65	24050	155.0	.42*
27	40	34	30	29	8	141	109	69	30738	175.3	.39*
28	48	47	41	39	35	210	40	34	16800	129.8	.26
29	43	35	27	21	15	141	109	70	30738	175.3	.40*
30	44	43	44	39	33	203	47	26	19082	138.0	.19
31	28	20	15	14	15	92	158	32	29074	170.3	.19
32	38	35	16	10	10	109	141	81	30738	175.3	.46*
33	31	38	26	13	11	119	131	65	31178	176.7	.37*
34	49	37	30	19	12	147	103	92	30282	174.0	.53*
35	44	39	43	36	16	178	72	59	25634	160.0	.37*
36	49	48	32	16	14	159	91	102	28938	170.0	.60*
37	49	45	26	9	6	135	115	122	31050	176.4	.69*
38	47	38	24	19	14	142	108	85	30672	175.0	.49*
39	21	6	9	4	5	45	205	34	13940	118.0	.29
40	36	20	17	16	13	102	148	82	30192	173.7	.47*

Table 1 (concl.).

Item	a	b	c	d	e	R	W	2a + b -d -2e	2RW	$\sqrt{2RW}$	r
41	42	32	27	15	8	124	126	85	31248	176.7	.48*
42	37	31	21	16	8	113	137	73	30962	176.0	.41*
43	42	37	32	16	13	140	110	79	30800	175.5	.45*
44	50	44	34	17	14	159	91	99	28938	170.0	.58*
45	5	1	2	9	3	20	230	-1	9200	96.0	-.01
46	47	36	24	24	8	139	111	90	30858	175.7	.51*
47	2	4	5	5	4	20	230	-5	9200	96.0	-.05
48	8	10	7	11	7	43	207	1	17802	133.7	.01
49	42	35	18	17	12	124	126	78	31248	176.7	.44*
50	50	44	31	19	11	155	95	103	29450	171.7	.60*
51	44	39	16	8	4	111	139	111	30858	175.7	.63*
52	2	5	11	5	3	26	224	-2	11648	108.0	-.02
53	37	28	14	7	5	91	159	85	28938	170.0	.50*
54	41	43	23	23	7	137	113	88	30962	176.0	.50*
55	49	40	29	24	9	151	99	96	29898	172.9	.56*
56	42	30	11	7	5	95	155	97	29450	171.8	.56*
57	36	33	14	6	8	97	153	83	29682	172.3	.48*
58	47	42	34	23	19	165	85	75	28050	167.5	.45*
59	48	40	21	16	10	135	115	100	31050	176.7	.57*
60	44	35	15	15	9	118	132	90	31154	176.8	.51*
61	42	19	19	16	8	104	146	71	30368	174.4	.41*
62	46	38	21	9	11	125	125	99	31250	176.8	.56*
63	27	15	4	9	0	55	195	60	21450	146.2	.41*
64	34	20	20	12	15	101	149	46	30098	173.0	.27
65	30	14	10	10	4	68	182	56	24754	157.2	.36*
66	44	26	24	11	12	117	133	79	31122	176.5	.45*
67	33	15	14	14	5	81	169	57	27378	165.2	.35
68	45	22	12	12	3	94	156	94	29328	171.3	.55*
69	38	21	12	6	4	81	169	83	27378	165.2	.50*
70	23	14	9	15	5	66	184	35	24288	155.7	.23
71	46	30	18	20	13	127	123	76	31242	176.8	.43*
72	30	17	6	4	2	59	191	69	22538	150.0	.46*
73	38	15	7	8	6	74	176	71	26048	161.3	.44*
74	36	12	8	13	3	72	178	65	25632	160.0	.41*
75	42	25	17	20	1	105	145	87	30450	174.5	.50*

* Items chosen for refined test.

Table 2. Order of difficulty of items on refined test.

Item number on Refined test	Item number on Original test	Number of right responses	r
1	8	216	.43
2	15	215	.36
3	16	202	.39
4	18	202	.38
5	19	189	.56
6	26	185	.42
7	12	181	.54
8	5	180	.44
9	35	178	.37
10	58	165	.45
11	20	161	.56
12	36	159	.60
13	44	159	.58
14	50	155	.60
15	23	152	.41
16	55	151	.56
17	34	147	.53
18	11	144	.41
19	38	142	.49
20	27	141	.39
21	29	141	.40
22	43	140	.45
23	46	139	.51
24	54	137	.50
25	37	135	.69
26	59	135	.57
27	71	127	.43
28	62	125	.56
29	41	124	.48
30	49	124	.44
31	22	123	.44
32	33	119	.37
33	60	118	.51
34	66	117	.45
35	42	113	.41

Table 2 (concl.).

Item number on Refined test	Item number on Original test	Number of right responses	r
36	51	111	.63
37	32	109	.46
38	75	105	.50
39	61	104	.41
40	40	102	.47
41	57	97	.48
42	56	95	.56
43	68	94	.55
44	53	91	.50
45	69	81	.50
46	73	74	.44
47	74	72	.41
48	65	68	.36
49	72	59	.46
50	63	55	.41

Table 3. Comparisons of r.

Question number	r for 250 high school and college students	r for 100 high school students	r for 90 college students
1	.48	.43	.31
2	.36	.44	.22
3	.39	.45	.17
4	.38	.28	.26
5	.56	.37	.33
6	.42	.31	.47
7	.54	.54	.24
8	.44	.45	.22
9	.37	.41	.25
10	.45	.30	.43
11	.56	.53	.48
12	.60	.56	.56
13	.58	.52	.39
14	.60	.51	.67
15	.41	.38	.40
16	.56	.50	.51
17	.53	.58	.46
18	.41	.44	.29
19	.49	.38	.48
20	.39	.47	.34
21	.40	.51	.34
22	.45	.44	.36
23	.51	.50	.49
24	.50	.36	.41
25	.69	.62	.58
26	.57	.48	.56
27	.43	.44	.53
28	.56	.63	.63
29	.48	.44	.43
30	.44	.43	.43
31	.44	.42	.46
32	.37	.26	.30
33	.51	.44	.57
34	.45	.45	.53
35	.41	.34	.31

Table 3 (concl.).

Question number	r for 250 high school and college students	r for 100 high school students	r for 90 college students
36	.63	.53	.64
37	.46	.48	.33
38	.50	.58	.66
39	.41	.52	.28
40	.47	.22	.38
41	.48	.45	.55
42	.56	.55	.52
43	.55	.50	.60
44	.50	.51	.55
45	.50	.53	.54
46	.44	.37	.53
47	.41	.29	.53
48	.36	.38	.44
49	.46	.56	.43
50	.41	.30	.49

RESULTS

The result of this study is the refined test in mathematics. Each item has an index of validity of .36 or more. The fifty item refined test is arranged in the order of difficulty based on 250 cases. The coefficient of reliability for fifty items in the original order was .904.

This problem included only the construction of the test. It must be clearly pointed out that the test is not presented for immediate use. Before it can be really useful with diagnosic or predictive values, it must be standardized.

The value of establishing local norms has been stressed in recent literature. The test could be used, if those using it desired to establish their own local norms before much significance could be attached to the scores of those taking the test. A correlation with some independent criteria is necessary before the test could be used with confidence as an indicator of mathematical aptitude.

CONCLUSION

In conclusion a copy of the refined test itself is presented. The directions to the student are given below:

Directions to the student - This is a test to see how much elementary mathematics you have learned. Read each question on the test sheet and compute the right answer on the blank paper provided. Note the number of the question and the letter preceding your chosen answer. Then draw a circle around that letter following the same number on the answer sheet. Make no marks on the question sheet. You may have as much time as you need.

Mathematics Test

1. $7 \times 0 =$ (a) 7, (b) 0, (c) 70, (d) -7, (e) .7.

2. The number of yards in 111 feet is (a) 39, (b) 37, (c) 43, (d) 41, (e) 36.

3. The common denominator for 1/4, 1/5, and 1/8 is (a) 120, (b) 60, (c) 40, (d) 20, (e) 32.

4. $.39\overline{)22.776}$ = (a) 58.4, (b) .584, (c) 5.84, (d) 584, (e) .0584.

5. $1\ 1/2 \div 2\ 2/3 =$ (a) 6/16, (b) 3/8, (c) 16/9, (d) 7/16, (e) 9/16.

6. $7.40 commission for selling a $37 article is (a) 2%, (b) 20%, (c) 21%, (d) 25%, (e) .02%.

7. $\sqrt{324} =$ (a) 13, (b) 15, (c) 17, (d) 12, (e) 18.

8. 5×5 may be expressed: (a) 5/5, (b) 5+5, (c) 5_2, (d) 5^2, (e) 2(5).

9. When f = 15i, then 5f + 8f + 10f = (a) 235i, (b) 355i, (c) 245i, (d) 340i, (e) 345i.

10. $6/a \cdot 7/b$ equals (a) $\frac{6a \cdot 7b}{a+b}$, (b) $\frac{42}{a+b}$, (c) $\frac{6b+7a}{ab}$, (d) $\frac{42}{ab}$, (e) $\frac{6a - 7b}{ab}$.

11. The perimeter of a plot 25 ft. by 16 ft. is (a) 200 ft., (b) 83 sq. ft., (c) 400 ft., (d) 82 ft., (e) 83 ft.

12. $a^x \cdot a^y =$ (a) ax+ay, (b) a^x+y, (c) $2a^x$, (d) axy, (e) a^{x-y}.

13. $x^2 = 2mn$, then x = (a) $\sqrt{2mn}$, (b) $2\sqrt{mn}$, (c) $m\sqrt{2n}$, (d) $2\sqrt{m/n}$, (e) $2\sqrt{\frac{n}{m}}$.

14. Solve for x: $x/2=5y^2$, (a) $10y^2$, (b) $\frac{5y^2}{-2}$, (c) $5y^2$, (d) $-5y^2$, (e) $-10y^2$.

15. 28 is what per cent of 560? (a) 2%, (b) 50%, (c) 20%, (d) 5%, (e) 40%.

16. $\frac{12x - 24x^2}{-6x}$ equals, (a) 2+4x, (b) 2-4x, (c) x+4x, (d) x-4x, (e) -2+4x.

17. Solve this expression $(a+b)(a-b)$ when $a=.03$, and $b=.02$: (a) .005, (b) .0005, (c) .05, (d) .5, (e) 5.

18. 2 2/5+4 1/3+6 1/8=(a) 12 107/120, (b) 12 5/8, (c) 12 4/5, (d) 12 109/120, (e) 12 103/120.

19. If $x=a/b$, then $b=$(a) a/x, (b) xa, (c) b/x, (d) bx, (e) $a+x$.

20. The square root of 203,401 is (a) 431, (b) 429, (c) 451, (d) 409, (e) 439.

21. A suit on sale with 1/3 off cost $24. The original price was (a) $36, (b) $35, (c) $32, (d) $30, (e) no correct answer.

22. $\frac{x_1}{x_2}=\frac{y2}{y1}$ then $x_1y_1=$ (a) x_2/y_2, (b) x_2y_2, (c) y_2/x_2, (d) x_2y_1, (e) y_1x_2.

23. What is the reciprocal of x/y? (a) $1/x$, (b) xy, (c) y/x, (d) $-x/y$, (e) $1/y$.

24. $\frac{12x-24x^2}{6x}$ equals, (a) $2x+4x$, (b) $2-4x$, (c) $2x-4$, (d) $2+4x$, (e) $2-4$.

25. $a^x/a^y=$ (a) $ax+ay$, (b) a^{x+y}, (c) $2a^x$, (d) axy, (e) a^{x-y}.

26. $ma-c=b$, then a equals (a) $\frac{b\ m}{c}$, (b) $\frac{m-b}{b}$, (c) $\frac{b+c}{m}$, (d) $b+m/c$, (e) $b+c/m$.

27. $m^5 \cdot m^{-10}$ equals (a) m^5, (b) m^{-5}, (c) m^{50}, (d) m^{10}, (e) m^{-5}.

28. $5/b=a/x$, then x equals, (a) $ab/5$, (b) $5/ab$, (c) $\frac{a+b}{5}$, (d) $\frac{5}{a+b}$, (e) $\frac{a+5}{b}$.

29. $x=m.n$, then $m=$(a) x/n, (b) xmn, (c) n/m, (d) nx, (e) xm.

30. If $m=a+b$, how much greater is m than b? (a) a, (b) $a-b$, (c) $b-a$, (d) $-a$, (e) $a+b$.

31. 6 is 6 1/4% of (a) 98, (b) 96, (c) 97, (d) 79, (e) 69.

32. The expression $3x+2mn^2$ is a (a) monomial, (b) polynomial, (c) binomial, (d) unknown, (e) trinomial.

33. $ma+c=b$, then a equals (a) $\frac{b-m}{c}$, (b) $\frac{m-b}{c}$, (c) $\frac{b+c}{m}$, (d) $b-c/m$, (e) $\frac{b-c}{m}$.

34. m^2-13m-48=0. The negative root is (a) -4, (b) -12, (c) no correct answer.

35. x=m/n, then m=(a) x/n, (b) n/x, (c) xn, (d) nx, (e) mx.

36. The quotient of x^2/x^6 is (a) x^4, (b) x^8, (c) x^{12}, (d) x^{-4}, (e) x^{-2}.

37. The expression $3x^2$ is a (a) monomial, (b) polynomial, (c) binomial, (d) unknown, (e) trinomial.

38. $\frac{12x-15x^3}{-3x}$ equals (a) $4+5x^2$, (b) $-4+5x^2$, (c) $4-5x^2$, (d) 4-5x, (e) 4+5x.

39. $\sqrt{\frac{16}{8}}$ equals (a) $2\sqrt{16}$, (b) $\sqrt{2}$, (c) 2, (d) 8, (e) $\sqrt{8}$.

40. x.y=m . 2πr, then x/m= (a) $\frac{2\pi r}{y}$, (b) $\frac{y}{2\pi r}$, (c) 2πry, (d) 2πrx, (e) 2 ry.

41. 6/a - 7/b equals (a) $\frac{6b-7a}{a-b}$, (b) $\frac{6b-7a}{ab}$, (c) $\frac{6a-7b}{ab}$, (d) $\frac{6a-7b}{a-b}$, (e) $\frac{1}{a-b}$.

42. 6/a+7/b equals (a) $\frac{6b+7a}{ab}$, (b) $\frac{6a+7b}{ab}$, (c) $\frac{13}{a+b}$, (d) $\frac{13ab}{ab}$, (e) $\frac{6a+7b}{a+b}$.

43. $\frac{n^2-7n+10}{n^2+4n-12}$ when reduced equals (a) $\frac{n+5}{n+6}$, (b) $\frac{n+5}{n-6}$, (c) $\frac{n-5}{n-6}$, (d) $\frac{n-5}{n+6}$, (e) $\frac{n-6}{n-5}$.

44. $6\sqrt{2} - \sqrt{2} + 4\sqrt{2}$ equals (a) $10\sqrt{2}$, (b) $10\sqrt{4}$, (c) $9\sqrt{2}$, (d) $2\sqrt{9}$, (e) $11\sqrt{2}$.

45. $\sqrt{75}$ simplified equals (a) $5\sqrt{3}$, (b) $3\sqrt{3}$, (c) $3\sqrt{5}$, (d) $3\sqrt{15}$, (e) $5\sqrt{5}$.

46. m^5+m^{-10} equals (a) m^{-5}, (b) m^{+5}, (c) m^{15}, (d) m^5, (e) no correct answer.

47. m^5-m^{-10} equals (a) m^{-5}, (b) m^5-m^{10}, (c) m^5+m^{10}, (d) m^5-m^{-10}, (e) m^5.

48. m^2-13m-48=0. The positive root is (a) 4, (b) 12, (c) 3, (d) 16, (e) 24.

49. m^5/m^{-10} equals (a) m^{-15}, (b) m^{-5}, (c) m^{15}, (d) m^5, (e) m^{-5}.

50. ma=5a-5, then a equals (a) $\frac{5}{m-5}$, (b) $\frac{5}{m+5}$, (c) $\frac{-5}{m+5}$, (d) $\frac{-5}{m-5}$, (e) $\frac{5}{m-5}$.

ACKNOWLEDGMENTS

The author wishes to acknowledge gratefully the assistance rendered by Dr. J. C. Peterson in the construction of the test, in giving advice on the systematic recording of data, and in the preparation of this thesis.

Acknowledgment is also given to Mr. H. D. Karns, Principal of the Junction City High School, and Mr. O. E. Hodgson, Principal of the Salina High School for their cooperation in permitting the testing of the senior students.

LITERATURE CITED

(1) Adkins, Dorothy C. and Herbert A. Toops.
Simplified formulas for item selection and construction. Psychometrika. 2(3). Sept. 1937.

(2) Bliss, Elisha F., Jr.
The difficulty of an item. Jour. Ed. Psychol. 20: 63-66. 1929.

(3) Henry, Lorne J.
A comparison of the difficulty and validity of achievement test items. Jour. Ed. Psychol. 25: 537-541. 1934.

(4) Kelly, Truman L.
Statistical method. New York. Macmillan. p. 193-194. 1932.

(5) Lentz, Theo. F., Jr., Bertha Hirshstein and F. H. Finch.
Evaluation of methods of evaluating test items. Jour. Ed. Psychol. 23: 344-350. 1932.

(6) Swineford, Francis.
Validity of test items. Jour. Ed. Psychol. 27: 68-78. 1936.