

Development of Computer Package Usage for Item Analysis

Boonrèang Kajornsin¹ and Preecha T. Tragul²

ABSTRACT

The objectives of this study are 1) to develop a microcomputer package for test item analysis called ITEMPC, 2) to develop a manual for using ITEMPC, 3) to analyze multiple choice test item response by using ITEMPC and 4) to compare ITEMPC results with SPSS/PC+.

ITEMPC is a microcomputer package that teachers can use to do multiple choice test item analysis. Information available from ITEMPC consists of 3 parts 1) test statistics, including mean, standard deviation, the K-R 20 reliability coefficient, frequency distribution and inter-correlation matrix between scale scores; 2) test item information, including frequencies and percentage of each selected option, item facilities or item difficulties, point-biserial correlations between the score on the individual item and the sum of the scores on the remaining items and flags for four different levels of significance testing of point-biserial correlations; and 3) individual information, including test scores and the responses of each item.

Seven entrance examination tests for master students of Department of Education at Kasetsart University are analyzed by using ITEMPC and SPSS/PC+. All the results from those two packages are the same.

KEY WORDS : item analysis, microcomputer software, item difficulty, item discrimination, reliability.

INTRODUCTION

One of the important problems in teaching-learning process across all levels of education is student achievement evaluation. At present, most teachers in Thailand usually not analyzed the quality of the achievement test used for evaluating their students' achievement. Therefore, sometimes the teachers use low quality test in terms of validity, reliability, item difficulty and item discrimination. The reason why teachers do not analyze the test quality is due to a lack of appropriate instrument that they can use easily, as well as a lack of simple and efficient technique of analysis.

ITEMANL package was developed in 1987 (Kajornsin, 1987) which was the software

program implemented on mainframe computer for doing multiple choice item analysis. It has been shown to be very useful for the teachers or instructors who teach in schools or universities where mainframe computers are available. The problem is that schools in Thailand do not have mainframe computers, so it is a burden for teachers to use ITEMANL to analyze their test. Now social and behavioral science researchers have used SPSS/PC+ (Norusis, 1990) for analyzing their data with a microcomputer with a hard disk. SPSS/PC+ version 4.0 can be used to determine reliability coefficient, item difficulty and item discrimination index, but in order to use SPSS/PC+ one must have a microcomputer with a hard

¹Dept. of Education, Faculty of Education, Kasetsart University, Bangkok 10900

²Office of University Computer Service, Kasetsart University, Bangkok 10900

disk. Most of schools in Thailand have micro-computers, but not many schools have micro-computer with hard disks. Therefore, SPSS/PC+ is rarely used in Thai schools. Due to this limitation, it is useful to develop a software program for item analysis that can be used with microcomputer without a hard disk to allow more opportunities for teachers to do item analysis and to improve the quality of their test.

OBJECTIVES

The objectives of this study were:

1. to develop a microcomputer package for test item analysis called ITEMPC,
2. to develop a manual for using ITEMPC,
3. to analyze multiple choice test item responses by using ITEMPC and
4. to compare ITEMPC results with SPSS/PC+.

METHODOLOGY

In this study, the item analysis program (ITEMPC) was written in FORTRAN. The ITEMPC package is an item analysis program for dichotomously scored multiple choice test items. This program consists of 3 parts. The first part is concerned with subject statistics which includes scores and item responses. The second part is contained test or scales statistics which includes mean, standard deviation, Kuder-Richardson formula 20 reliability coefficient (Cited in Kajornsinsin, 1990: 163), frequency distribution and inter-correlation matrix between scales if more than one scale is defined. The final part deals with item statistics which includes absolute number and percentage selecting each alternative, item facility, unbiased pointbiserial correlation coefficient (Cited in Kajornsinsin, 1990:121) for each alternative and flags for four different levels of significance of the point-biserial correlation coefficient. Point-biserial correlation coefficient is used to determine the item discrimination index. The ITEMPC package can be run

on any kinds of microcomputer with or without hard disk.

After the completion of the ITEMPC development a manual for using ITEMPC was developed by Boonreang Kajornsinsin. To determine the efficiency of the ITEMPC, seven entrance examination tests for master student of Department of Education at Kasetsart University in 1990 were analyzed by using ITEMPC and SPSS/PC+. The values of item facility, point biserial correlation coefficient of item discrimination index, Kuder-Richardson formula 20 reliability coefficient for each scale, scale mean and standard deviation were determined by using both packages. By using SPSS/PC+, the command of RELIABILITY was used to determine the values of item facility point biserial correlation coefficient, reliability coefficient, mean and standard deviation of each scale. The values of corrected item-total correlation that obtained from SPSS/PC+ are the values of point-biserial correlation coefficient or the values of the correlation coefficient between the score on the individual item and the sum of the scores on the remaining items. (Norusis, 1990:B-190)

RESULTS

1. The researchers obtain item analysis computer program namely ITEMPC that can be run on any kinds of microcomputer. By using ITEMPC package 3 parts output will be obtained; the first part dealing with subject statistics, the second part dealing with scale statistics and the final part dealing with item statistics. The example of printout is showed in the appendix.

2. A manual for using ITEMPC composes of 2 parts; the first part shows how to set up the data set, the second part displays the example of data set and output obtained from ITEMPC, including the interpretation of the result.

3. The results of item analysis of seven entrance examinations for master students of the Department of Education at Kasetsart University are shown in table 1 to table 3.

Table 1 Number of items classified by subject matter and item facility.

Subject matter	Item facility					Total
	< 20%	20-39%	40-60%	61-80%	> 80%	
Aptitude test part 1	7	10	15	6	2	40
Aptitude test part 2	2	11	17	5	5	40
Aptitude test part 3	3	11	16	7	2	39
Sciences part 1	8	3	8	7	14	40
Sciences part 2	5	10	13	6	6	40
Sciences part 3	-	20	14	6	-	40
Social studies	22	29	33	21	15	120
Thai	21	22	26	7	4	80
English part 1	3	2	2	12	11	30
English part 2	7	-	11	6	6	30
English part 3	9	11	17	2	1	40
Research and evaluation part 1	3	5	12	7	3	30
Research and evaluation part 2	5	10	3	8	4	30
Administration	11	16	16	18	9	70

Table 2 Number of items classified by subject matter and item discrimination index.

Subject matter	Item discrimination index			Total
	< .20	.20-.39	> .40	
Aptitude test part 1	38	2	-	40
Aptitude test part 2	12	20	8	40
Aptitude test part 3	4	23	12	39
Sciences part 1	23	14	3	40
Sciences part 2	22	17	1	40
Sciences part 3	13	18	9	40
Social studies	85	30	5	120
Thai	56	20	4	80
English part 1	14	12	4	30
English part 2	16	7	7	30
English part 3	21	14	5	40
Research and evaluation part 1	13	10	7	30
Research and evaluation part 2	27	3	-	30
Administration	64	6	-	70

Table 1 shows that some items of all seven tests should be revised because they are too easy or too difficult. For example, Aptitude test part 1, 7 items are too difficult and 2 items are too easy.

Table 2 shows that some items of all seven tests should be revised because they

have low quality in terms of discrimination power. For example, Aptitude test part 1, the discrimination index of 38 items are too low.

Table 3 shows that Aptitude test part 1, Research and evaluation part 2 and Administration test have low quality in terms of reliability,

Table 3 Test reliability coefficient and test statistics classified by subject matter.

Subject matter	Number of items	Reliability coefficient	Mean	Standard deviation
Aptitude test part 1	40	.362	17.194	3.445
Aptitude test part 2	40	.788	19.646	5.960
Aptitude test part 3	39	.848	18.095	6.985
Sciences part 1	40	.582	22.759	3.781
Sciences part 2	40	.561	19.611	4.127
Sciences part 3	40	.801	16.704	6.282
Social studies	120	.683	55.915	8.262
Thai	80	.612	29.750	6.117
English part 1	30	.645	20.771	3.482
English part 2	30	.745	15.800	4.757
English part 3	40	.657	15.886	4.757
Research and evaluation part 1	30	.731	15.393	4.555
Research and evaluation part 2	30	.303	14.750	2.710
Administration	70	.328	34.723	4.392

all of these tests should be revised.

4. The values of individual statistics, scale statistics and item statistics obtained from ITEMPC and SPSS/PC+ are the same.

CONCLUSION

Analysis of item responses using ITEMPC provide the information of 1) individual statistics which includes scores and item responses; 2) scale statistics which includes mean, standard deviation, Kuder-Richardson formula 20 reliability coefficient, frequency distribution and inter-correlation matrix between scales if more than one scale is defined; 3) item statistics which includes absolute number and percentage selecting each alternative, item facility, item discrimination index which is determined by unbiased point-biserial correlation coefficient for each alternative and flags for four different levels of significance of the point-biserial correlation coefficient. The results of individual statistics, scale statistics and item statistics obtained from ITEMPC and SPSS/PC+ are the same. Therefore, we can conclude that ITEMPC is an efficient computer program for analyzing the multiple choice items. Besides its efficiency, three other advantages resulting

from using ITEMPC are: First, ITEMPC output is easy to read and understand; second, ITEMPC provides the item discrimination for the distractors while SPSS/PC+ provides only the value of item discrimination of the correct alternative; third, ITEMPC can be used with a microcomputer with or without hard disk while SPSS/PC+ cannot.

RECOMMENDATION

The test to be analyzed by the ITEMPC package should have not more than 200 items, 15 scales and 10 different response sets of items.

LITERATURE CITED

- Kajornsinsin, Boonreang. 1987. *Item Analysis by Using the ITEMANL Package*. Paper to be presented at the 13th International Association for Educational Assessment Conference.
- Kajornsinsin, Boonreang. 1990. *Educational Research Methodology*. Bangkok: Physics Center Publishing Company.
- Norusis, Marija J. 1990. *SPSS/PC+ Statistics 4.0 for the IBM PC/XT/AT and PS/2*. Chicago : SPSS Inc.

APPENDIX 1

STATISTICAL FORMULAS

The computing formulas for item facility, point-biserial correlation coefficient and Kuder-Richardson formula 20 reliability coefficient are follows:

Item facility = (Number of correct responses / Total number of responses) X 100

Point-biserial correlation coefficient of item i

$$= \sqrt{pq} (\bar{X}_p - \bar{X}_q) / s$$

where \bar{X}_p is the mean score on the test excluding item i of examinees who passed the item,

\bar{X}_q is the mean score on the test excluding item i of examinees who failed the item,

p is proportion of correct responses,

q is proportion of incorrect responses.

S is standard deviation of total scores excluding item i of all examinees.

Kuder-Richardson formula 20 reliability coefficient

$$= k/(k-1)(1 - (\sum pq)/S^2)$$

where k is the number of items,

p is proportion of correct responses

q is proportion of incorrect responses

S^2 is variance of total scores of all examinees.

APPENDIX 2

EXAMPLE OF ITEMPC PRINTOUT

KASETSART UNIVERSITY FACULTY OF EDUCATION PROGRAM-PROGRAM ITEM ANALYSIS
USING WITH MICRO COMPUTER

ITEM ANALYSIS FOR RESEARCH AND EVALUATION 1991 (2534)

NUMBER OF ITEMS = 60
NUMBER OF SCALES = 2
DATA FORMAT(I6,2X,60I1)

ITEMS SCORED ON VARIOUS SCALES

ITEM NUMBER	CORRECT RESPONSE	SCALE NUMBER	SCALE LABEL
1	4	1	RES.
2	3	1	RES.
3	4	1	RES.
4	3	1	RES.
5	4	1	RES.
6	2	1	RES.
7	2	1	RES.
8	1	1	RES.
9	2	1	RES.
10	1	1	RES.
11	2	1	RES.
12	1	1	RES.
13	2	1	RES.
14	3	1	RES.
15	3	1	RES.
16	2	1	RES.
17	3	1	RES.
18	3	1	RES.
19	1	1	RES.
20	2	1	RES.
21	2	1	RES.
22	2	1	RES.
23	3	1	RES.
24	3	1	RES.
25	2	1	RES.
26	3	1	RES.
27	1	1	RES.
28	1	1	RES.
29	1	1	RES.
30	4	1	RES.
31	2	2	EVAL.

SERIAL NUMBER	SCALE LABELS AND SCORES				ITEM RESPONSES FROM TO		
585297	RES. 13 EVAL. 14	*****	0	1	40		
+X++++XXXXXX+XXX+XX+X+X+XX+X+XXXXXXXXXX+							
585297	***** 0 ***** 0	*****	0	41	60		
XX+X++X++XX++++X++++							
585279	RES. 14 EVAL. 14	*****	0	1	40		
+++X++X+XXXX+X+X+XXXX++XXXX+X+XXX+XXX++							
585279	***** 0 ***** 0	*****	0	41	60		
XX+X++XXX+X++++X++XX							
585151	RES. 23 EVAL. 17	*****	0	1	40		
+X++++++X++++X++++++X+X+XX++++XX+X+XXX+X							
585151	***** 0 ***** 0	*****	0	41	60		
+X++++XX+XX++++++X							
585133	RES. 17 EVAL. 15	*****	0	1	40		
++X+++X+XX++++++XX++XX+XXX+XXX++XXX+XX							
585133	***** 0 ***** 0	*****	0	41	60		
XX++++X++XXX++X++X							
585044	RES. 19 EVAL. 12	*****	0	1	40		
++++++XXXX+XX+XX++X++++XX++XXXX++XXXX							
585044	***** 0 ***** 0	*****	0	41	60		
X++++XXX+X++XX+XX+XX							
585024	RES. 24 EVAL. 18	*****	0	1	40		
++++++XX++X++++++X++XX++++XX+X++++++							
585024	***** 0 ***** 0	*****	0	41	60		
X++X+++X+XX++X++XX+X							
585005	RES. 6 EVAL. 13	*****	0	1	40		
XX+XXXXXXXX+X+XXXXXXXX+XXXXX+X+XXXXX+XXXX							
585005	***** 0 ***** 0	*****	0	41	60		
XX++++XX++X++X+X+++X							
585004	RES. 17 EVAL. 16	*****	0	1	40		
+++X+XXXX+XX++X++X++XXX+X+XX+X+XXXX+							
585004	***** 0 ***** 0	*****	0	41	60		
+X+++XXX++X++X+X++++							
584960	RES. 17 EVAL. 16	*****	0	1	40		
+X++++X+XX++XX+X++++XX++XXX+X+X++X+XXX+X							
584960	***** 0 ***** 0	*****	0	41	60		
+X++++XX+XX++X++XX++							
584929	RES. 23 EVAL. 15	*****	0	1	40		
++++++X++++X+X+X++XX++X+XX+X+X+XXX							
584929	***** 0 ***** 0	*****	0	41	60		
+X++++XX++++X+X+XXX							
584883	RES. 24 EVAL. 15	*****	0	1	40		
++++++XX++++++X++X++X+X+XX+X+X++XX							
584883	***** 0 ***** 0	*****	0	41	60		
+X++++XXX+X+XX+X+X++							

ITEM RESPONSE KEY

+= CORRECT ANSWER

*= NO ANSWER

X= INCORRECT ANSWER

ITEM ANALYSIS FOR RESEARCH AND EVALUATION 1991 (2534)

FREQUENCY COUNT OF SCORES

SCALE = RES.

MAXIMUM SCORE ATTAINABLE = 30

SCORE	FREQ
0	0
1	0
2	0
3	0
4	0
5	0
6	2
7	0
8	0
9	3
10	5
11	8
12	10
13	6
14	7
15	6
16	4
17	5
18	7
19	4
20	3
21	3
22	1
23	6
24	4
25	0
26	0
27	0
28	0
29	0
30	0

ITEM ANALYSIS FOR RESEARCH AND EVALUATION 1991 (2534)

FREQUENCY COUNT OF SCORES

SCALE = EVAL.

MAXIMUM SCORE ATTAINABLE = 30

SCORE	FREQ
0	0
1	0
2	0
3	0
4	0
5	0
6	0
7	0
8	1
9	1
10	4
11	4
12	6
13	12
14	11
15	9
16	15
17	7
18	9
19	2
20	1
21	2
22	0
23	0
24	0
25	0
26	0
27	0
28	0
29	0
30	0

ITEM ANALYSIS FOR RESEARCH AND EVALUATION 1991 (2534)

POPULATION = 84 CANDIDATES

STATISTICS

SCALES	MEAN SCORE	STANDARD DEVIATION	K.R.20 REL COEFF

RES.	15.393	4.555	.731
EVAL.	14.750	2.710	.303

ITEM ANALYSIS FOR RESEARCH AND EVALUATION 1991 (2534)

INTERCORRELATION COEFFICIENTS FOR SCALE SCORES

RES. (1) 1.000 .392
 EVAL. (2) .392 1.000

ITEM ANALYSIS FOR RESEARCH AND EVALUATION 1991 (2534)

ITEM ANALYSIS DATA FOR ITEM 1
 ITEM IS ON SCALE 1 (RES.)

NUMBER SELECTING EACH OPTION

1	2	3	4*	5	N.A.
4	9	28	41	0	0
(4%)	(10%)	(33%)	(48%)	(0%)	(0%)

FACILITY = 48.81%

POINT BISERIAL CORRELATIONS WITH SCALE SCORES (UNBIASED)

1	2	3	4*	5	N.A.
-.09	-.24+	-.30++	.40++++	.00	.00

KEY TO SYMBOLS

+ = PROBABILITY LESS THAN .05
 ++ = PROBABILITY LESS THAN .01
 +++ = PROBABILITY LESS THAN .001
 ++++ = PROBABILITY LESS THAN .0001

ITEM ANALYSIS FOR RESEARCH AND EVALUATION 1991 (2534)

ITEM ANALYSIS DATA FOR ITEM 2
 ITEM IS ON SCALE 1 (RES.)

NUMBER SELECTING EACH OPTION

1	2	3*	4	5	N.A.
16	16	39	13	0	0
(19%)	(19%)	(46%)	(15%)	(0%)	(0%)

FACILITY = 46.43%

POINT BISERIAL CORRELATIONS WITH SCALE SCORES (UNBIASED)

1	2	3*	4	5	N.A.
-.34+++	-.04	.19(.1)	.00	.00	.00