DOCUMENT RESUME

ED 270 456 TM 860 296

AUTHOR	Reckase, Mark D.; Ackerman, Terry A.
TITLE	Building a Test Using Items That Require More than One Skill to Determine a Correct Answer.
SPONS AGENCY	Naval Personnel and Training Research Lab., San Diego, Calif.
PUB DATE	Apr 86
CONTRACT	N00014-85-C-0241
NOTE	38p.; Paper presented at the Annual Meeting of the American Educational Research Association (70th, San Francisco, CA, April 16-20, 1986).
PUB TYPE	Speeches/Conference Papers (150) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Achievement Tests; College Entrance Examinations; *Estimation (Mathematics); Goodness of Fit; High Schools; *Item Analysis; *Latent Trait Theory; Mathematical Models; Mathematics Tests; Psychometrics; Research Methodology; Sample Size; Tables (Data); *Test Construction; *Test Items
IDENTIFIERS	Ability Parameters; ACT Assessment; Item Dimensionality; Item Discrimination (Tests); *Multidimensional Approach; *Unidimensionality (Tests)

ABSTRACT

This paper demonstrates the relationship between the concept of unidimensionality and direction of an item in a multidimensional space. The basic premise is that if items that measure in the same direction are combined to form a test, that test will meet the item response theory requirements of unidimensionality. This will be true even if the items measuring in the same direction measure more than one psychological construct. A form of the ACT Mathematics Usage Test was analyzed using the multidimensional extension of the two-parameter logistic model to determine the direction for each item using the multidimensional difficulty formulated by Reckase. Using the direction, three unidimensional sets of items were formed and one multidimensional item set. The performance of 1,000 examinees on these items sets was analyzed using LOGIST 5 to determine the fit of the three-parameter logistic model to the data and the relationship of the unidimensional ability estimates and item parameter estimates to the multidimensional counterparts. Overall, the results strongly support the conception of unidimensionality suggested by a common direction in the multidimensional space for a set of items and the use of multidimensional difficulty statistics in forming unidimensional item sets. (PN)

Building a Test Using Items that Require More than One Skill to Determine a Correct Answer

Mark D. Reckase

Terry A. Ackerman

The American College Testing Program

ED270456

Running head: BUILDING A TEST REQUIRING MORE THAN ONE SKILL

Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 1986. This research was supported by Contract No. N00014-85-C-0241, NR153-531 with the Personnel and Training Research Programs of the Office of Naval Research.

Building a Test Using Items that Require More than One Skill to Determine a Correct Answer

Many, if not all, test items require more than one skill to determine the correct answer. For example, many mathematics problems require some verbal skill to determine what is required by the problem and mathematical skill to solve the problem. Even vocabulary items might require knowledge in several content domains to determine the correct answer choice. Yet, tests are constructed regularly from items that require more than one skill and examinees are rank ordered on the scores that are a function of combinations of these skills. The more complex the function of the skills required to relate the skills to the total score on the test, the more difficult is the task of interpreting the score.

With the development of item response theory (IRT) (Birnbaum, 1968; Rasch, 1960; Lord, 1980), the problem of analyzing the results of tests composed of items that require more than one skill for a correct response became more critical because most IRT models assume that the construct being measured is unidimensional. However, the concept of unidimensionality required by IRT is not the same as the commonly held conception of unidimensionality. The IRT definition states that a test is unidimensional if all persons with the same estimate of ability have the same probability of a correct response for each item. This definition does not require that the estimate of ability be a function of a single psychological trait or construct. As an alternative, each item could require the same combination of traits or constructs. Thus, from an IRT perspective, a test would be unidimensional even though the items were very complex in terms of the skills required for solution as long as all of the items required the same combination of skills.

This paper formalizes the above concept of unidimensionality by considering the relationship between a description of a set of test items in a multidimensional latent space and their performance when considered as a unidimensional test. Based upon the theory presented, sets of items that will function as unidimensional tests can easily be identified using the results of a multidimensional IRT analysis. This approach will be demonstrated using items from a mathematics achievement test.

Theoretical Background

The basic assumption underlying the approacn taken in this paper is that each test item is best at measuring a specific combination of skills. One test item may require a lot of mathematical computation skill and only a little verbal reasoning skill, while another item may require just the opposite combination. The particular combination of skills required for the solution of an item can be determined by computing the multidimensional item difficulty (MID) of the item (Reckase, 1985). This statistic indicates the combination of skills that form a multidimensional space for which the test item provides the best discrimination.

MID is described by two pieces of information: the direction from the origin of the multidimensional space to the point in the space where the item is most discriminating, and the distance from the origin to that point. For example, suppose that the relationship between the skills in a two-dimensional space and the probability of a correct response on a test item is given by the surface shown in Figure 1a. This item is most discriminating in a direction of 34.7° from the θ_1 axis. The distance to the most discriminating point in that direction is 1.58 units. This information is shown graphically in Figure 1b.

Insert Figure 1 about here

If the relationship between the skills in the multidimensional space and the probability of correct response to a test item can be represented by the multidime sional extension of the two-parameter logistic (M2PL) model, the MID information can be obtained from the equation presented in Reckase (1985). If the M2PL model is given in the form

$$P\,(x_{ij} = 1 \mid a_i,\, d_i,\, \theta_j) = \frac{e^{a_i' \theta_i + d_i}}{1 + e^{a_i' \theta_j + d_i}} \tag{1}$$

where x_{ij} is the 0, 1 score on item i by person j,

a_i is a vector of discrimination parameters, (a_{i1}, a_{i2}, ..., a_{in}),

d_i is a scalar value related to the item difficulty,

and θ_j is a vector of ability parameters, then the direction of the item in terms of direction cosines is

$$\cos \alpha_{ik} = \frac{a_{ik}}{\left[\sum_{k=1}^{n} a_{ik}^2\right]^{1/2}} \tag{2}$$

and the signed distance is

$$D_i = \frac{-d_i}{\left[\sum_{k=1}^{n} a_{ik}^2\right]^{1/2}} \qquad (3)$$

where n is the number of dimensions. For the example given in Figure 1, $a_{i1} = 1.3$, $a_{i2} = .9$, $d_i = -1.0$, $\cos \alpha_{i1} = .822$, $\cos \alpha_{i2} = .569$, and $D_i = 1.58$.

When items have the same direction of maximum discrimination in the multidimensional space, lines of equal probability on the item characteristics surfaces will be parallel to each other. This implies that regions can be defined in the space such that all persons in the region have the same probabilities of correct response to all the items. If all persons in one of these "iso-probability" regions are assigned the same ability estimate, the set of items meet the definition of unidimensionality given by Lord and Novick (1968). Thus, a set of items with the same direction will function as if they were unidimensional even though they may require more than one skill to obtain a correct solution. Items that have different directions will form a multidimensional set from the IRT perspective.

In order to test this theory, a set of items that require more than one skill for a correct response was analyzed using an approach based on the M2PL model and the MID statistics. This analysis and the results are described in the next two sections of this paper.

Analysis of the ACT Mathematics Usage Test

The procedure to be followed to determine whether selecting items on the basis of the MID would yield unidimensional subsets of items was to analyze a test to estimate the parameters of the M2PL model, compute the MID estimates for each item, sort the items according to direction, and then analyze the item sets using the three-parameter logistic (3PL) model. The results of the 3PL analysis would then be studied to determine whether they supported the conclusion that unidimensional item sets had been formed.

The test data used for these analyses were obtained from an administration of the ACT Mathematics Usage Test in February 1983. A systematic sample of responses from 1,000 students was selected from the total number of students who took the test at that time. The ACT Assessment Mathematics Usage Test is a 40 item, multiple choice measure of general achievement in high school mathematics. A more complete description of the test is given in Figure 2.

Insert Figure 2 about here

The estimates of the M2PL item parameters for the items on the test were determined using the MAXLOG program (McKinley & Reckase, 1983). This program estimates the parameters using a joint maximum likelihood procedure. For this demonstration, a two-dimensional solution was used so that the results could be readily represented graphically. The results of this analysis are given in Table 1.

Insert Table 1 about here

The parameter estimates from the M2PL model were used to compute the direction and distance for each of the items on the test. These statistics are also presented in Table 1. The directions and distances for the items are represented as vectors in Figure 3. Because of the overlap of the vectors near the origin, the items located within ±.5 on each dimension are shown in Figure 3b while the rest of the items are shown in Figure 3a.

Insert Figure 3 about here

Four sets of items were selected from the 40 items on the ACT Mathematics Usage Test for this study. Set 1 included all items within 10° of the 10° direction from Dimension 1. Set 2 included all items within 10° of the 45° direction from Dimension 1. Set 3 included all items within 10° of the 70° direction from Dimension 1. Set 4 included items that were approximately evenly spaced between 0° and 90° from Dimension 1. Set 4 was included as a part of this study to serve as a basis for comparison for the other item sets since it is a truly multidimensional item set. An example item from each of Sets 1, 2, and 3 is given in Figure 4.

Insert Figure 4 about here

Table 2 presents a summary of these four sets of items. The mean and standard deviation of the directions for each item set is given in the table along with the number of items in each set. Each item set is shown graphically in Figure 5. Note that item Sets 2 and 4 have about the same mean direction, but Set 4 has a much larger standard deviation.

Insert Table 2 about here

Insert Figure 5 about here

Figures 6a, 6b, 6c, and 6d show the characteristics of the four item sets in another way. The figures show the amount of information provided by each item set in directions that are incremented by 10°. Item Set 1 provides most of its information in a direction parallel to Dimension 1. Item Set 2 provides most information at a 45° angle. This angle is equivalent to an equally weighted composite of abilities. Item Set 3 provides most information in a direction slightly clockwise from 90°. Thus, the information tends to be concentrated along Dimension 2. The amount of information provided by item Set 4 varies with the position in the θ-space. In the region around (0,-3) the information is greatest parallel to Dimension 1. At (2,1) the information is greatest along Dimension 2. At (-1,1) the information is spread over all directions. This information structure supports the multidimensional nature of this set of items.

Insert Figure 6 about here

Each of these four item sets was analyzed using the LOGIST 5 (Wingersky, Barton, & Lord, 1982) 3PL estimation program. There was some concern about using LOGIST for this purpose since some of the item sets are fairly small (Set 3 has eight items). However, the estimates of the parameters for all four item sets seem to have converged properly, so the results of the program were used in the study. Still, the small number of items should be considered when interpreting the results.

For each of the item sets, the item parameter estimates were used to test the goodness of fit of the 3PL model to the data and to obtain ability parameter estimates. The ability parameter estimates were then correlated to determine whether they matched the pattern predicted by the average direction of the item set in the θ-space. The a-parameter estimates were compared to the slope of the item characteristic surface in the average direction of item set.

Results

Goodness of Fit

The 3PL item parameter estimates and the results of Bock's (1972) goodness of fit chi-square test for each item set are presented in Table 3. Notice that Sets 1 and 2 were fit very well by the 3PL model. If a .01 level of significance is used to reject the hypothesis of fit to compensate for the large sample size and the multiple significance tests, no items are rejected for Set 1 and three items are rejected for Set 2. Item Sets 3 and 4 showed serious deviations from fit. In both cases the 3PL model was rejected for most of the items in the item sets.

Insert Table 3 about here

Ability Estimates

For each of the sets of items, an ability estimate was computed for each person. A number-correct score was also computed for each item set as an alternative ability estimate, since the IRT-based estimates could be unstable for the small item sets. The estimates for the two dimensions in the multidimensional latent space were also obtained from the MAXLOG program. The intercorrelations between these ability estimates are given in Table 4.

Insert Table 4 about here

A comparison of the correlations between the ability estimates obtained from the first three item sets shows that the ranking of these correlations is exactly what would be expected based on the orientation of those item sets in the multidimensional space. A simplex pattern is evident with the item sets that have adjacent directions having the higher correlations. This is true for both the raw scores and the θ-estimates.

The correlations between the estimates from the first three item sets and the dimensions of the multidimensional space also have the predicted pattern of relationships. Estimates from Set 1 correlate most highly with Dimension 1, Set 3 correlates most highly with Dimension 2, but at a lower level than Set 1 with Dimension 1, and Set 2 correlates with both dimensions. The raw scores have higher correlations with the dimensions than the θ-estimates suggesting possible estimation problems occurring in the application of LOGIST to the small item sets.

The ability estimates from Set 4 correlate with the estimates from the other three sets, as they should since Set 4 tends to measure in all directions in the space. One anomaly in the correlations with the Set 4 estimates is present, however. The Set 4 estimates correlated .50 with Dimension 1 and only .08 with Dimension 2. A possible explanation of this result might be that the items measuring Dimension 2 in Set 4 were fairly difficult and did not differentiate among individuals in the sample (see Figure 4d). Further research will be needed to check this hypothesis.

Item Discrimination

If the unidimensional ability estimate scale obtained from LOGIST is a linear function of the dimensions in the ability space, then the a-parameter estimates from the unidimensional analysis should be related to the slope of the item response surface along the unidimensional scale. To determine if this were so, the correlation was computed between the a-parameter estimates and the slope estimates for each item in the average direction for each item set. These correlations are given in Table 5 along with the correlation with the a-parameter estimates from the multidimensional model.

Insert Table 5 about here

Only for Set 1 was the correlation between the a-parameter estimates and the slope estimates higher than the correlation with the M2PL a-parameter estimates. This implies that either the unidimensional scale is not a linear function of the dimensions in the multidimensional space, the estimates of the parameters are unstable, or the M2PL model is inappropriate.

To further probe the discrimination of the items in the sets, the average point-biserial correlation between the item responses in a set and the score on the item set minus the item in question was computed. These average correlations are given in Table 6. For the first two item sets, the mean point-biserial was significantly higher than for the last two sets. Set 4 would be expected to have a lower average discrimination because it is very heterogenous in content. Set 3, however, should have a higher average point-biserial. The observed findings could be a result of the difficulty of Set 3. The items in that set have an average p-value of .28. This is substantially harder than the other sets. This set of items appears to be too difficult for many of the examinees in the sample.

Insert Table 6 about here

Discussion

The purpose of this paper has been to demonstrate the relationship between the concept of unidimensionality and direction of an item in a multidimensional space. The basic premise is that if items that measure in the same direction are combined to form a test, that test will meet the IRT requirements of unidimensionality. This will be true even if the items measuring in the same direction measure more than one psychological construct.

In order to demonstrate this method for forming sets of items that are unidimensional in the IRT sense, a form of the ACT Mathematics Usage Test was analyzed using the multidimensional extension of the two-parameter logistic model to determine the direction for each item using the multidimensional difficulty formulated by Reckase (1985). Using the direction, three unidimensional set of items were formed and one multidimensional item set. The performance of 1,000 examinees on these item sets was analyzed using LOGIST 5 to determine the fit of the three-parameter logistic model to the data and the relationship of the unidimensional ability estimates and item parameter estimates to the multidimensional counterparts.

The results of the analyses showed that two of the three unidimensional item sets yielded responses that were consistent with a unidimensionality assumption: the sets were fit well by a unidimensional model, the correlations with each other and the dimensions of the space had the predicted pattern, and discrimination statistics suggested a homogeneous item set.

Item Set 3 did not fit the pattern. However, this item set was both short and difficult. The results for that set could possibly be explained by these factors. The multidimensional item set yielded results that matched the predicted characteristics.

Overall, the results presented in this paper strongly support the conception of unidimensionality suggested by a common direction in the multidimensional space for a set of items and the use of multidimensional difficulty statistics in forming unidimensional item sets.

Further research is still needed, however. The M2PL model, which does not have a lower asymptote parameter, was used in this paper. A multidimensional model with a lower asymptote parameter is needed for use with multiple choice items. Estimation procedures for the parameters of such a model are available in programs like TESTFACT (Wilson, Wood & Gibbons, 1984), but their usefulness for computing multidimensional difficulty has not yet been determined.

The sample size requirements for multidimensional estimation also need to be determined and the sensitivity of the unidimensional procedures to spread in direction needs to be studied. Still, the use of multidimensional difficulty to form unidimensional item sets seems to be a very promising procedure.

References

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick, Statistical theories of mental test scores (pp. 397-479). Reading, Mass.: Addison-Wesley.

Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. Psychometrika, 37, 29-51.

Lord, F. M. (1980). Applications of item response theory to practical testing problems. Hillsdale, NJ: Laurence Erlbaum.

Lord, F. M. & Novick, M. R. (1968). Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley.

McKinley, R. L. & Reckase, M. D. (1983). MAXLOG: A computer program for the estimation of the parameters of a multidimensional logistic model. Behavior Research Methods & Instrumentation, 15(3), 389-390.

Rasch, G. (1960). Probabilistic models of some intelligence and attainment tests. Copenhagen: The Danish Institute for Educational Research.

Reckase, M. D. (1985). The difficulty of test items that measure more than one ability. Applied Psychological Measurement, 9(4), 401-412.

Wilson, D. T., Wood, R. & Gibbons, R. T. (1984). TESTFACT: Test scoring, item statistics, and item factor analysis. Mooresville, IN: Scientific Software.

Wingersky, M. S., Barton, M. A., & Lord, F. M. (1982). LOGIST user's guide. Princeton, NJ: Educational Testing Service.

Table 1

Item Parameters, Directions and Distances for the Items in the ACT Assessment Mathematics Usage Test

Item Number	a_{i1}	a_{i2}	d_i	$\cos \alpha_{i1}$	$\cos \alpha_{i2}$	α_{i1}	α_{i2}	D_i
1	1.81	.86	1.46	.90	.43	25	65	-.73
2	1.22	.02	.17	1.00	.02	1	89	-.14
3	1.57	.36	.67	.97	.22	13	77	-.42
4	.71	.53	.44	.80	.60	37	53	-.50
5	.86	.19	.10	.98	.21	12	78	-.11
6	1.72	.18	.44	.99	.10	6	86	-.25
7	1.86	.29	.38	.99	.15	9	81	-.20
8	1.33	.34	.69	.97	.25	14	76	-.50
9	1.19	1.57	.17	.60	.80	53	37	-.09
10	2.00	.00	.38	1.00	.00	0	90	-.19
11	.87	.00	.03	1.00	.00	0	90	-.03
12	2.00	.98	.91	.90	.44	26	64	-.41
13	1.00	.89	-.49	.75	.66	42	48	.37
14	1.22	.14	.54	.99	.11	7	83	-.44

(table continues)

Item Number	a_{i1}	a_{i2}	d_i	cos α_{i1}	cos α_{i2}	α_{i1}	α_{i2}	D_i
15	1.27	.47	.29	.94	.35	20	70	-.21
16	1.35	1.15	-.21	.76	.65	40	50	.12
17	1.06	.45	.08	.92	.39	23	67	-.07
18	1.92	.00	.12	1.00	.00	0	90	-.06
19	.96	.22	-.30	.97	.22	13	77	.30
20	1.20	.12	-.28	.99	.10	6	84	.23
21	1.41	.04	-.21	.99	.03	2	88	.15
22	1.54	1.79	.02	.65	.76	49	41	-.01
23	.54	.23	-.69	.92	.39	23	67	1.18
24	1.53	.48	-.83	.95	.30	17	73	.52
25	.72	.55	-.56	.79	.61	37	53	.62
26	.51	.65	-.49	.62	.79	52	38	.59
27	1.66	1.72	-.38	.69	.72	46	44	.16
28	.69	.19	-.68	.96	.27	15	75	.95
29	.88	1.12	-.91	.62	.79	52	38	.64
30	.68	1.21	-1.08	.49	.87	61	29	.78
31	.24	1.14	-.95	.21	.98	78	12	.82

(*table continues*)

Item Number	a_{i1}	a_{i2}	d_i	cos α_{i1}	cos α_{i2}	α_{i1}	α_{i2}	D_i
32	.51	1.21	-1.00	.39	.92	67	23	.76
33	.76	.59	-.96	.79	.61	38	52	1.00
34	.01	1.94	-1.92	.01	1.00	90	0	.99
35	.39	1.77	-1.57	.22	.98	78	12	.87
36	.76	.99	-1.36	.61	.79	52	38	1.09
37	.49	1.10	-.81	.41	.91	66	24	.67
38	.29	1.10	-.99	.25	.97	75	15	.87
39	.48	1.00	-1.56	.43	.90	64	26	1.41
40	.42	.75	-1.61	.49	.87	61	29	1.87

Table 2

Characteristics of Each Item Set

Item Set	$\bar{p}$	Average Direction	Standard Deviation of Direction	Number of Items
1	.51	7.0	5.7	14
2	.41	45.3	6.6	11
3	.28	68.8	7.2	8
4	.43	43.8	29.0	9

Table 3

3PL Item Parameter Estimates and Goodness of Fit Statistics for the Four Item Sets.

Item	Item	Parameter Estimates			Goodness of Fit Statistics	
Set	Number	a	b	c	χ^2	P
	2	.65	.12	.14	7.75	.36
	3	.89	-.25	.13	6.58	.47
	5	.52	.22	.14	13.65	.06
	6	1.05	-.14	.08	12.35	.09
	7	1.63	.12	.18	6.70	.46
	8	.67	-.39	.14	5.16	.64
1	10	1.43	-.18	.00	9.01	.25
	11	.43	.33	.14	10.20	.18
	14	.73	-.24	.14	13.99	.05
	18	.95	-.03	.03	10.30	.17
	19	.92	.97	.23	6.11	.53
	20	1.35	.68	.21	10.08	.18
	21	1.06	.45	.14	3.45	.84
	28	1.70	1.33	.24	12.54	.08

(table continues)

Item Set	Item Number	Parameter Estimates a	b	c	Goodness of Fit Statistics χ^2	P
	4	.31	-2.01	.00	80.38	.00
	9	.74	-.57	.00	37.94	.00
	13	1.51	.65	.20	9.65	.21
	16	1.66	.16	.13	16.09	.02
	22	1.61	-.03	.11	28.61	.00
2	25	1.20	1.12	.26	10.63	.16
	26	.58	1.23	.23	11.87	.11
	27	1.44	.18	.09	15.90	.03
	29	1.20	.81	.13	10.58	.16
	33	1.59	1.20	.19	8.95	.26
	36	.70	1.28	.07	14.47	.04

(table continues)

Item	Item	Parameter Estimates			Goodness of Fit Statistics	
Set	Number	a	b	c	χ^2	P
	30	1.66	.42	.10	51.87	.00
	31	.27	.28	.00	79.68	.00
	32	.29	.33	.00	55.65	.00
3	35	1.82	.90	.16	42.52	.00
	37	.57	-.05	.00	54.95	.00
	38	.88	.97	.20	70.25	.00
	39	1.33	1.40	.16	31.56	.00
	40	2.00	1.85	.18	13.07	.07

(table continues)

Item Set	Item Number	Parameter Estimates a	b	c	Goodness of Fit Statistics χ^2	P
	2	.39	-.74	.00	66.86	.00
	3	1.22	-.74	.00	94.40	.00
	4	.43	1.17	.00	21.12	.00
	13	2.00	.72	.20	24.31	.00
4	17	1.13	.56	.30	32.10	.00
	26	1.19	.89	.21	27.29	.00
	30	1.51	1.18	.17	24.39	.00
	34	2.00	1.71	.16	11.49	.12
	38	2.00	1.50	.23	9.10	.25

Note: All χ^2 statistics have seven degrees of freedom.

Table 4

Correlations between Ability Estimates

Ability Estimate	1	2	3	4	5	6	7	8	9	10
1. Raw Score 1		61	43	64	78	35	23	50	92	05
2. Raw Score 2			61	78	42	68	37	48	62	61
3. Raw Score 3				64	26	35	62	26	35	74
4. Raw Score 4					46	53	34	66	61	50
5. Theta 1						28	13	43	75	-03
6. Theta 2							29	47	37	42
7. Theta 3								21	20	46
8. Theta 4									50	08
9. Dim 1										-04
10. Dim 2										

Note: Correlations are presented without the leading decimal point. Extreme theta estimates ($\theta > 3$, $\theta < -3$) have been deleted.

Table 5

Correlations between the Unidimensional a-Parameter Estimates and the Slope and the Multidimensional a-Parameter Estimates

Item		Multidimensional a-Parameters	
Set	Slope	a_1	a_2
1	.35	.31	.10
2	.46	.58	.35
3	.11	.31	.06
4	.59	-.48	.78

Table 6

Mean Item Discrimination Values for Each Item Set

Item Set	Average $r_{pt.bis}$	Standard Deviation $r_{pt.bis}$
1	.39	.08
2	.40	.09
3	.32	.04
4	.29	.05

Figure Captions

Figure 1. Item response surface and difficulty vector for an item with parameters $\underline{a}_{i1} = 1.3$, $\underline{a}_{i2} = .9$, and $\underline{d}_i = -1.0$.

Figure 2. Content specifications for the ACT Mathematics Usage Test.

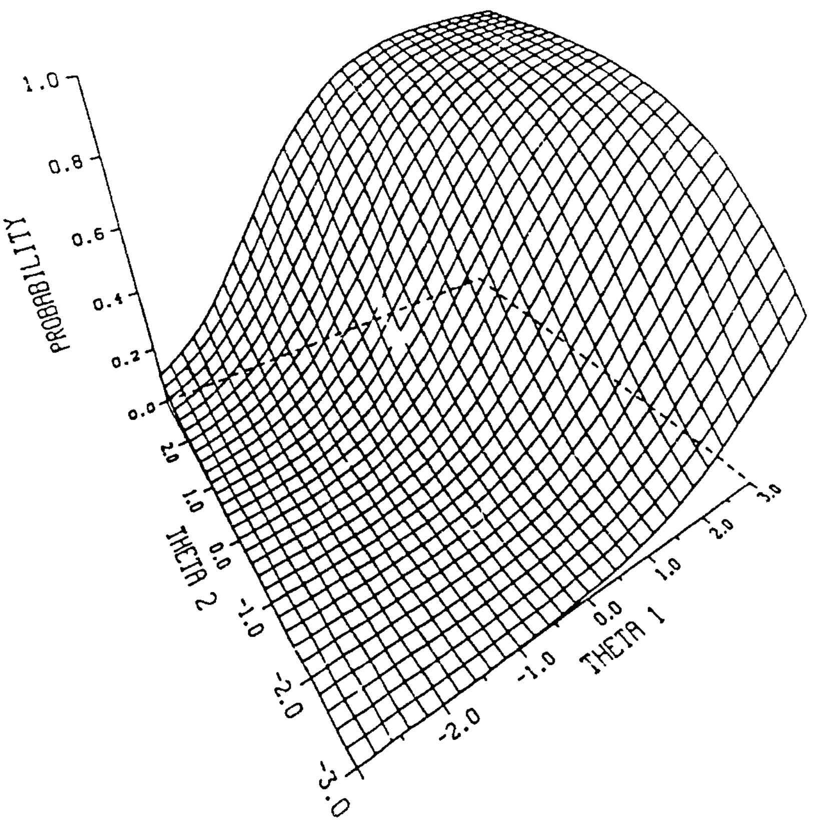
Figure 3. Multidimensional difficulty vectors for the ACT Mathematics Usage Test, Form 24B.

Figure 4. Example test items from Sets 1, 2, and 3.

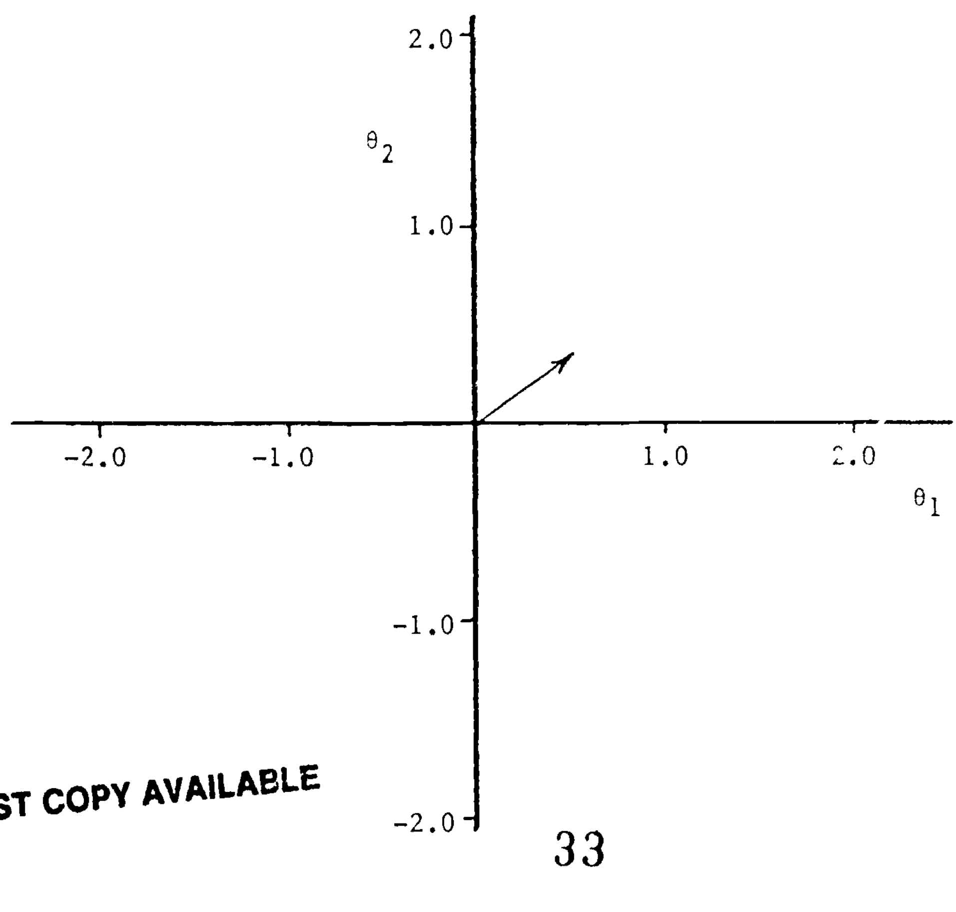
Figure 5. Multidimensional item difficulty vectors for item Sets 1, 2, 3, and 4.

Figure 6. Information plots for each of the four item sets.

a.



b.



BEST COPY AVAILABLE

ACT MATHEMATICS USAGE TEST

Description of the test. The Mathematics Usage Test is a 40-item, 50-minute test that measures the students' mathematical reasoning ability. It emphasizes the solution of practical quantitative problems that are encountered in many postsecondary curricula and includes a sampling of mathematical techniques covered in high school courses. The test emphasizes quantitative reasoning, rather than memorization of formulas, knowledge of techniques, or computational skill. Each item in the test poses a question with five alternative answers, the last of which may be "None of the above."

Content of the test. In general, the mathematical skills required for the test involve proficiencies emphasized in high school plane geometry and first- and second-year algebra. Six types of content are included in the test. These categories and the approximate proportion of the test devoted to each are given below.

Mathematics Content Area	Proportion of Test	Number of Items
a. Arithmetic and Algebraic Operations	.10	4
b. Arithmetic and Algebraic Reasoning	.35	14
c. Geometry	.20	8
d. Intermediate Algebra	.20	8
e. Number and Numeration Concepts	.10	4
f. Advanced Topics	.05	2
Total	1.00	40

a. *Arithmetic and Algebraic Operations.* The items in this category explicitly describe operations to be performed by the student. The operations include manipulating and simplifying expressions containing arithmetic or algebraic fractions, performing basic operations in polynomials, solving linear equations in one unknown, and performing operations on signed numbers.

b. *Arithmetic and Algebraic Reasoning.* These word problems present practical situations in which algebraic and/or arithmetic reasoning is required. The problems require the student to interpret the question and either to solve the problem or to find an approach to its solution.

c. *Geometry.* The items in this category cover such topics as measurement of lines and plane surfaces, properties of polygons, the Pythagorean theorem, and relationships involving circles. Both formal and applied problems are included.

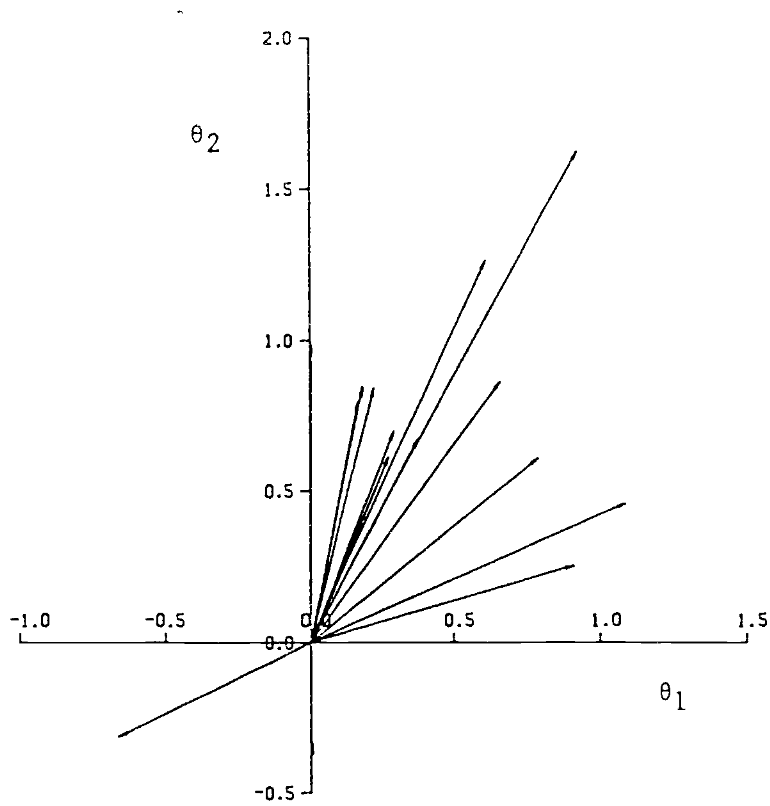
d. *Intermediate Algebra.* The items in this category cover such topics as dependence and variation of quantities related by specific formulas, arithmetic and geometric series, simultaneous equations, inequalities, exponents, radicals graphs of equations, and quadratic equations.

e. *Number and Numeration Concepts.* The items in this category cover such topics as rational and irrational numbers, set properties and operations, scientific notation, prime and composite numbers, numeration systems with bases other than 10, and absolute value.

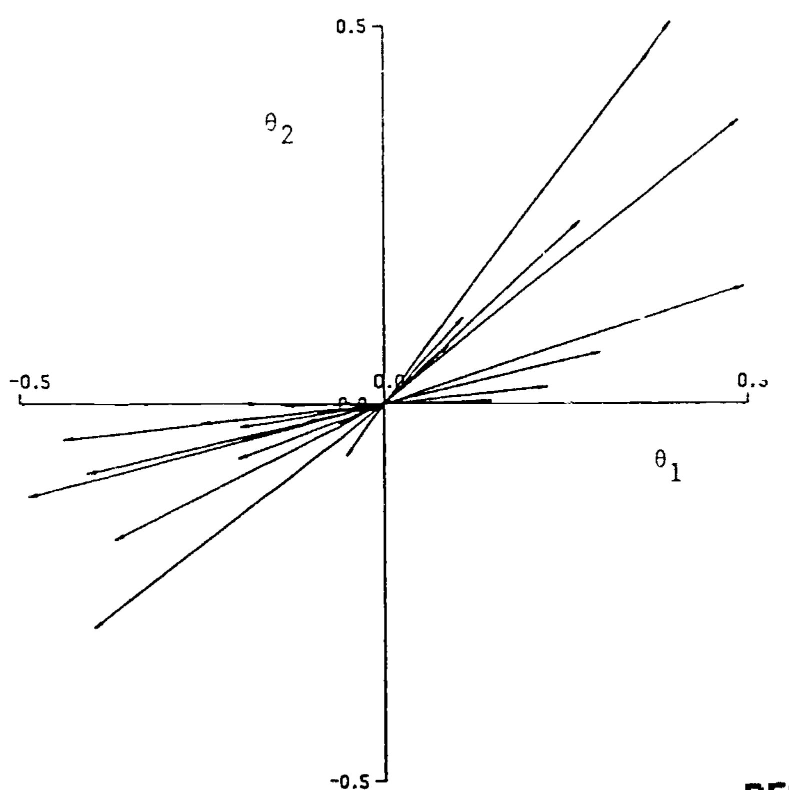
f. *Advanced Topics.* The items in this category cover such topics as trigonometric functions, permutations and combinations, probability, statistics, and logic Only simple applications of the skills implied by these topics are tested.

FIGURE 3

a.



b.



BEST COPY AVAILABLE

: 1

10. A bread recipe calls for $\frac{1}{2}$ cup of butter and $3\frac{1}{2}$ cups of flour. Using this recipe to make enough bread for a party, John will need $1\frac{1}{2}$ cups of butter. How many cups of flour will he need?

F. $4\frac{1}{2}$

G. $5\frac{1}{2}$

H. $7\frac{1}{2}$

J. $9\frac{1}{2}$

K. $10\frac{1}{2}$

t 2

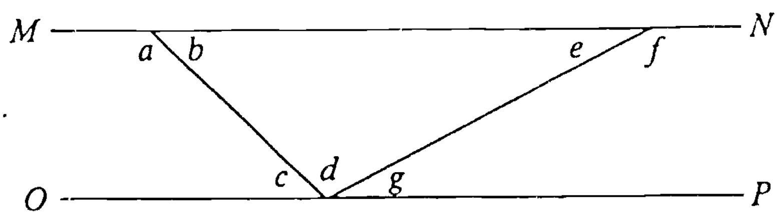
27. Which of the following equations expresses the relationship shown in the table below?

x	0	2	4	6	8	10
y	4	7	10	13	16	19

A. $y = 2x$
B. $y = \quad x + 4$
C. $y = \quad x + 9$
D. $2y = 3x + 4$
E. $2y = 3x + 8$

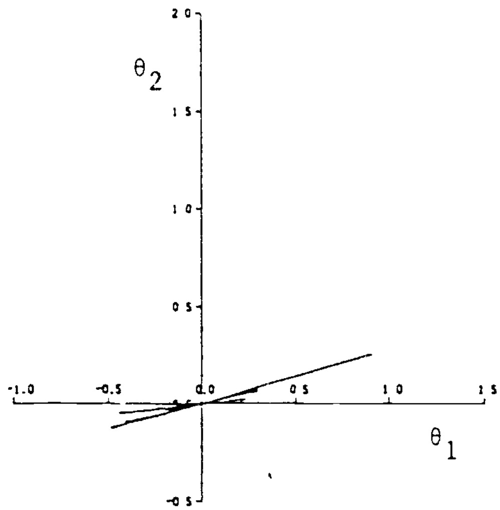
t 3

35. In the figure below, $\overline{MN} \parallel \overline{OP}$ and a, b, c, d, e, f, and g are the measures, in degrees, of their respective angles. Which of the following statements is NOT necessarily true?

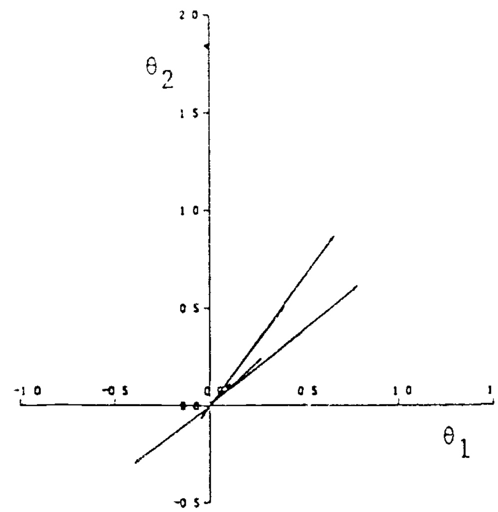


A. $b = 180° - d - c$
B. $e = 180° - d - c$
C. $a = 180° - c$
D. $f = 180° - g$
E. $g = 18$[illegible] $- f$

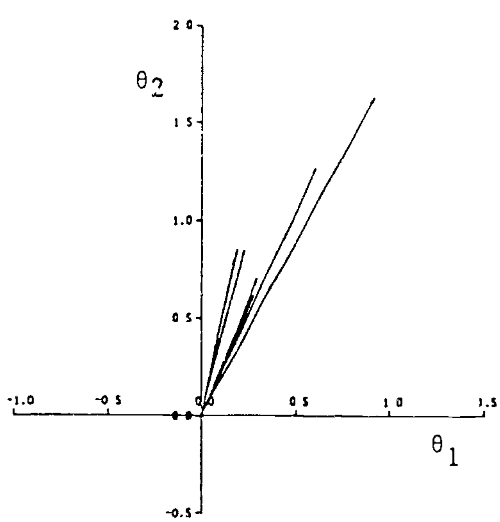
a. Set 1



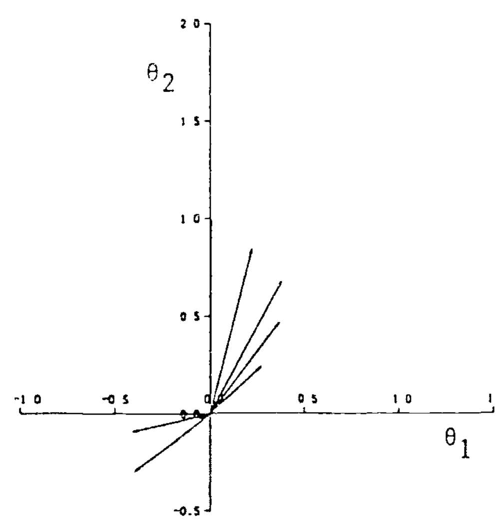
b. Set 2



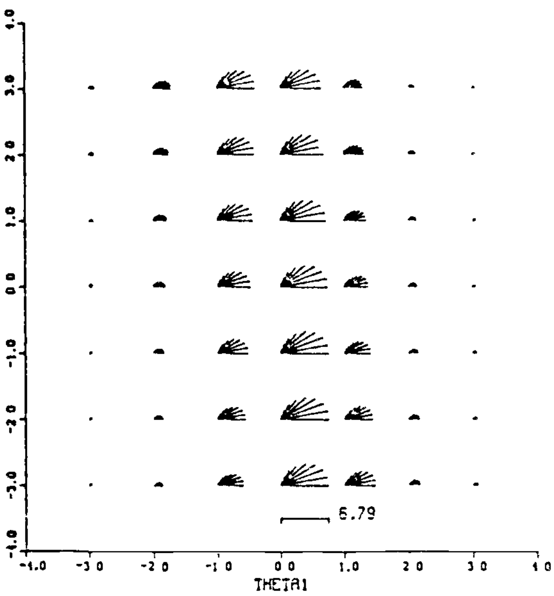
c. Set 3



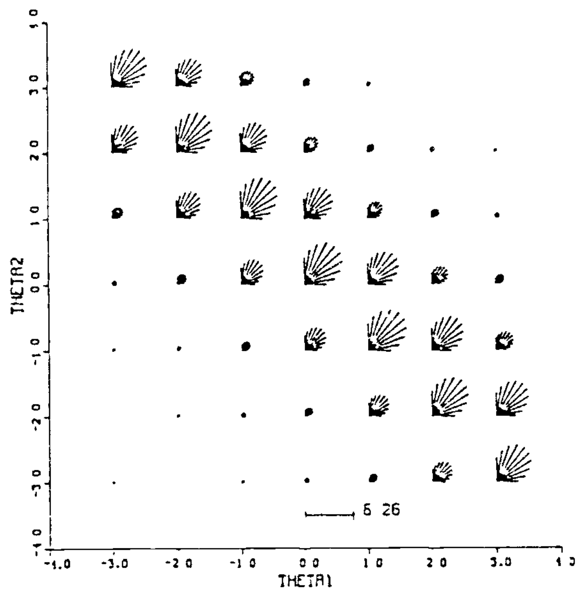
d. Set 4



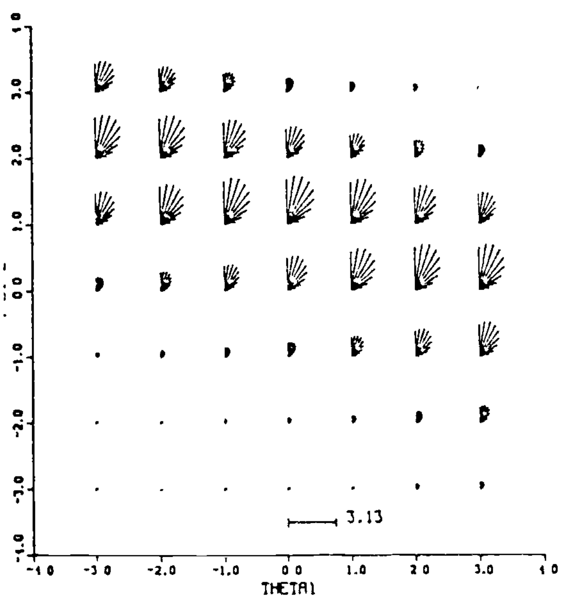
. Set 1



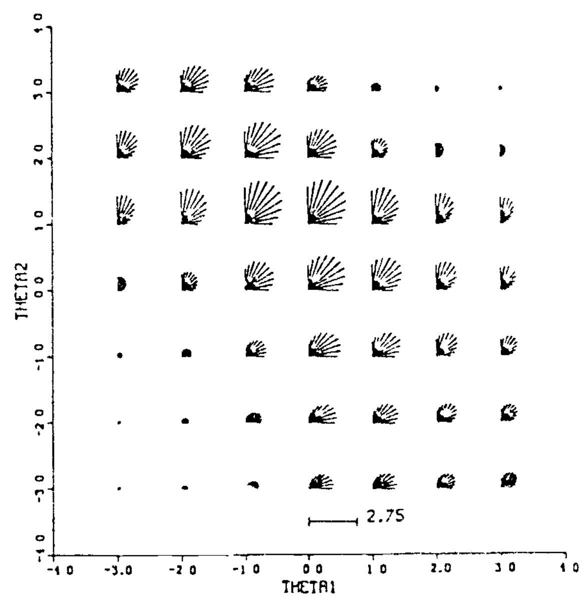
b. Set 2



. Set 3



d. Set 4



BEST COPY AVAILABLE