DOCUMENT RESUME

ED 306 238 TM 013 001

AUTHOR Reckase, Mark D.; And Others

TITLE Similarity of the Multidimensional Space Defined by Parallel Forms of a Mathematics Test.

PUB DATE Mar 89

NOTE 26p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, March 27-31, 1989).

PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.

DESCRIPTORS *Achievement Tests; Factor Analysis; High Schools; *High School Students; *Item Analysis; Latent Trait Theory; *Mathematics Tests; Multidimensional Scaling; Test Construction; *Test Format

IDENTIFIERS Dimensional Analysis; *Mathematics Usage Test (ACT); *Parallel Test Forms

ABSTRACT

The purpose of the paper is to determine whether test forms of the Mathematics Usage Test (AAP Math) of the American College Testing Program are parallel in a multidimensional sense. The AAP Math is an achievement test of mathematics concepts acquired by high school students by the end of their third year. To determine the dimensionality of the multidimensional space, each test form was factor analyzed at content area and item levels, and the number of dimensions needed to define the space was determined. A multidimensional item response theory analysis (MIRT) determined where in space each item provided the most information. A sample of 2,500 individuals for each form was taken from the equating administration in 1988 of five forms of the AAP Math: 29B, 29C, 29D, 29E, and 29F. A third step was to compare the information provided by each form at 49 points in the space defined by item-person interactions. Factor analysis showed that the tests had a very dominant first factor and a possible weaker second factor. MIRT analysis showed a fairly clear distinction between story-problem items and those only requiring manipulation, but the constructs measured by these item types were highly correlated. The information comparison indicated that the test forms all had the same basic structure. Considering that these test forms were not constructed with multidimensional considerations in mind, they are very similar. Eleven tables and five graphs present study data. (SLD)

Similarity of the Multidimensional Space Defined by Parallel Forms of a Mathematics Test

Mark D. Reckase

Tim Davey

Terry Ackerman

ACT

Paper presented at the meeting of the American Educational Research Association, San Francisco, March, 1989.

The Mathematics Usage Test (AAP Math) in the ACT Assessment Program (The American College Testing Program, 1988) is an achievement test which measures the mathematics concepts acquired by high school students up to the end of their third year of high school. It does not measure achievement in a particular course, but samples from the content taught in courses usually offered in grades seven through eleven. As is typical of most standardized achievement tests, forms of AAP Math are constructed according to a set of content and statistical specifications with the intent that the test forms meet the criteria for strict parallelism. To the credit of the test development staff, they regularly meet that goal to a high degree in a classical test theory sense. However, achievement tests that are constructed with an emphasis on content specifications are likely not to be unidimensional and it is uncertain whether the current test construction process yields tests that are parallel in a multidimensional sense when that is not specifically stated as a requirement in the test development process. That is, do all test forms measure all of the dimensions in the multidimensional space defined by the content domain in the same way, or does one form emphasize one dimension while others emphasize other dimensions.

The purpose of this paper is to determine whether a set of forms of AAP Math are parallel in a multidimensional sense even though that has not been a specific requirement of the test construction process. If they are, the current specifications are sufficient to produce strongly parallel forms. If they are not, more specific test specifications may be required to improve the parallelism of test forms. A secondary purpose of the paper is to describe a set of procedures that can be used to determine whether test forms are multidimensionally parallel.

Design of the Study

In order to determine whether test forms are parallel in a multidimensional sense, two factors must be considered. First, the dimensionality of the multidimensional space must be determined. Second, the information provided by the test for each linear composite of abilities in the space, here called the information structure, must be compared across forms. If the dimensionality and the information structure are the same for the test forms, they are multidimensionally parallel.

In order to determine the dimensionality of the space, each form of the test was factor analyzed at both the content area and the item level and the number of dimensions needed to define the space was determined. Once the dimensionality of the space was determined, a multidimensional IRT (MIRT) analysis was performed to determine where in the space each item provided the most information and the linear composite of abilities best measured by the item. The linear composite of abilities that defined the unidimensional score on the test, called the reference composite (Wang, 1987), was also determined.

Since there is some rotational indeterminacy in the MIRT solutions for each form, the MIRT solutions were rotated to yield the same reference composite as one of the forms arbitrarily selected as a base form. The implied assumptions are that each of the test forms is measuring the same dominant dimension and that this dimension is reflected in the reference composite. Following this rotation to a common reference composite, the multidimensional information functions were computed and compared across forms to determine whether they had the same information structure.

Data

The data used in this study was obtained from the equating administration of five forms of AAP Math: 29B, 29C, 29D, 29E and 29F. For the purposes of this paper they will be referred to as forms B through F. These forms were produced in the 1987-88 production year and were equated in Fall, 1988. A sample of 2500 individuals for each form was taken from the equating sample. These samples were considered randomly equivalent since the forms were distributed in spiralled fashion. This was an operational administration of the forms so examinees were motivated to do well.

The content specifications for AAP Math consist of six content areas and a number of test questions assigned to the content areas. Table 1 summarizes these content specifications. For this paper the content areas will be abbreviated as follows: Arithmetic and Algebraic Operations, AAO; Arithmetic and Algebraic Reasoning, AAR; Geometry, G; Intermediate Algebra, IA; Number and Numeration Concepts, NNS; and Advanced Topics, AT. It was a requirement that all forms have the specified number of items for each content area.

Insert Table 1 about here

The statistical specifications for the forms place constraints on the distribution of p-values and biserial correlation discrimination indices. The p-values for a form should have an average as close to .5 as possible and should range from .3 to .8 according to a specified distribution. Biserial correlations should be above .3 and should have an equal mean value across forms. When forms are constructed, content considerations take precedence, followed by item difficulty specifications, and then discrimination specifications. No attempt is made to match the difficulty or discriminating power of content areas across forms, although content areas tend to sort themselves on difficulty due to the inherent complexity of the material (i.e., algebra is generally more difficult than arithmetic). The ACT Assessment Program Technical Manual (The American College Testing Program, 1988) provides further information about the test construction process.

Results

To describe the parallelism of the test forms in a classical psychometric sense, a standard item and test analysis was performed on each test form. The results of this analysis are summarized in Tables 2 and 3. Given the sample size involved, some significant differences may exist in these form statistics, but in general, the forms appear very similar on the basis of classical statistics. They also clearly meet the statistical specifications set out for the test forms. Of course, the content specifications were also exactly met.

Insert Tables 2 & 3 about here

Two different factor analyses were performed on each test form to determine the appropriate number of dimensions to use for the MIRT analysis. The first analysis, a principal factor analysis, was performed on the Pearson product movement correlations between the number correct scores on each of the content areas. Thus, the analysis was performed on a six by six correlation matrix, with the variables being content areas. The correlation matrices for all five forms are given in Table 4.

Insert Table 4 about here

While there is some variation in the correlations between the content areas, particularly when the number of items in the content area is small (AT, 2; AAO, 4; NNS, 4), there are also noticeable regularities. The highest correlation with AAR is most frequently G. AAO tends to correlate most highly with IA. The magnitudes of the correlations are uniformly large.

Table 5 shows the six eigenvalues and the factor loadings on the first factor for each of the test forms. These results show that all of the forms have content areas that are dominated by one factor that accounts for 88% or more of the variance and that IA loads highest on the first factors on all analyses. NNS and AT have the lowest loadings for all analyses. In general, these results would suggest that AAP Math is highly unidimensional despite the multiple content areas.

Insert Table 5 about here

The item factor analysis was performed on the matrix of interitem tetrachoric correlations. The principal factor method was performed using the largest correlation with the item as a communality estimate. The eigenvalues greater than one for each form from this analysis are given in Table 6. The eigenvalues support a latent space dominated by one factor with possibly a second, minor factor. These results, and the results of the content area factor analyses, indicate that the data from the test forms have a dominant first factor, but that there may be a significant second factor. Therefore, it was decided to perform a two dimensional MIRT analysis on the data.

Insert Table 6 about here

The MIRT analysis was performed using the NOHARM program (Fraser, 1983) specifying a two dimensional solution. This program estimates the parameters of the multidimensional normal ogive model given by the formula

$$P(\underline{u}_{ij} = 1 | a_i, \underline{c}_i, \underline{d}_i, \theta_j) = \underline{c}_i + (1 - c) \int_{\gamma_{ij}}^{\infty} \phi(a_i'\theta_j + \underline{d}_i)\, d\theta_j$$

where ϕ ( ) is the normal density function. The parameters of this model can be converted to estimates of logistic item parameters by dividing them by 1.7. The d-parameter estimates from this model are presented for each item in Table 7 and the a-parameter estimates are presented in Table 8. A c-parameter of .16 was assumed for all analyses.

Insert Tables 7 and 8 about here

The values obtained from the NOHARM program are not directly comparable because of the rotational indeterminacy of the solution. In order to obtain comparable solutions, the reference composite for form B was arbitrary selected as a reference direction in the space and the solutions for the other four tests were orthogonally rotated so that they all had the same direction for their reference composites. The angles with the coordinate axes for the initial orientation of the reference composites are given in Table 9. Following the rotation, all reference composites had the orientation given by Form B. Note that the axes for Forms C and D were interchanged as well as rotated since the relationship between content and the two dimensions were mirror images for these two forms when compared to the other three forms.

Insert Table 9 about here

The rotated solution for each of the tests is given in Table 10. This table provides the distance from the origin (D) and angle from θ_2 (α_2) that defines the most discriminating direction for each item. These statistics are described in more detail in Reckase (1985). Table 11 provides average value of D and α_2 for each content area.

Insert Tables 10 and 11 about here

From an analysis of the results in these tables, it can be seen that the angles with the coordinate axes vary substantially across items. For example, the first AAO items on Forms C and D measure mostly θ_2 while many of the AAR items measure mostly θ_1 (are 90° from θ_2). A review of the text of these items shows that the former items are mainly arithmetic computation items. The latter items are story problems with significant amounts of text and little advanced mathematics. Thus, the items seem to vary on verbal problem solving versus computational and symbol manipulation skills. The AAR items, uniformly across forms, have the highest average angular distance from θ_2. This indicates the AAR items, in general, have the greatest requirement for verbal problem solving skills. The IA items are quite different than the AAR items with an angular difference between the two content areas of 25° to 30°. Such an angular difference would result in a correlation between true scores of about .86 suggesting why the factor analysis results implied a strongly unidimensional test. However, the results of the MIRT analysis imply that distinctly different skills are being assessed by the test even though the skills are highly correlated.

The purpose of the last set of analysis was to determine whether the forms measured the various skills with the same precision and emphasis. To determine whether the test forms were multidimensionally parallel, the multidimensional information was determined for each form for ten different directions in the space (ten different linear weightings of θ_1 and θ_2). The information in each of these directions is shown by the length of a line in that direction at each of 49 points in the two-dimensional ability space. Figure 1 gives the graphic representation of the information for Forms B through F. If the forms are multidimensionally parallel, the information plots for these forms should all look the same.

Insert Figure 1 about here

A review of the information plots shows substantial similarity and also some notable differences. The graphs show that all of the forms are most informative for persons above the mean, which was arbitrarily set at (0,0). Forms B, D and F provide somewhat more information than the other two forms. These forms also have higher KR-20 values (see Table 2). Form B seems to provide more information for the high ability examinees than the other forms, but overall the forms seem very similar in the pattern of information they provide.

Discussion and Conclusion

Achievement tests are typically designed to measure a complex of skills related to a curriculum area. These tests are inherently multidimensional in what they measure. Yet, a single score is often reported to summarize an examinees performance on such a test. This score may have slightly different meanings at different points on the score scale depending on the relationship between the contents and the difficulty of the items that measure it. Even though a set of test forms measure more than one ability, they still can be considered parallel if the interrelationship between the skills measured is the same across the forms. In fact, if the items on the test measure the same linear combination of skills, the test will be indistinguishable from a unidimensional test.

It was the purpose of the research reported here to investigate the multidimensional parallelism of five forms of the AAP Math test. Three steps were followed to investigate the multidimensional parallelism. First the tests were factor analyzed to determine the dimensionality of the space defined by the interaction of the examinees with the items. This analysis showed that the test forms had a very dominant first factor, but that a weaker second factor might be present. Therefore, further analyses were performed assuming a two dimensional space.

The second step was to perform a MIRT analysis on each of the five test forms. This analysis showed that there is a fairly clear distinction between the arithmetic items that are in story problem format and those items that only require computation or formula manipulation. However, the constructs measured by these two different types of items are highly intercorrelated.

The third step in the investigation was to compare the information provided by each form at 49 points in the space defined by the item/person interactions. The information comparison showed that the test forms all had the same basic information structure, but that there were some differences in the forms. Three forms seemed to provide more information than the other two and one of the forms seemed to provide more information for high ability examinees. Whether those differences are of practical consequence is unknown. But considering that these forms were not constructed with multidimensional considerations in mind, they are amazingly similar.

As always, this research points out the need for further research. Means are needed to determine how much of a difference in the information structure is important and how to better represent the differences. The information graphs contain substantial amounts of information, but they are difficult to compare directly. These topics will be the focus of future work. The main value of this paper, however, is that it is a beginning at building a methodology for studying multidimensional parallelism.

References

The American College Testing Program (1988). ACT Assessment Program Technical Manual. Iowa City, IA: Author.

Fraser, C. (1983). NOHARM II: A Fortran program for fitting unidimensional and multidimensional normal ogive models of latent trait theory. Armidale, Australia: University of New England, Centre for Behavioral Studies.

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

Wang, M. (1987, April). Estimation of ability parameters from response data to items that are precalibrated with a unidimensional model. Paper presented at the meeting of the American Educational Research Association, Washington, D.C.

Table 1

Specifications for the AAP 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.

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.

Table 2
Summary of Test Form Characteristics

	Statistic						
Form	$\bar{x}$	$\hat{s}_x$	$\bar{p}$	$\hat{s}_p$	$\bar{r}_{bis}$	$\hat{s}_r$	KR-20
B	18.8	8.7	.47	.13	.58	.10	.90
C	19.4	8.4	.49	.15	.56	.09	.89
D	19.3	8.7	.48	.15	.59	.11	.90
E	19.2	8.2	.48	.16	.56	.12	.89
F	20.5	8.9	.51	.14	.59	.10	.91

Note: Sample size is 2500 for each form.

Table 3
Item Statistic Distribution
for Each Form

Range of Values	p-values B	C	D	E	F	Range of Values	r_{bis} B	C	D	E	F
.00 - .19	0	0	0	1	0	.00 - .19	0	0	0	0	0
.20 - .39	13	14	14	13	9	.20 - .29	0	0	0	1	0
.40 - .59	18	16	16	13	17	.30 - .39	1	2	3	3	1
.60 - .79	8	10	10	13	13	.40 - .49	6	7	5	5	7
.80 - 1.00	1	0	0	0	1	.50 - 1.00	33	31	32	31	32

Table 4
Correlations Between Content Areas
for AAP Math Forms B Through F

Content Area	Form	G	NNS	IA	AAR	AT
AAO	B	.58	.49	.60	.61	.48
	C	.54	.48	.65	.58	.34
	D	.57	.46	.63	.59	.38
	E	.57	.54	.61	.59	.35
	F	.60	.50	.60	.57	.41
G	B		.49	.63	.66	.48
	C		.49	.61	.61	.37
	D		.50	.67	.66	.46
	E		.54	.63	.61	.40
	F		.51	.70	.67	.48
NNS	B			.56	.51	.41
	C			.53	.48	.29
	D			.55	.52	.34
	E			.57	.55	.34
	F			.53	.49	.38
IA	B				.60	.50
	C				.63	.38
	D				.65	.46
	E				.60	.39
	F				.62	.48
AAR	B					.50
	C					.43
	D					.43
	E					.38
	F					.46

Table 5
Eigenvalues and First Factor Loadings
for the AAP Math Forms

Eigenvalue Number	Form B	C	D	E	F
1	3.71	3.51	3.65	3.59	3.69
2	.60	.76	.69	.74	.64
3	.55	.56	.56	.48	.55
4	.43	.47	.44	.43	.44
5	.39	.38	.35	.40	.38
6	.33	.33	.31	.35	.29
Content Area	**First Factor Loadings**				
AAO	.75	.74	.72	.74	.72
G	.77	.74	.79	.76	.82
NNS	.65	.63	.64	.70	.64
IA	.79	.81	.82	.79	.81
AAR	.79	.78	.79	.76	.77
AT	.62	.49	.54	.49	.58

Table 6
Eigenvalues Greater than One
for the Principle Factor Analysis
of the Interitem Tetrachoric Correlations

	Form				
	B	C	D	E	F
	16.95	15.53	17.14	15.86	16.87
Eigenvalues	1.60	1.66	1.47	1.68	1.31
	1.05	1.12		1.18	
				1.01	

Table 7
d-parameter Estimates
for the Five AAP Math Forms

Item Content	Form B	Form C	Form D	Form E	Form F
AAO	.35	.90	1.17	.47	.50
AAO	-.16	.00	-.12	.31	1.00
AAO	-.24	-.19	.10	.31	.19
AAO	-2.50	-1.30	-1.11	-1.45	-.41
G	.80	-.28	-.05	-.29	.22
G	-.54	-.16	-.65	.54	-.58
G	-.43	-.60	-7.44	.08	-.62
G	-.70	-2.62	-3.50	-2.39	-.78
G	-14.10	-.64	-.77	-1.04	-.45
G	-1.66	-1.80	-.93	-.96	-1.95
G	-9.71	-1.14	-1.30	-1.60	-1.66
G	-6.43	-8.65	-2.38	-1.74	-2.37
NNS	.36	.02	.42	.59	.76
NNS	-.29	.36	.16	.15	.44
NNS	-.58	-.49	-.60	-.60	-.14
NNS	-.64	-1.54	-.94	-1.45	-1.43
IA	.12	.71	1.01	.30	.13
IA	.33	.03	.22	.29	.08
IA	-.25	.31	.11	.04	.14
IA	-2.18	-.31	.06	-.32	-.27
IA	-.50	-.20	-.43	-.90	-.53
IA	-7.81	-.63	-1.12	-.98	-1.34
IA	-1.17	-.74	-4.10	-.96	-2.28
IA	-.93	-1.27	-1.26	-15.08	-1.26
AAR	.57	.59	.85	.80	1.01
AAR	-.20	.46	.76	.77	.23
AAR	.36	.21	.30	.11	.21
AAR	-.28	.14	.02	-.09	-.06
AAR	-.22	.06	-.10	-.04	.14
AAR	.14	-.56	.03	.23	-.25
AAR	-.45	-.28	-.38	-.25	-.68
AAR	-.36	-.45	.41	-.53	-.09
AAR	-1.02	-.94	-1.06	-.94	-.31
AAR	-.84	-1.26	-.66	-1.27	-.54
AAR	-.62	-.93	-1.49	-1.57	-1.10
AAR	-1.10	-1.71	-1.28	-1.40	-4.48
AAR	-1.50	-1.49	-2.30	-5.56	-1.36
AAR	-4.93	-1.69	-2.89	-2.41	-1.69
AT	.13	-.15	-.78	-.12	.36
AT	-.90	.14	-1.30	-.69	-.95

Table 8
a-parameter Estimates
for the Five AAP Math Forms

Item Content	Form B a_1	Form B a_2	Form C a_1	Form C a_2	Form D a_1	Form D a_2	Form E a_1	Form E a_2	Form F a_1	Form F a_2
AAO	1.02	0.00	.91	0.00	1.44	.00	1.08	0.00	1.43	.00
AAO	1.34	.37	1.84	.76	1.08	.76	.80	.69	.60	1.28
AAO	1.01	.30	.70	.91	.84	.26	1.21	.62	.74	.96
AAO	2.67	1.58	1.35	.70	1.29	1.73	1.36	1.05	1.11	.67
G	.40	-.13	.87	.64	.51	.41	.86	.23	.80	.59
G	1.24	.20	.71	1.28	2.14	1.35	.38	.35	1.02	1.31
G	1.25	.37	.92	.44	5.55	4.23	1.11	1.11	.68	.61
G	.76	.20	2.54	2.20	2.38	2.79	4.00	2.35	.73	.54
G	10.51	3.98	.32	.28	.83	.84	1.10	.26	.76	.54
G	1.48	.42	.81	.42	.30	.51	.61	.25	1.68	1.67
G	8.42	1.95	.86	.44	1.03	.91	1.29	.14	2.14	2.52
G	6.06	3.26	3.53	3.48	2.22	1.43	1.24	.17	2.64	2.76
NNS	.84	.22	.56	.27	.72	.41	1.05	.62	.56	.80
NNS	.47	.19	.67	.42	.66	.49	.97	.76	.57	.55
NNS	.87	.60	.60	.52	.26	.29	.99	.44	.50	.55
NNS	.79	.49	.84	.49	.91	.66	.67	.11	.73	.72
IA	.94	.72	.90	.30	1.68	.44	.71	.65	.56	.82
IA	.94	.61	.60	1.18	1.11	.42	.62	.86	.78	1.26
IA	.61	.37	.80	.36	.85	.40	2.06	1.84	.42	.48
IA	3.00	2.36	1.46	.55	1.21	.50	.96	1.08	1.12	1.11
IA	1.07	.74	1.22	.74	1.13	.62	1.03	.98	.70	1.14
IA	7.51	4.02	1.16	.63	1.37	.96	1.12	.84	1.43	1.53
IA	1.05	.60	1.12	.48	4.07	2.68	.86	.57	2.01	1.81
IA	.83	.57	.98	.54	1.01	.90	6.34	2.25	.54	.47
AAR	.81	-.31	.34	.78	.49	.65	.53	.49	.46	.46
AAR	.48	.11	.41	.40	.83	.53	.86	.22	.81	.23
AAR	.94	-.27	.37	.80	.54	.53	1.15	.10	.72	.32
AAR	.67	.09	.49	.69	.70	1.15	.39	.26	.71	.15
AAR	.96	-.24	.46	.52	.64	.73	.62	.32	.45	.25
AAR	1.55	.11	.63	1.52	.52	.56	.65	.17	.80	.35
AAR	.69	.24	.37	.63	.79	1.40	1.35	-.05	1.42	.37
AAR	.66	-.19	1.07	.92	.61	1.19	.37	.07	.50	.30
AAR	.85	.25	.52	.98	.21	.77	.46	-.10	.45	.23
AAR	.76	-.01	.44	1.77	.33	.58	1.30	.02	1.37	.39
AAR	.89	-.00	.42	.78	.44	1.09	1.18	.43	1.30	.49
AAR	1.56	.19	.53	.93	.50	.66	1.25	.01	4.24	2.96
AAR	1.16	.92	1.13	.74	1.39	1.60	4.88	.80	1.21	.45
AAR	3.76	-.10	.56	.91	.76	1.91	1.46	-.34	1.25	.68
AT	.88	-.01	.31	.54	1.40	.65	.66	.42	.46	.25
AT	1.39	.87	.41	.67	.39	.81	.64	.10	1.19	1.29

Table 9
Angle of the Reference Composite
with each Axis for each AAP Math Form

Axis	Form				
	B	C	D	E	F
θ_1	31	47	49	24	41
θ_2	59	43	41	66	49

Table 10
Distance and Direction for the Items
on Each Test Form after Rotation
to the Form B Reference Composite

Item Content	Form B D	Form B α_2	Form C D	Form C α_2	Form D D	Form D α_2	Form E D	Form E α_2	Form F D	Form F α_2
AAO	-.34	90	-.99	16	-.81	18	-.44	83	-.35	99
AAO	.12	75	-.00	39	.10	53	-.29	42	-.71	35
AAO	.22	73	.16	69	-.11	35	-.23	56	-.16	47
AAO	.81	60	.85	44	.51	71	.84	45	.32	69
G	-1.91	108	.26	52	.08	57	.32	68	-.22	53
G	.43	81	.11	77	.26	50	-1.04	40	.35	47
G	.33	74	.59	42	1.20	61	-.05	38	.68	57
G	.88	75	.78	57	.96	67	.52	52	.86	63
G	2.35	48	1.50	58	.65	63	.92	69	.48	64
G	1.08	74	1.96	44	1.58	77	1.45	61	.82	55
G	1.98	67	1.18	43	.94	59	1.23	76	.50	50
G	1.16	54	1.74	61	.90	51	1.39	75	.62	53
NNS	-.42	75	-.03	42	-.51	47	-.49	52	-.78	44
NNS	.57	67	-.46	49	-.19	54	-.12	45	-.55	56
NNS	.55	55	.62	57	1.53	65	.55	59	.20	52
NNS	.69	58	1.58	46	.83	54	2.14	73	1.40	55
IA	-.10	52	-.75	34	-.58	33	-.31	40	-.13	43
IA	-.30	57	-.03	79	-.19	39	-.27	29	-.06	41
IA	.35	59	-.35	40	-.12	43	-.01	41	-.22	50
IA	.57	52	.20	37	-.04	41	.22	34	.17	55
IA	.38	55	.14	48	.34	47	.64	39	.40	41
IA	1.29	48	.48	45	.67	53	.70	46	.64	52
IA	.97	60	.61	40	.84	51	.93	49	.84	57
IA	.93	56	1.14	45	.93	59	3.00	56	1.75	58
AAR	-.66	111	-.70	83	-1 04	71	-1.11	40	-1.56	55
AAR	.41	77	-.80	61	-.77	50	-.87	68	-.27	83
AAR	-.36	105	-.24	81	-.40	62	-.10	78	-.27	75
AAR	.41	82	-.17	71	-.01	76	.19	49	.08	87
AAR	.22	104	-.09	65	.11	66	.06	56	-.28	70
AAR	-.09	86	.34	84	-.04	65	-.34	68	.29	76
AAR	.62	71	.38	76	.24	78	.18	85	.47	85
AAR	.53	106	.32	57	.31	81	1.43	72	.16	69
AAR	1.16	74	.84	78	1.34	93	2.01	95	.62	72
AAR	1.10	91	.69	92	.99	79	.98	82	.38	84
AAR	.70	90	1.05	78	1.28	86	1.26	63	.79	79
AAR	.70	83	1.60	76	1.55	71	1.12	82	.87	64
AAR	1.01	52	1.10	50	1.08	67	1.12	73	1.05	79
AAR	1.31	92	1.59	75	1.41	86	1.60	96	1.18	71
AT	-.15	90	.24	76	.50	43	.16	50	-.69	71
AT	.55	58	-.17	75	1.45	82	1.06	73	.54	52

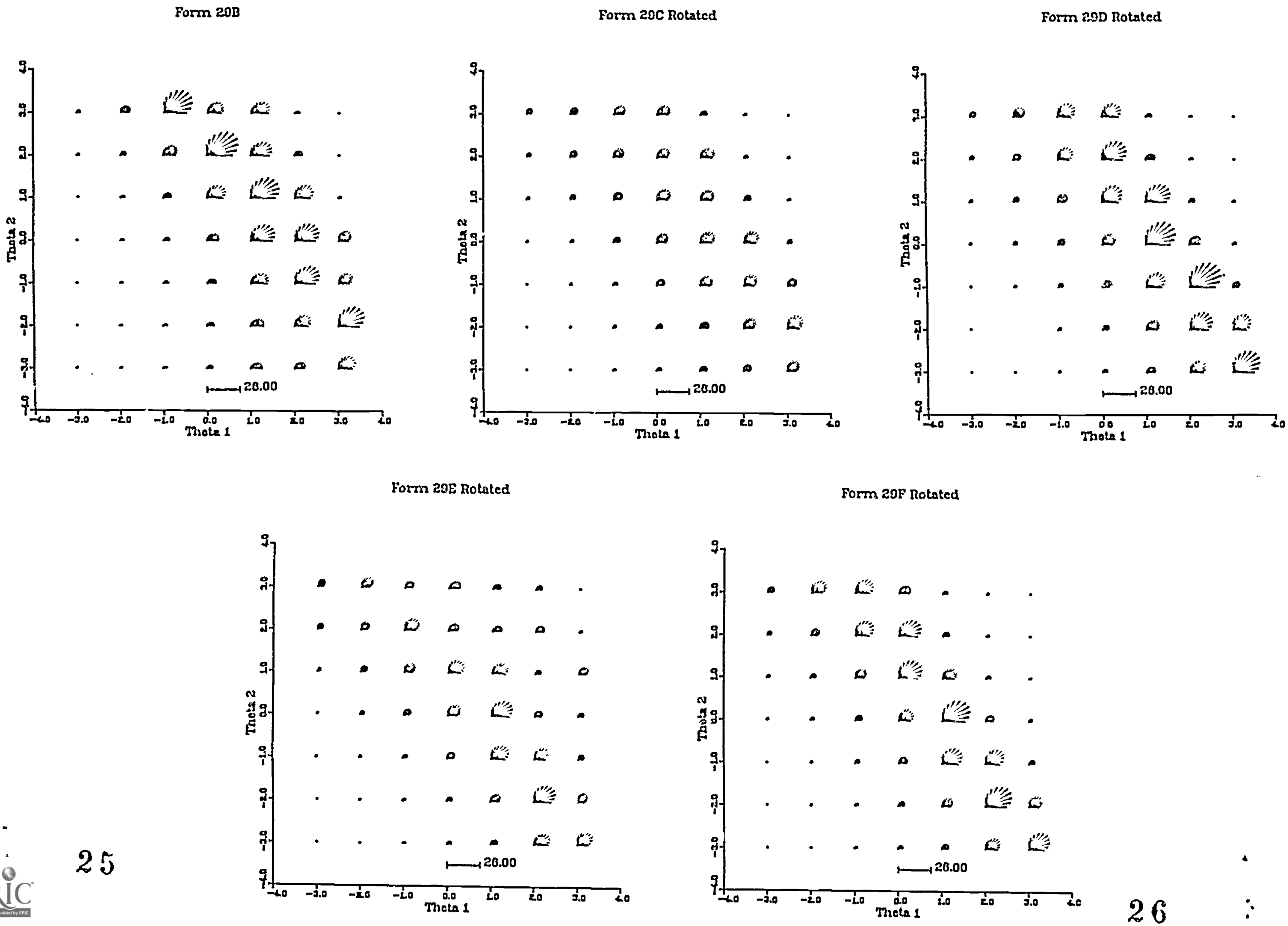
Table 11
Summary of Rotated MIRT Parameters
by Form and Content Area

Item Content	Form									
	B		C		D		E		F	
	D	α_2	D	α_2	D	α_2	D	α_2	D	α_2
AAO (4)*	.20	74	.00	42	-.08	44	-.03	56	-.22	63
G (8)	.79	73	1.02	54	.82	61	.59	60	.51	55
NNS (4)	.35	64	.43	48	.42	55	.52	57	.07	52
IA (8)	.51	55	.18	46	.23	46	.61	42	.42	50
AAR (14)	.50	87	.42	73	.43	74	.54	72	.25	75
AT (2)	.20	74	.04	76	.98	62	.61	62	-.08	62

* Number of items

Figure Caption

Figure 1. Information plots for the AAP Mathematics Forms 29B - 29F.


Form 20B
Form 20C Rotated
Form 20D Rotated
Form 20E Rotated
Form 20F Rotated
Theta 1
Theta 2
20.00
-4.0 -3.0 -2.0 -1.0 0.0 1.0 2.0 3.0 4.0