DOCUMENT RESUME

ED 281 872 TM 870 269

AUTHOR DeAyala, R. J.; Koch, William R.

TITLE Computerized Adaptive Testing: A Comparison of the Nominal Response Model and the Three Parameter Logistic Model.

PUB DATE Apr 87

NOTE 20p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Washington, DC, April 21-23, 1987).

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

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS *Adaptive Testing; *Comparative Testing; *Computer Assisted Testing; *Computer Simulation; Estimation (Mathematics); *Latent Trait Theory; Mathematical Models; Mathematics Tests; Predictive Measurement; Test Items

IDENTIFIERS *Nominal Response Model; Three Parameter Model

ABSTRACT

A nominal response model-based computerized adaptive testing procedure (nominal CAT) was implemented using simulated data. Ability estimates from the nominal CAT were compared to those from a CAT based upon the three-parameter logistic model (3PL CAT). Furthermore, estimates from both CAT procedures were compared with the known true abilities used to generate the simulated data. Results showed that the nominal CAT's ability estimates were highly correlated with those of the 3PL CAT as well as with the true abilities. Furthermore, the nominal CAT had a significantly higher association with negative true thetas than did the 3PL CAT, and it also had significantly lower standard errors of estimate than did the 3PL CAT. However, the nominal model-based CAT had difficulty estimating positive thetas and had a poor convergence rate. In contrast, the 3PL CAT had a high convergence rate and its performance was not affected by whether the true abilities were positive or negative. Potential reasons for the nominal CAT's high nonconvergence rate as well as implications for computerized adaptive testing were discussed. (Author)

ABSTRACT

Computerized Adaptive Testing: A Comparison of the Nominal Response Model and the Three Parameter Logistic Model

R. J. DeAyala and William R. Koch, The University of Texas at Austin

A nominal response model-based computerized adaptive testing procedure (nominal CAT) was implemented using simulated data. Ability estimates from the nominal CAT were compared to those from a CAT based upon the three-parameter logistic model (3PL CAT). Furthermore, estimates from both CAT procedures were compared with the known true abilities used to generate the simulated data. Results showed that the nominal CAT's ability estimates were highly correlated with those of the 3PL CAT as well as with the true abilities. Furthermore, the nominal CAT had a significantly higher association with negative true Θs than did the 3PL CAT and it also had significantly lower standard errors of estimate than did the 3PL CAT. However, the nominal model-based CAT had difficulty estimating positive Θs and had a poor convergence rate. In contrast, the 3PL CAT had a high convergence rate and its performance was not affected by whether the true abilities were positive or negative. Potential reasons for the nominal CAT's high nonconvergence rate as well as implications for computerized adaptive testing were discussed.

Computerized Adaptive Testing: A Comparison of the Nominal Response Model and the Three Parameter Logistic Model

OBJECTIVES

The objectives of this research were (a) to develop and implement a computerized adaptive testing (CAT) procedure based on Bock's (1972) nominal response model and (b) to compare the nominal-based CAT's performance in ability estimation with that of a more traditional three-parameter logistic (3PL) model-based CAT.

Computerized adaptive testing methods have primarily been based on two item response models : the three-parameter and one-parameter logistic models. Because both these models are based on dichotomously scored item responses, they ignore potentially useful information about the examinee contained in incorrect responses. The existence of information in incorrect responses has been demonstrated in several studies, e.g., Levine and Drasgow (1983), Thissen (1976).

Given the existence of information in incorrect responses as well as item response models which can access this information, it is reasonable to try to utilize this information for CAT ability estimation. Specifically, the ability estimate for an examinee who has enough partial knowledge to select an alternative which is almost correct should be higher than the ability estimate for the examinee who selects an incorrect alternative which is attractive only to the examinee of very low ability. By utilizing the information in incorrect responses one can administer items which are more appropriate for the examinee who responded incorrectly. The assessment of this partial knowledge requires a polychotomous model.

The potential advantages of a nominal-based CAT are the capability of using items in a CAT system which do not need to be scored dichotomously and the possibility of greater efficiency in ability estimation due to using all the information provided by the examinee's response.

METHOD

Programs : Two CAT programs were written, one program was based on the 3PL model (called 3PL CAT), whereas the other was based on the nominal response model (called the nominal CAT). Both programs used maximum likelihood estimation of ability and item selection was on the basis of item information. The adaptive testing simulation was terminated when either of two criteria were met : a maximum of thirty items was reached or when there were no more items in the item pool which had at least a predetermined minimum amount of information. The initial ability estimate for an examinee was the theta corresponding to the mode of the total test information distribution. Each CAT program was run on a simulated data set where the Θs which produced the response strings were known.

Data : A data set of 1093 examinees was created from five administrations of The College Board's[1] Achievement Test in Mathematics, Level I, at the University of Texas at Austin. This data set (called the original-1093) contained only individuals who answered at least 80% of the questions and the last question.

To work within the constraints of the calibration program, MULTILOG 5 (Thissen, 1986), one item was omitted from the data set and the 5-choice items of the Math Level I test were collapsed into 4-choice items. The item omitted had the lowest joint traditional difficulty and discrimination values of all the items on the test. MULTILOG 5 was used to calibrate the data for both the 3PL and nominal response models. It should be noted that the above constraints in MULTILOG are currently being removed.

Two simulated data sets were created based on the item parameters from the calibration of the original-1093 data set. The item parameter estimates for both the 3PL and nominal response model from the test were doubled to produce a 98-item pool for each model. By sampling from a normal distribution, 1000 examinees (their z-values were considered to be true Θs) were generated. These true Θs plus the 98 item parameter estimates for the 3PL and for the nominal response model were used to generate binary and non-binary response strings with a random error component for each simulated examinee, respectively. The former of these generated data sets will be referred to as Sim-3PL and the latter as Sim-Nom. Each simulation data set was subsequently calibrated to obtain item parameter estimates to be used in the simulated data set CAT runs.

RESULTS AND CONCLUSIONS

Table 1 shows the mean and standard deviation (S.D.) of the number correct (NC) score and coefficient alpha (alpha) for the various data sets and Table 2 presents the principal axes analysis of these data. As can be seen from these tables, all the data sets may be considered to be highly reliable. Further, although none of the data sets had a first factor which accounted for a large percentage of the total variance, each data set's first factor did account for a large proportion of the common variance. It was concluded that the

---

[1] The authors wish to thank The College Board for granting permission to use the Mathematices Level I Achievement test data.

data sets did not seriously violate the unidimensionality assumption and that the simulated data had desirable properties.

Insert Table 1 about here

Insert Table 2 about here

Table 3 presents statistics on the item pool information for the Θ range -4.0 to 4.0 for the simulation data sets. From the literature on the nominal response model (e.g., Bock, 1972; Thissen, 1976), it was expected that the nominal model would provide more information than the 3PL model for Θs less than 0.0. The entries for the nominal response model in Table 3 as well as Figure 1 are consistent with this belief. Although the simulation data sets did not contain extremely informative items, the nominal response model, provided, on the average, twice as much information per item for the simulated data as did the 3PL model (0.20 and 0.10, respectively). In fact, the total information computed over the -4.0 to 4.0 Θ range for the nominal response model was twice that of the 3PL model (Total Information = 1637.2 and 806.99, respectively). However, at approximately Θ = 1.5 and above the 3PL model provided more information than the nominal model.

Insert Table 3 about here

Insert Table 4 about here

Tables 4 and 5 present the analysis of the simulated data's estimated Θs for convergent cases for the 3PL CAT and nominal CAT, respectively. As these tables show, estimated Θs from both CATs are highly related with and accounted for 88% of the variability of the true Θs; Pearson Product Coefficients (r) were r = 0.94 for the 3PL CAT and r = 0.94 for the nominal CAT. These tables also show that nonconvergence of the Θ estimates was a problem with the nominal CAT, but not with the 3PL CAT. Inspection of the estimated Θ range for the two CATs for corresponding convergent cases (Table 6) shows that they shared approximately the same lower bound, but their upper bound differed considerably, with the nominal CAT's upper bound being less than that of the 3PL, 2.63 vs. 3.80, respectively.

Insert Table 5 about here

Insert Table 6 about here

Comparing the 3PL CAT and the nominal CAT for those Θ estimates which were convergent in both CATs (jointly convergent), one sees that the estimates from each CAT were highly related with another, r = 0.85, although they were significantly different from one another ($t = -6.6$, $df = 999$, $p = 0.0001$; see Table 6). Furthermore, for these jointly convergent cases Tables 7 and 8 show that each CAT demonstrated a great deal of association with the true Θs, r = 0.91 for the 3PL CAT and 0.93 for the nominal CAT. Further, the 3PL CAT Θ estimates were not significantly different from the true Θs, although the nominal CAT Θ estimates were significantly different from the true Θs.

Insert Table 7 about here

Insert Table 8 about here

Analysis of the CATs' Θ estimates for true Θs above and below 0 are presented in Tables 9 - 18. As can be seen from Table 9 for the 3PL CAT using its convergent cases only, the average Θ estimate was significantly different from the true Θs' average, although the two were significantly related (r = 0.89). This same pattern occurred for the nominal CAT for its convergent cases (Table 10) and also when the jointly convergent cases were used (Tables 11 and 12). The 3PL and nominal CATs' Θ estimates for the jointly convergent cases were only moderatedly related (r = 0.62) and were significantly different from one another

(Table 13). Further, the association between the 3PL CAT's Θ estimates and their true values wa significantly greater than that of the nominal CAT's ($t = 2.33$, $p = 0.01$, $df=263$).

Insert Table 9 about here

Insert Table 10 about here

Similar results were obtained for the true Θs below zero for the 3PL and nominal CATs using their respective convergent cases (Tables 14 and 15). However, when only the jointly convergent cases were used, neither the 3PL CAT nor the nominal CATs' average Θ estimate were significantly different from the average true Θ (Table 16 and Table 17); both CATs' Θ estimates were significantly related to the true Θs (3PL CAT : $r = 0.84$; nominal CAT : $r = 0.95$). Further, from Table 18 it can be seen that the Θ estimates from the two CATs were not significantly different from each other and were highly related to one another ($t = 0.34$, $p = 0.74$, $df = 265$; $r = 0.78$). A significance test for the correlations of each CAT's Θ estimates with the true Θs showed that the nominal CAT was significantly more strongly related with the true Θs than was the 3PL CAT ($t = -13.68$, $p = 0.0005$, $df = 442$).

Insert Table 11 about here

Insert Table 12 about here

In short, on the basis of the Θ estimates/true Θs' correlations as well as from the similarity of the Θ estimates' means and the true Θs' mean, both the 3PL CAT and the nominal CAT performed better for true Θs below 0 than for those above 0.

Insert Table 13 about here

Table 19 shows descriptive statistics on the standard error of the estimate of Θ (SEE) for Sim-3PL and Sim-Nom. In all cases except one, the average SEE and the mode SEE for the nominal CAT was smaller than that of the 3PL CAT.

Insert Table 14 about here

Figures 2 and 3 show plots of the SEE against the true Θs for the jointly convergent cases of the 3PL CAT and the nominal CAT, respectively. As can be seen from both these figures, the smallest SEEs are to be found around the peak of the total test information for each CAT's item pool. As one proceeds away from each pool's respective points of maximum information, the SEEs increase. In general, it appears that even with extreme cases, the nominal CAT has lower SEEs than does the 3PL CAT. A t-test showed that the nominal CAT's SEEs were significantly lower than those of the 3PL CAT ($t = 40.5$, $p = 0.0001$).

Insert Table 15 about here

Figures 4 and 5 show plots of the differences between the true Θs and each CAT's estimated Θs for the jointly convergent cases. From Figure 4 it can be seen that the plot of these differences for the 3PL CAT is relatively symmetrically distributed about a difference of 0, regardless of the true Θ value. In contrast, Figure 5 shows a parabolic pattern for the differences computed for the nominal CAT. As can be seen from this figure, the nominal CAT tends to underestimate extreme true Θs and overestimate true Θs (i.e., in absolute terms) in the range -1.25 to 1.25 (approximately). The overestimation of Θ is believed to be a result of administering items which although they met the minimum information criterion contained little information and resulted in a degradation of the Θ estimate. For instance, suppose five items were administered which had information values of 1.7, 1.[illegible], 1.8, 1.3, and 1.1 for a Θ estimate = 0.5. However, with a minimum information cutoff set at 0.20, another, say, ten items with information values in the 0.20 to 0.30 would be administered. Because these additional items contain very little information about the Θ in question they would, in all probability, introduce degradation in the ability estimate. Therefore, it appears that the minimum information cutoff for item selection may have been set to low.

Insert Table 16 about here

Analysis of the true Θs for the nonconvergent nominal CAT cases (Figure 6) showed that the majority (93.8% of 276 cases) of these cases fell above 0.25 and 56.5% occurred above 0.75. This corresponds with the above mentioned analyses of Θ estimates and their true values when the true Θs where above and below 0. In short, the nominal CAT experience a great deal of trouble estimating Θs in the positive true Θ range and virtually no trouble estimating negative true Θs. The average and mode number of items administered for these nonconvergent cases was 3.

Insert Table 17 about here

Given the mode number of items administered for the nominal CAT, the nonconvergent cases can be attributed to premature termination of the CAT program. This premature termination is believed to be a result of two factors. The first factor is a lack of appropriate items to be administered and is a reflection of the low information content of each item pool. A second factor potentially responsible for premature termination may be the algorithm used for providing a Θ estimate when MLE was not possible. The method used was to calculate for all Θs in a given Θ range the probability of responding in the category the examinee chose. The new Θ estimate for the examinee was the Θ corresponding to the largest probability. This method was used until it was possible to perform MLE, i.e., the examinee responded in a different category.

Insert Table 18 about here

Although in the design of the nominal CAT alternative techniques for dealing with this special case were considered, it was felt that the above mentioned technique used the potential of the nominal model to a fuller extent than the other methods considered. Given the nonconvergence problem this decision may not have been the best.

Insert Table 19 about here

To summarize, both CATs' Θ estimates were found to be highly related to the true Θs and to one another. It was found that the 3PL CAT had no systematic tendency to either under- or overestimate the true Θs. In contrast, the nominal CAT tended to underestimate extreme true Θs and overestimate true Θs in the range 0.5 to 1.25. In general, both CATs performed better for true Θs below 0 than for those above 0. For negative true Θs and using either the CAT's corresponding convergent cases or the jointly convergent cases, the nominal CAT had a stronger association with the true Θs than did the 3PL CAT. Further, for true values less than 0 both CATs' mean Θ estimates were not significantly different from either each other or from the true values' mean. Analysis of the standard errors of the Θ estimates for each CAT showed that on the average the nominal CAT had lower SEEs than the 3PL CAT.

Nonconvergence of Θ estimates was a far greater problem for the nominal CAT than for the 3PL CAT. Analysis of the nonconvergent cases for the nominal CAT revealed that the nominal CAT had difficulty with positive true Θs and virtually no problem estimating negative true Θs. The nonconvergent cases for both CATs were attributed to premature termination of the CAT as a result of the low information content of the items in the pool as well as the algorithm used for providing a Θ estimate when no MLE was possible. Given the total test information for each simulated data's item pool, it was expected that nonconvergence would be a greater problem for the 3PL CAT than for the nominal CAT. Future research may identify alternative explanation(s) or solution(s) for the nonconvergent problem with the nominal CAT.

EDUCATIONAL IMPORTANCE

A nominal model-based CAT would allow the use of both polychotomous and dichotomously scored items. These polychotomous items represent a new domain of items which may be used in adaptive testing environments. Further, new and innovative item types may be developed and applied through a nominal-based CAT. These new item types may provide a means of merging computer assisted instruction with computerized adaptive testing.

Although the nominal model does not require scoring the examinee's response, there still may exist the tendency to guess. This tendency to guess will most likely be exhibited when inappropriately difficult items are administered to the examinee. Under the nominal CAT procedure, examinees who select different incorrect alternatives will receive different ability estimates and therefore will most likely be administered different next items. This is not possible with a CAT based on a dichotomously scored model. By providing more accurate ability estimates for persons who respond incorrectly, items may be selected which are more appropropriate for the examinee's current ability estimate. This will probably result in a reduction of the likelihood of guessing as well as less anxiety for the examinee.

Table 1 : Descriptive Statistics of data sets

Data Set	N	# Items	# Opt.	Mean No. Correct	S.D.	Alpha
Original 1093.	1093	49	4	27.1	7.4	0.85
Sim-3PL	1000	98	4	56.6	14.7	0.92
Sim-Nom	1000	98	4	59.8	15.4	0.92

Table 2 : Factor Analyses

						Factor 1 Var. Acctd	
Data Set	N	# Items	# Opt.	Factor 1	Factor 2	Total	Common
Original 1093	1093	49	4	7.23	1.31	14.8%	75.8%
Sim-3PL	1000	98	4	11.76	2.06	12.0%	85.1%
Sim-Nom	1000	98	4	22.42	6.17	22.9%	70.5%

Table 3 : Test Information for item pools - Simulated data sets, 98 items for the Θ range -4.0 to 4.0.

Data Set	Θ Peaked	S.D.	Min	Max	Avg Info/Item	Total Info
Sim-3PL	1.8	7.46	1.52	35.04	0.10	806.99
Sim-Nom	0.4	13.63	2.30	41.30	0.20	1637.2

Table 4 : Θ Analysis from CAT run on simulated data, 3PL convergent cases only, N = 976

Data Set	Θ Mean	S.D	r	t	p	Min	Max
Sim-3PL	0.12	1.20	0.94	-4.72	0.0001	-3.94	3.80
True Θs	0.06	0.99				-2.86	3.23

Table 5 : Θ Analysis from CAT run on simulated data, Nominal convergent cases only, N = 724

Data Set	Θ Mean	S.D	r	t	p	Min	Max
Sim-Nom	-0.13	1.31	0.94	-7.47	0.0001	-3.98	2.63
True Θs	0.29	0.87				-2.86	2.70

Table 6 : Θ Analysis from CAT run on simulated data, comparison of both models jointly convergent cases only, N = 711

Data Set	Θ Mean	S.D	r	t	p	Min	Max
Sim-3PL	-0.25	1.03	0.85	-6.66	0.0001	-3.94	3.80
Sim-Nom	-0.08	1.27				-3.98	2.63

Table 7 : Θ Analysis from CAT run on simulated data, 3PL & Nominal jointly convergent cases only, N = 711

Data Set	Θ Mean	S.D	r	t	p	Min	Max
Sim-3PL	-0.25	1.03	0.91	-0.85	0.3930	-3.94	3.80
True Θs	-0.27	0.84				-2.86	2.70

Table 8 : Θ Analysis from CAT run on simulated data, 3PL & Nominal jointly convergent cases only, N = 711

Data Set	Θ Mean	S.D	r	t	p	Min	Max
Sim-Nom	-0.08	1.27	0.93	-8.48	0.0001	-3.98	2.63
True Θs	-0.27	0.84				-2.86	2.70

Table 9 : Θ Analysis from CAT run on simulated data, 3PL convergent cases only; true Θs GE 0.0, N = 521

Data Set	Θ Mean	S.D	r	t	p	Min	Max
Simulated 3PL	0.94	0.81	0.89	-9.60	0.0001	-0.92	3.80
True Θs	0.78	0.62				0.0	3.23

Table 10 : Θ Analysis from CAT run on simulated data, Nominal convergent cases only; true Θs GE 0.0, N = 267

Data Set	Θ Mean	S.D	r	t	p	Min	Max
Simulated Nominal	1.11	0.48	0.71	-22.25	0.0001	0.24	2.63
True Θs	0.55	0.57				0.0	2.70

Table 11 : Θ Analysis from CAT run on simulated data, 3PL & Nominal jointly convergent cases only; true Θs GE 0.0, N = 266

Data Set	Θ Mean	S.D	r	t	p	Min	Max
Simulated 3PL	0.65	0.75	0.86	-4.11	0.0001	-0.92	3.80
True Θs	0.55	0.56				0.0	2.70

Table 12 : Θ Analysis from CAT run on simulated data, 3PL & Nominal jointly convergent cases only; for true Θs GE 0.0, N = 266

Data Set	Θ Mean	S.D	r	t	p	Min	Max
Simulated Nominal	1.10	0.48	0.71	-22.52	0.0001	0.24	2.63
True Θs	0.55	0.56				0.0	2.70

Table 13 : Θ Analysis from CAT run on simulated data, comparison of both models jointly convergent cases only; true Θs GE 0.0, N = 266

Data Set	Θ Mean	S.D	r	t	p	Min	Max
Simulated 3PL	0.65	0.75	0.62	-12.75	0.0001	-0.92	3.80
Simulated Nominal	1.10	0.48				0.24	2.63

Table 14 : Θ Analysis from CAT run on simulated data, 3PL convergent cases only; true Θs LT 0.0, N = 455

Data Set	Θ Mean	S.D	r	t	p	Min	Max
Simulated 3PL	-0.82	0.80	0.84	2.17	0.03	-3.94	0.95
True Θs	-0.78	0.59				-2.86	-0.002

Table 15 : Θ Analysis from CAT run on simulated data, Nominal convergent cases only; true Θs LT 0.0, N = 457

Data Set	Θ Mean	S.D	r	t	p	Min	Max
Simulated Nominal	-0.85	1.07	0.95	2.75	0.006	-3.98	0.622
True Θs	-0.78	0.59				-2.86	-0.002

Table 16 : Θ Analysis from CAT run on simulated data, 3PL & Nominal jointly convergent cases only; true Θs LT 0.0, N = 445

Data Set	Θ Mean	S.D	r	t	p	Min	Max
Simulated 3PL	-0.79	0.77	0.83	1.77	0.08	-3.94	0.95
True Θs	-0.75	0.56				-2.86	-0.002

Table 17 : Θ Analysis from CAT run on simulated data, 3PL & Nominal jointly convergent cases only; true Θs LT 0.0, N = 445

Data Set	Θ Mean	S.D	r	t	p	Min	Max
Simulated Nominal	-0.80	1.03	0.95	1.90	0.6	-3.98	0.62
True Θs	-0.75	0.56				-2.86	-0.002

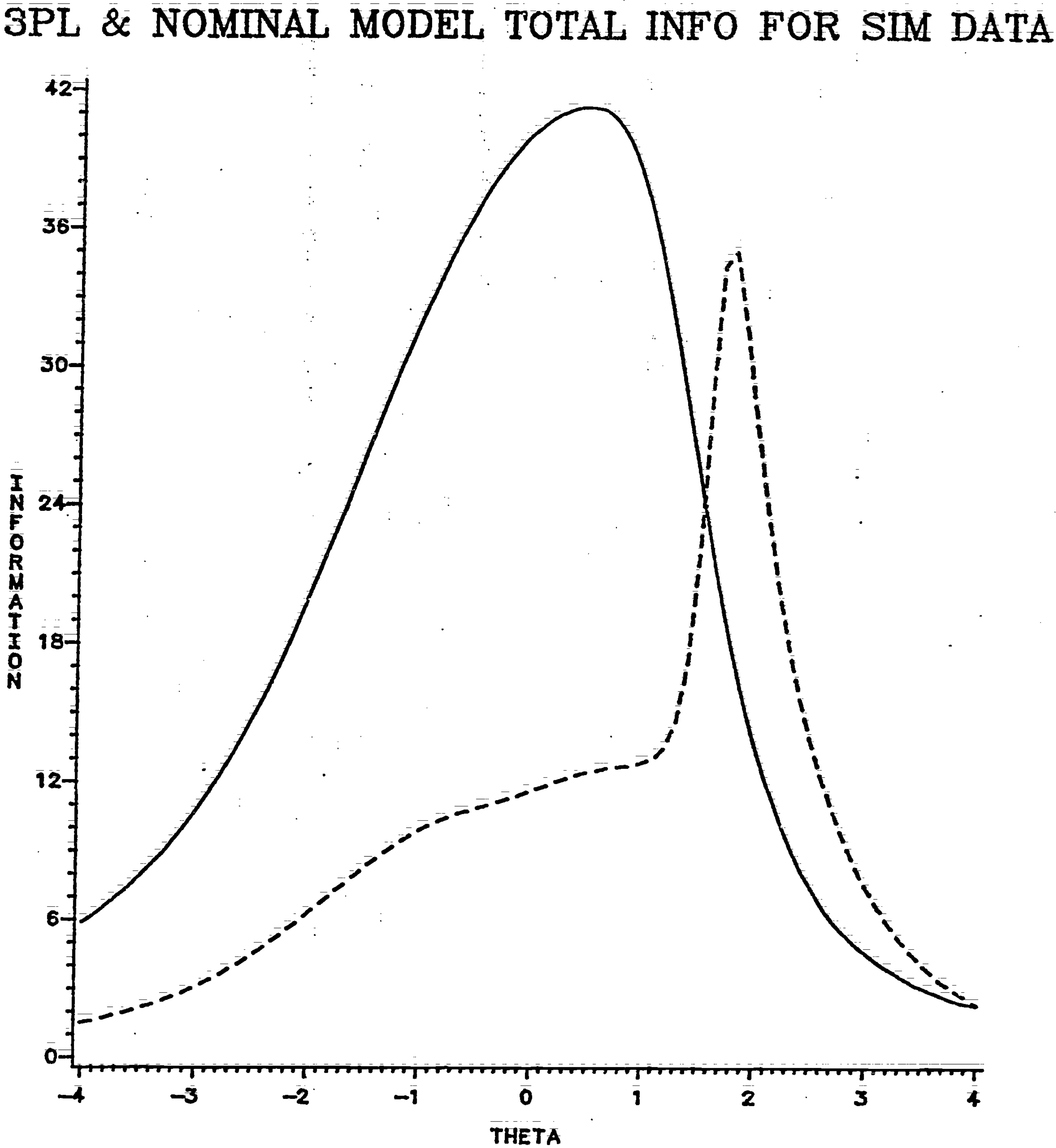
Table 18 : Θ Analysis from CAT run on simulated data, comparison of both models jointly convergent cases only; true Θs LT 0.0, N = 445

Data Set	Θ Mean	S.D	r	t	p	Min	Max
Simulated 3PL	-0.79	0.77	0.78	0.34	0.74	-3.94	0.95
Simulated Nominal	-0.80	1.03				-3.98	0.62

Table 19 : Descriptive Statistics : Standard Errors from Θ estimates for CAT runs with simulated data sets.

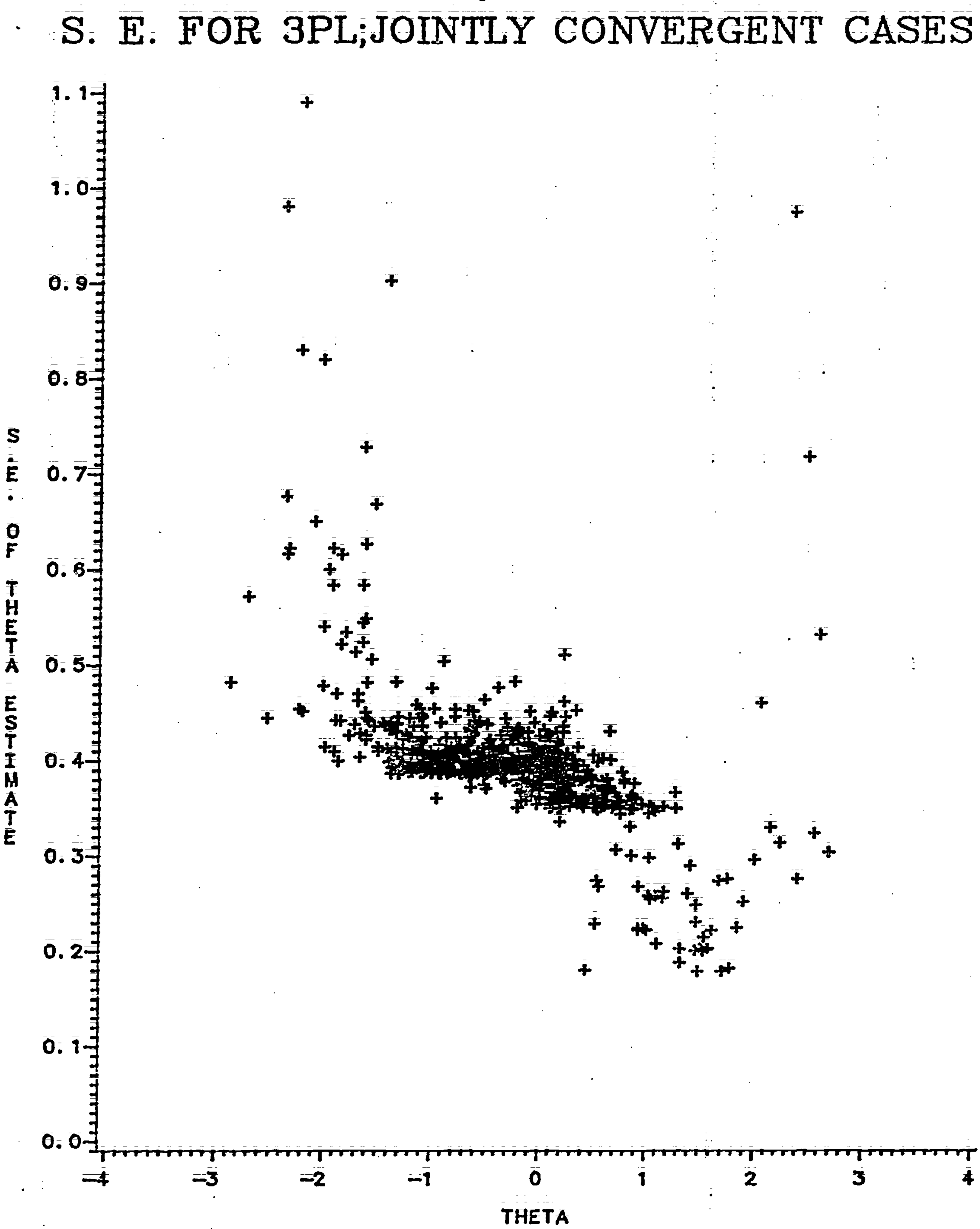
Model for CAT	N	Mean	Mode	S.D.	Min	Max
3PL	976	0.38	0.40	0.09	0.18	1.09 (convergent cases)
Nominal	724	0.24	0.19	0.11	0.19	1.05 (convergent cases)
3PL	711	0.40	0.40	0.08	0.18	1.09 (joint convergence)
Nominal	711	0.23	0.19	0.10	0.19	1.05 (joint convergence)

Figure 1



DASHED LINE : 3PL MODEL
SOLID LINE : NOMINAL MODEL

Figure 2



S. E. FOR 3PL;JOINTLY CONVERGENT CASES
S.E. OF THETA ESTIMATE
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-4
-3
-2
-1
0
1
2
3
4
THETA

Figure 3

S.E. FOR NOMINAL;JOINTLY CONVERGENT CASES

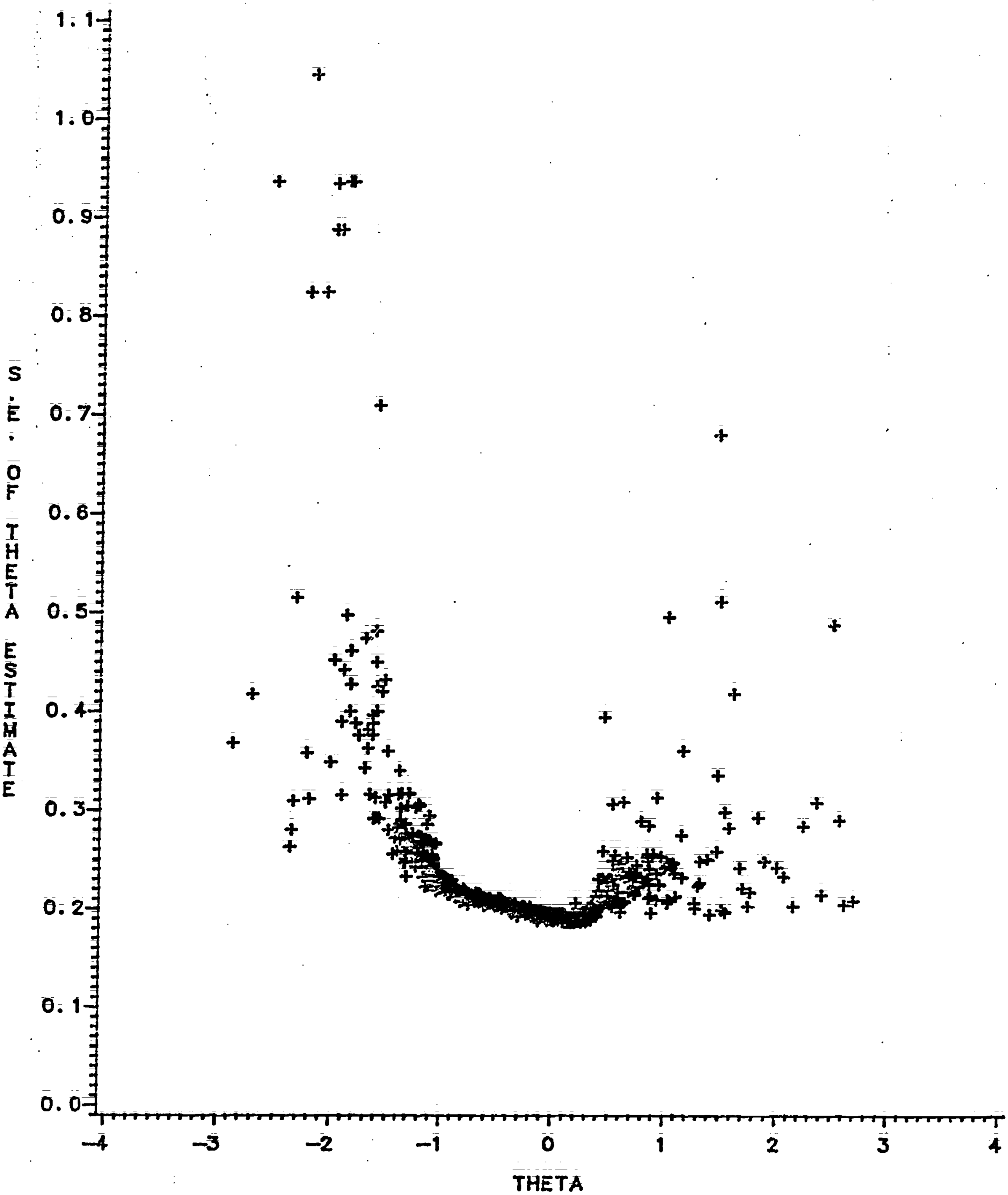


Figure 4

TRUE THETA – 3PL EST.; JOINTLY CONVERGENT

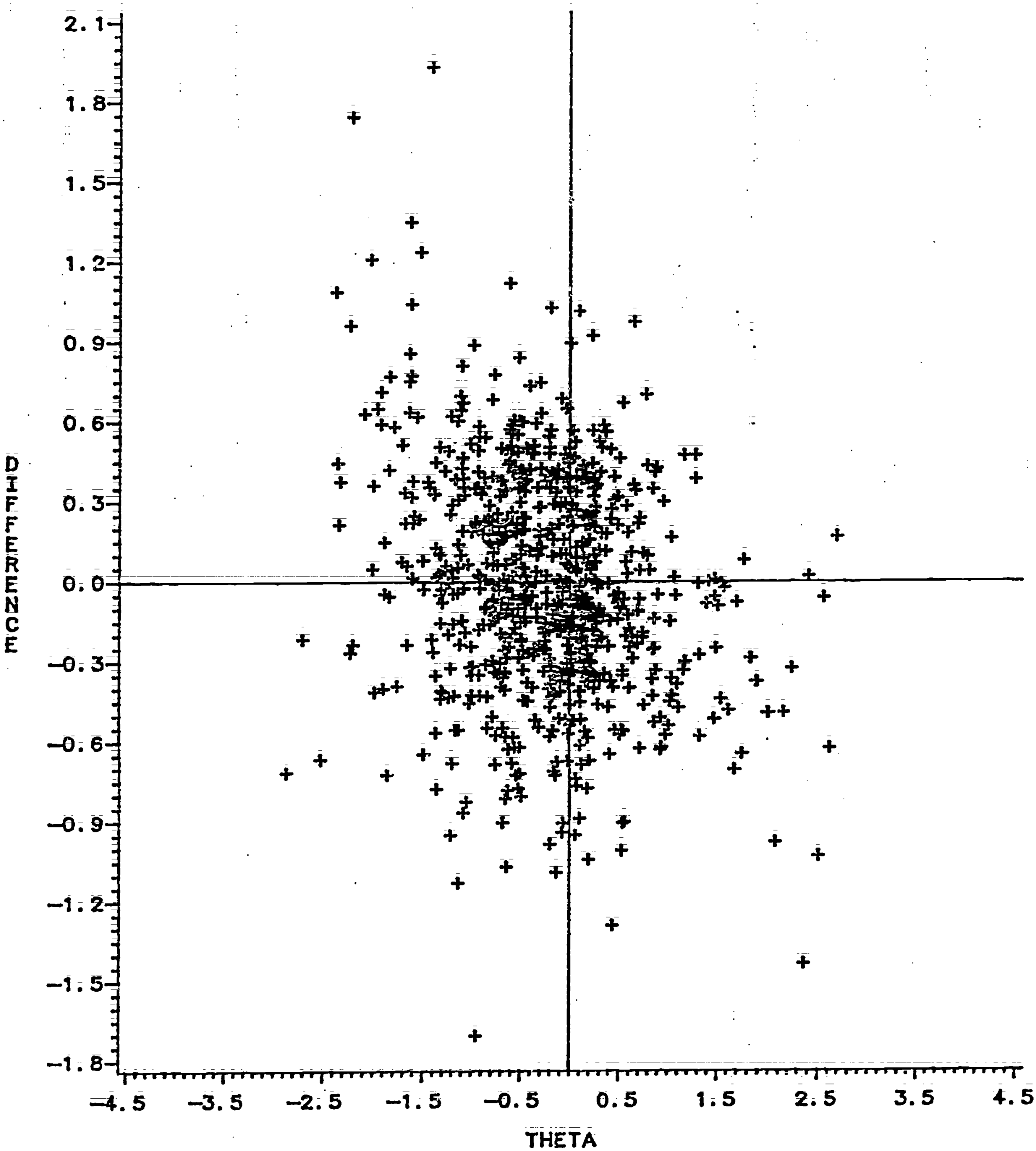


Figure 5

TRUE THETA – NOMINAL EST.; JOINTLY CONVERGENT

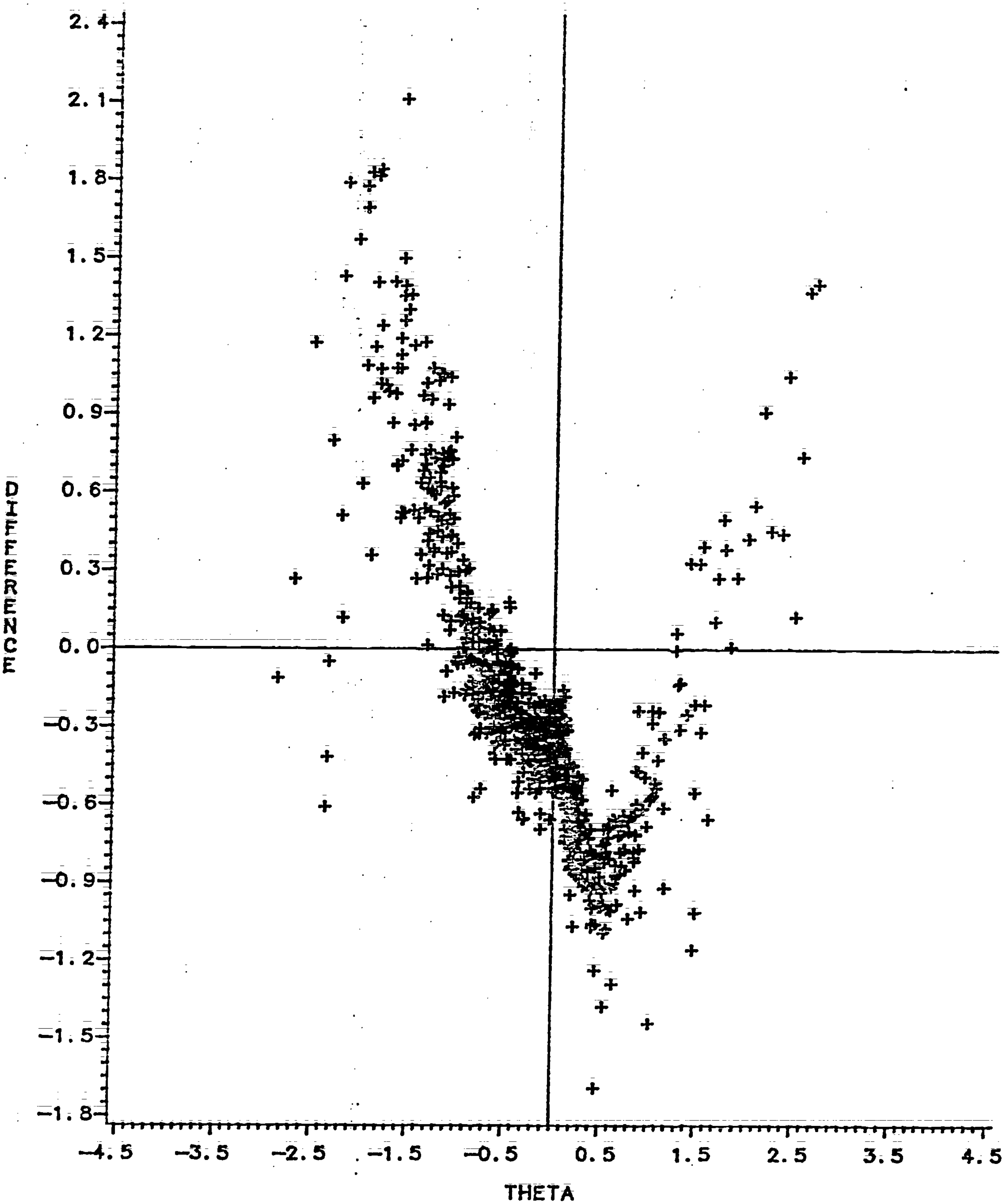
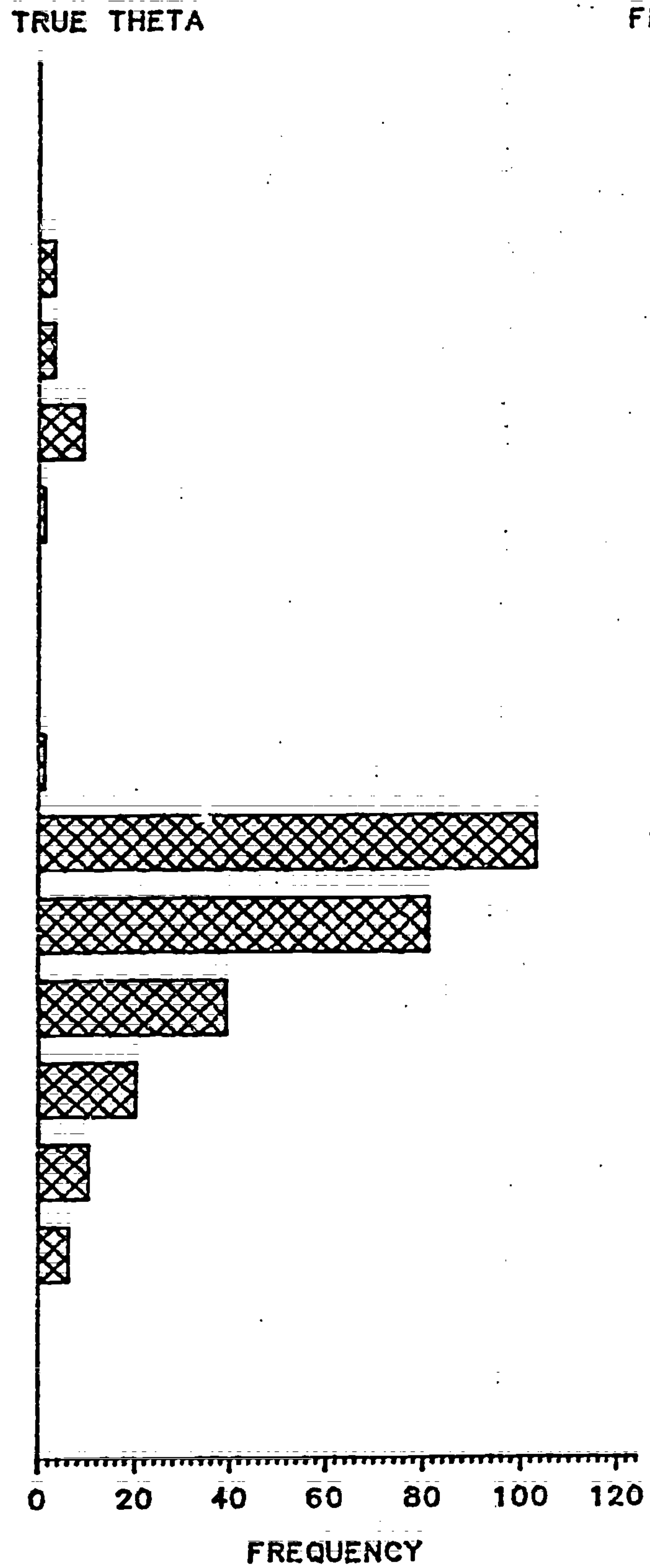


Figure 6

FREQ OF TRUE THETA FOR NOMINAL NONCONVERGENT CASES



MIDPOINT	FREQ	CUM. FREQ	PERCENT	CUM. PERCENT
-4.0	0	0	0.00	0.00
-3.5	0	0	0.00	0.00
-3.0	3	3	1.09	1.09
-2.5	3	6	1.09	2.17
-2.0	9	15	3.26	5.43
-1.5	1	16	0.36	5.80
-1.0	0	16	0.00	5.80
-0.5	0	16	0.00	5.80
0.0	1	17	0.36	6.16
0.5	103	120	37.32	43.48
1.0	81	201	29.35	72.83
1.5	39	240	14.13	86.96
2.0	20	260	7.25	94.20
2.5	10	270	3.62	97.83
3.0	6	276	2.17	100.00
3.5	0	276	0.00	100.00
4.0	0	276	0.00	100.00

REFERENCES

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

Levine, M., & Drasgow, F. (1983). The relation between incorrect option choice and estimated ability. Educational and Psychological Measurement, 43, 675-685.

Thissen, D.M. (1976). Information in wrong responses to Raven's Progressive Matrices. Journal of Educational Measurement, 13, 201-214.

Thissen, D.M. (1986). MULTILOG-User's Guide (Version 5). Scientific Software, Inc., Mooresville, IN.