Universidad de Costa Rica

Instituto de Investigación para el Mejoramiento de la Educación Costarricense

Maestría en Evaluación Educativa

IIMEC

Análisis de la Calidad Técnica Pruebas de Bachillerato Fórmula 11 Español

Elaborado por:

Marta E. Picado Mesén

Karla Salguero Moya

Diciembre 1998

PRESENTACIÓN

Este informe es presentado por el IIMEC al Ministerio de Educación Pública y al Departamento de Control de Calidad de ese Ministerio, como parte del análisis de calidad técnica de la prueba de Bachillerato. Corresponde a un análisis de psicométrico de la fórmula número 11 de la prueba de Español, aplicada en 1997.

El análisis se realizó haciendo uso de la Teoría Clásica del Test, pero para efecto de este trabajo, además se aplicó un modelo de Teoría de Respuesta a los Itemes para aportar acerca de las propiedades psicométricas de los ítemes que componen la prueba.

Este estudio es de naturaleza preliminar y se considera necesario continuar los análisis a mayor profundidad, en coordinación con el IIMEC, el Ministerio de Educación y la Maestría de Evaluación Educativa de la Universidad de Costa Rica.

OBJETIVOS

1. Realizar un análisis de factores para detectar evidencia de validez de constructo en la prueba de Español de aplicó el MEP a los estudiantes de Undécimo año en 1997.
2. Analizar bajo la Teoría de Respuesta al Ítem, aspectos de calidad técnica de la mencionada prueba que incluya técnicas para determinar la confiabilidad de dicha prueba, así como las propiedades psicométricas individuales para cada uno de los ítemes que la componen.
3. Detectar por medio de la Teoría de Respuesta al item, posibles sesgos por categorías de población, comparando estudiantes provenientes de colegios públicos contra estudiantes provenientes de colegios privados y comparando estudiantes de colegios rurales frente a estudiantes de colegios urbanos.

II

METODOLOGÍA

El objeto evaluado

Los resultados obtenidos de la fórmula 11 de la prueba de Español que aplicó el Ministerio de Educación Pública a la población estudiantes de undécimo año del país, como prueba de bachillerato en 1997.

Descripción del instrumento

La fórmula 11 de la prueba de español está constituida por dos partes.

- La primera, responde a la producción textual, con un instructivo para el estudiante con 14 indicaciones, se destacan cuatro temas para realizar el trabajo de producción textual, además cuenta con tres hojas para que se trabaje en borrador y otras tres páginas para escribir la composición que se calificará, y por último, la página con la tabla para calificar la "producción Textual", con los criterios y los puntajes de esta parte.
- La segunda parte, corresponde a los ítemes de selección, hay una hoja de instrucciones para el estudiante, y luego están 56 ítemes de un punto cada uno, de estos se excluyó del análisis el número 55, por haber sido eliminado por el Ministerio de Educación Pública.

Cada ítem de selección cuenta con cuatro opciones de respuestas, una respuesta clave y tres distractores.

Descripción del procedimiento metodológico general para efectuar el análisis

Se presenta a continuación una descripción de todas las fases seguidas para efectuar el análisis con la TRI, que incluye desde la conversión de la base de datos hasta la obtención del análisis sustantivo y las conclusiones.

Etapa previa a la TRI:

Se abre la base de datos en el programa Excel, versión 5.0.

Se selecciona de la base general una muestra de la población estudiantil que respondió la prueba de español en 1997.

La muestra corresponde a los 41165 estudiantes que respondieron la fórmula 11 de la prueba de español en el año de 1997.

- Se abre la base de datos de la muestra escogida en el programa SPSS.

  Se dicotomiza el patrón de respuesta para cada una de las variables. Se asigna el valor

  "1" para las respuestas correctas y,

  "0" para las respuestas incorrectas, los espacios en blanco, valores perdidos.
- Se realizó el análisis de factores para determinar la unidimensionalidad del instrumento, optando por una rotación oblicua, al suponer asociación entre los factores con respecto al constructo.
- Se dicotomizaron los ítemes para el análisis de sesgo quedando de la siguiente manera:

Variable tipo de Colegio de Procedencia:

1= colegios públicos incluyendo los académicos diurnos.

0= colegios privados, en el que se incluyó colegios privados, semiprivados y académicos diurnos.

Variable zona del Colegio de Procedencia:

1= colegios urbanos públicos incluyendo los académicos diurnos.

0= colegios rurales, se incluyó los académicos diurnos.

Variable Sexo:

0= Hombres

1= Mujeres

- Se calculó con el paquete SPSS la confiabilidad del instrumento utilizando el alfa de Cronbach, el cual se define como una medida para determinar la confiabilidad o grado de precisión de una prueba. El instrumento debe poseer un coeficiente de confiabilidad de .7 como mínimo para fines de investigación y de .9 para toma de decisiones.

- Se procedió a dividir la base de datos en seis archivos, dos para las categorías tipo de Colegios, dos para las categorías zona de Colegios y dos para el sexo, se incluyen los "valores perdidos".

Etapa de trabajo con la TRI:

- Se transfieren los datos al paquete BILOG, incorporando los aspectos necesarios para que pueda funcionar el análisis: conversión del ancho de columna desde el SPSS.
- Se crea el archivo de instrucciones. En este caso se optó por un "modelo de 3 parámetros", por ser este el de mejor se ajuste en el procedimiento.
- Se abre el archivo de datos y se corren las tres fases del análisis de la prueba.
- Se interpreta el modelo de la TRI y se comparan los resultados obtenidos en ambas teorías.
- Se realizó el análisis de la función de información de todos los ítemes.
- Se realizó el análisis individual de los 55 ítemes, de la función de información, la discriminación, la dificultad y el azar presentado.
- Se realizaron las conclusiones en los niveles de aprendizaje y sustantiva.

Etapa de trabajo con el análisis de Sesgo

- Se transfieren los datos al paquete BILOG desde el SPSS.
- Se crean dos archivos de instrucciones. Uno para colegios públicos y otro para colegios privados.
- Se mantuvo el modelo de 3 parámetros.
- A partir del archivo de las instrucciones se corren las fases del análisis de la prueba cognoscitiva.
- Posteriormente, se abre el archivo con extensión .plt para obtener los gráficos en cada categoría (Tipo, Zona y Sexo).
- Se obtienen los índices de dificultad, de discriminación y de azar de los ítemes

- Se procede a obtener gráficos de dispersión de las categorías de análisis con el uso del paquete SPSS.
- Se calcula una correlación simple (bivariable) en cada una de las categorías con el SPSS.
- Ambos criterios se utilizan para seleccionar de manera más objetiva, los ítemes que presentan sesgo.
- Consultando la prueba, se revisan aquellos ítemes que presentan probabilidad de sesgo.
- A partir de lo anterior, construimos las respectivas conclusiones.

Análisis Clásico de la Prueba

Análisis Exploratorio de Factores:

De acuerdo con Nunnally (1967 p. 546), el análisis factorial es una serie amplia de procedimientos diseñados para lograr una variedad de fines, entre los que destacan: el agrupamiento de variables, la determinación de cuáles variables y qué tan fuertemente pertenecen a cuál grupo, cuántas dimensiones se necesitan para explicar las relaciones entre las variables, entre otras.

Particularmente, el análisis de factores, como señala Nunnally, es un instrumento básico para explicar constructos, de manera que se puede establecer cómo se agrupan los ítemes de una prueba de acuerdo con las dimensiones que subyacen en ella.

En el caso del presente trabajo, se aplicó el análisis de factores en forma exploratoria dado que no hay un marco teórico que permita identificar claramente dimensiones subyacentes del constructo principal.

El análisis de factores constituye una técnica estadística multivariable que se emplea para evidenciar la validez de constructo, es decir verificar la unidimensional de la prueba.

Dadas las circunstancias de la inexistencia del marco teórico que definieran claramente los factores, nos decidimos por el tipo de análisis exploratorio. Para poder determinar cuántos factores están presentes en la prueba, se construyó el gráfico

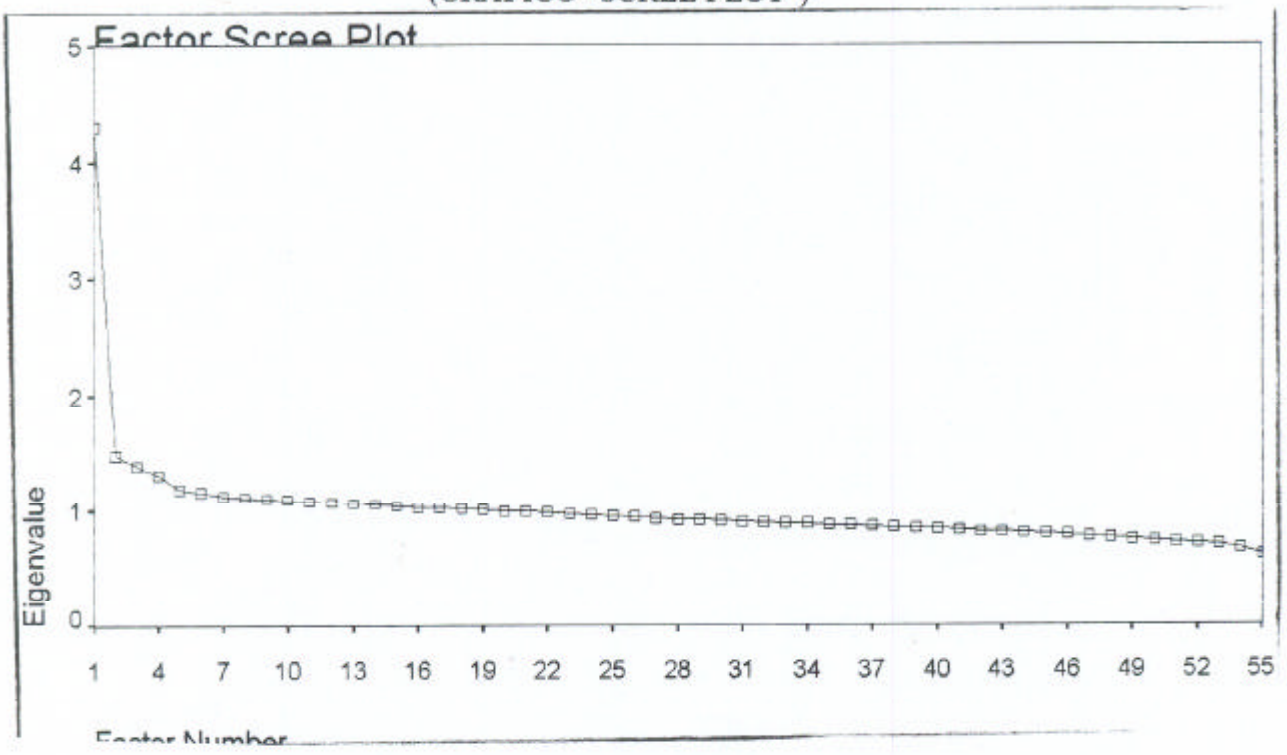
denominado "Scree Plot" que compara el número de factores contra el valor característico asociado a cada factor, es una medida de su importancia relativa. Cuanto mayor sea el valor característico, mayor será el porcentaje de la varianza de las variables (ítemes) explicado por ese factor. Típicamente el factor 1 tiene mayor valor característico, le sigue en importancia el factor 2 y así sucesivamente.

Para la determinación del número de factores se observa en dónde se estabiliza la solución en el "Scree Plot". En el gráfico esta estabilización se visualiza en el punto en el cual se dibuja una especie de "codo" o esquina. Los factores que se definen después de ese "codo" no tendrán mayor importancia relativa, en términos de varianza explicada, por tanto, no vale la pena tomarlos en cuenta para la solución final". Se determinó que la prueba tiene un sólo factor a partir del gráfico"Scree Plot", como se observa seguidamente:

GRÁFICO No 1

PRUEBA DE ESPAÑOL DE UNDECIMO AÑO

(GRÁFICO "SCREE PLOT")



Para la interpretación de los factores se toma como base la carga factorial de los ítemes, la cual debe ser como mínimo .30 para que se consideren como representativas del factor en el cual se ubican.

Como no se contó con el marco teórico que respalda esta prueba, y por ser una prueba de conocimiento de español, se determinó que es español lo que representa el factor

Factor 1. Español

Se presentan a continuación las cargas factoriales asociadas al factor 1, donde se describen estas cargas con el subrayado y negrilla en valores cercanos al 0.30. El porcentaje de varianza explicada es de 7,8%.

Apartir del análisis de los ítemes que cargan en el factor 1, se detectó que muchos de estos representan el área de Comprensión de lectura, lo cual podría permitir considerar que el factor podría representar este constucto.

TABLA 1

CARGA FACTORIAL DE LA PRUEBA DE ESPAÑOL

PARA UNDECIMO AÑO

Factor 1

Item	Factor 1	Item	Factor 1
R01	0,18409	R29	**0,45084**
R02	0,27527	R30	0,27038
R03	0,14959	R31	**0,35452**
R04	0,27120	R32	**0,34355**
R05	0,28692	R33	0,20221
R06	**0,36911**	R34	0,22948
R07	0,27748	R35	**0,35678**
R08	**0,44175**	R36	**0,42084**
R09	0,21943	R37	0,28856
R10	**0,34132**	R38	0,27346
R11	0,19936	R39	0,16899
R12	0,11980	R40	0,21323
R13	0,09297	R41	**0,41145**
R14	0,27016	R42	0,16865
R15	**0,40764**	R43	0,15863
R16	**0,31710**	R44	0,25970
R17	0,28365	R45	0,12676
R18	**0,32038**	R46	**0,31465**
R19	0,29197	R47	0,14260
R20	0,27648	R48	0,20521
R21	**0,30453**	R49	0,07871
R22	0,21306	R50	**0,35681**
R23	0,27538	R51	**0,30651**
R24	0,09025	R52	0,21793
R25	0,21401	R53	**0,40357**
R26	0,23291	R54	**0,46744**
R27	0,13986	R56	0,21521
R28	0,15787		

TABLA 2

CARGA FACTORIAL DE LOS ÍTEMES,

SEGÚN FACTOR

(0,30)

Itemes	06	08	10	15	16	18	21	29	31	32	35	36	41	50	51	53	54

De las tablas anterior se desprende que el factor 1 está representado por un 30,90% o sea en solo 17 de los 55 ítemes de la prueba..

Los ítemes que no alcanzaron la carga factorial mínima (.30) son los siguientes: 1, 2, 3, 4, 5, 7, 9, 11, 12, 13, 14, 17, 19, 20, 22, 23, 24, 25, 26, 27, 28, 30, 33, 34, 37, 38, 39, 40, 42,43, 44, 45, 46, 47, 48, 49, 52 y 56, los cuales representan el 69,09% del total de ítemes de la prueba o sea 38 de 55 ítemes analizados.

Alfa de Cronbach:

El Alfa de Cronbach es una medida para determinar la confiabilidad o grado de precisión de una prueba. El instrumento debe poseer un coeficiente de confiabilidad de .7 como mínimo para fines de investigación y de .9 para toma de decisiones según Nunnally (1995, p.282)."El coeficiente alfa por lo común proporciona una buena estimación de la confiabilidad porque el muestreo del contenido es la principal fuente de error de medición para constructos estáticos y demás debido a que es sensible al "muestreo" de factores así como el contenido del reativo"

En el caso de la fórmula 11 de la prueba de español aplicada en 1997 se obtuvo un alfa de .7539 lo que indica que la prueba podria mejorarse.

Discriminación:

El índice de discriminación es la capacidad que tiene cada ítem para poder distinguir entre sujetos con puntajes altos y bajos. De acuerdo con la teoría, los ítemes deben tener un índice de discriminación como mínimo de .30.

A continuación se presentan los resultados del SPSS, en relación con la confiabilidad y la discriminación de la prueba en estudio:

TABLA 3

RELIABILITY ANALYSIS - SCALE (ALPHA)

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Alpha if Item Deleted
R01	37.5975	38.1715	.1454	.7528
R02	37.5231	37.8076	.2082	.7501
R03	37.1341	39.0086	.1129	.7530
R04	37.6544	37.7353	.2193	.7497
R05	37.4790	37.7621	.2210	.7496
R06	37.4622	37.3763	.2902	.7467
R07	37.1773	38.5134	.2142	.7506
R08	37.4435	37.1063	**.3421**	.7446
R09	37.7604	38.0292	.1835	.7511
R10	37.4783	37.4626	.2723	.7474
R11	37.6839	38.0680	.1662	.7519
R12	37.7908	38.5767	.0930	.7546
R13	37.1566	39.0491	.0708	.7538
R14	37.4509	37.8196	.2158	.7498
R15	37.4123	37.3024	**.3167**	.7457
R16	37.3142	37.8609	.2528	.7485
R17	37.1655	38.5983	.2061	.7509
R18	37.3586	37.8102	.2427	.7488
R19	37.2900	38.0658	.2243	.7496
R20	37.3924	37.9362	.2085	.7501
R21	37.3857	37.7927	.2365	.7490
R22	37.2593	38.4411	.1613	.7517
R23	37.3610	37.9855	.2091	.7501
R24	37.3205	38.8051	.0653	.7552
R25	37.2979	38.3680	.1585	.7519
R26	37.2200	38.4710	.1802	.7512
R27	37.1864	38.8700	.1020	.7532
R28	37.3332	38.4741	.1256	.7531
R29	37.2818	37.5475	**.3391**	.7458
R30	37.1578	38.6608	.1988	.7512
R31	37.1948	38.2365	.2679	.7490
R32	37.3500	37.7731	.2532	.7484
R33	37.2943	38.4088	.1517	.7521
R34	37.5613	37.9963	.1747	.7516
R35	37.6469	37.3926	.2762	.7472
R36	37.5625	37.0725	**.3288**	.7449
R37	37.4234	37.8248	.2206	.7496
R38	37.3982	37.8932	.2147	.7498
R39	37.1478	38.8985	.1331	.7526
R40	37.4608	38.1218	.1624	.7520
R41	37.2708	37.7057	**.3142**	.7467
R42	37.4956	38.3235	.1246	.7536
R43	37.5217	38.3436	.1194	.7539
R44	37.7045	37.8829	.1996	.7505
R45	37.5930	38.4725	.0963	.7549

(CONTINUACION TABLA 3)

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Alpha if Item Deleted
R46	37.4864	37.6558	.2380	.7489
R47	37.3783	38.4920	.1117	.7538
R48	37.4059	38.2273	.1534	.7523
R49	37.3195	38.7880	.0688	.7551
R50	37.1785	38.3280	.2672	.7493
R51	37.5500	37.6957	.2250	.7494
R52	37.4553	38.1425	.1596	.7521
R53	37.3850	37.4192	**.3052**	.7463
R54	37.2569	37.5081	**.3578**	.7455
R56	37.5586	38.0201	.1708	.7517

Reliability Coefficients

N of Cases = 4169.0 N of Items = 55

Alpha = .7539

El indice de discriminación de cada uno de los ítemes por medio de la Teoria Clásica se lee en la columna "Corrected item-test correlation", donde se encuentra que los ítemes 8, 15, 29, 36, 41, 53, 54 cumplen con el requisito del estandar establecido. Los ítemes restantes deberian ser revisados.

Indice de dificultad

A continuación se presentan las salidas obtenidas en la primera fase del Bilog, que brindan el índice de dificultad de la prueba bajo la teoria clásica:

TABLA 4

CLASSICAL ITEM STATISTIC FOR SUBTEST COGNITIV

ITEM	NAME	NUMBER TRIED	NUMBER RIGHT	PERCENT	LOGIT/1.7	ITEM*TEST PEARSON	CORRELATION BISERIAL
1	0001	4169.0	2074.0	0.497	-0.01	0.145	0.182
2	0002	4169.0	2384.0	0.572	0.17	0.208	0.262
3	0003	4169.0	4006.0	0.961	1.88	0.113	0.257
4	0004	4169.0	1837.0	0.441	-0.14	0.219	0.276
5	0005	4169.0	2568.0	0.616	0.28	0.221	0.281
6	0006	4169.0	2638.0	0.633	0.32	0.290	0.371
7	0007	4169.0	3826.0	0.918	1.42	0.214	0.388
8	0008	4169.0	2716.0	0.651	0.37	0.342	0.441
9	0009	4169.0	1395.0	0.335	-0.40	0.183	0.238
10	0010	4169.0	2571.0	0.617	0.28	0.272	0.347
11	0011	4169.0	1714.0	0.411	-0.21	0.166	0.210
12	0012	4169.0	1268.0	0.304	-0.49	0.093	0.122
13	0013	4169.0	3912.0	0.938	1.60	0.071	0.140
14	0014	4169.0	2685.0	0.644	0.35	0.216	0.277
15	0015	4169.0	2846.0	0.683	0.45	0.317	0.414
16	0016	4169.0	3255.0	0.781	0.75	0.253	0.354
17	0017	4169.0	3875.0	0.929	1.52	0.206	0.391
18	0018	4169.0	3070.0	0.736	0.60	0.243	0.327
19	0019	4169.0	3356.0	0.805	0.83	0.224	0.322
20	0020	4169.0	2929.0	0.703	0.51	0.209	0.275
21	0021	4169.0	2957.0	0.709	0.52	0.237	0.313
22	0022	4169.0	3484.0	0.836	0.96	0.161	0.241
23	0023	4169.0	3060.0	0.734	0.60	0.209	0.281
24	0024	4169.0	3229.0	0.775	0.73	0.065	0.091
25	0025	4169.0	3323.0	0.797	0.80	0.158	0.226
26	0026	4169.0	3648.0	0.875	1.14	0.180	0.289
27	0027	4169.0	3788.0	0.909	1.35	0.102	0.179
28	0028	4169.0	3176.0	0.762	0.68	0.126	0.173
29	0029	4169.0	3390.0	0.813	0.87	0.339	0.492
30	0030	4169.0	3907.0	0.937	1.59	0.199	0.390
31	0031	4169.0	3753.0	0.900	1.29	0.268	0.458
32	0032	4169.0	3106.0	0.745	0.63	0.253	0.344
33	0033	4169.0	3338.0	0.801	0.82	0.152	0.217
34	0034	4169.0	2225.0	0.534	0.08	0.175	0.219
35	0035	4169.0	1868.0	0.448	-0.12	0.276	0.347
36	0036	4169.0	2220.0	0.533	0.08	0.329	0.413
37	0037	4169.0	2800.0	0.672	0.42	0.221	0.287
38	0038	4169.0	2905.0	0.697	0.49	0.215	0.282
39	0039	4169.0	3949.0	0.947	1.70	0.133	0.276
40	0040	4169.0	2644.0	0.634	0.32	0.162	0.208
41	0041	4169.0	3436.0	0.824	0.91	0.314	0.463
42	0042	4169.0	2499.0	0.599	0.24	0.125	0.158
43	0043	4169.0	2390.0	0.573	0.17	0.119	0.151
44	0044	4169.0	1628.0	0.391	-0.26	0.200	0.254
45	0045	4169.0	2089.0	0.501	0.00	0.096	0.121
46	0046	4169.0	2537.0	0.609	0.26	0.238	0.302

TABLA 4 (cont)

CLASSICAL ITEM STATISTIC FOR SUBTEST COGNITIV

ITEM	NAME	NUMBER TRIED	NUMBER RIGHT	PERCENT	LOGIT/1.7	ITEM*TEST PEARSON	CORRELATION BISERIAL
47	0047	4169.0	2988.0	0.717	0.55	0.112	0.149
48	0048	4169.0	2873.0	0.689	0.47	0.153	0.201
49	0049	4169.0	3233.0	0.775	0.73	0.069	0.096
50	0050	4169.0	3821.0	0.917	1.41	0.267	0.481
51	0051	4169.0	2272.0	0.545	0.11	0.225	0.283
52	0052	4169.0	2667.0	0.640	0.34	0.160	0.205
53	0053	4169.0	2960.0	0.710	0.53	0.305	0.405
54	0054	4169.0	3494.0	0.838	0.97	0.358	0.537
55	0055	4169.0	2236.0	0.536	0.09	0.171	0.214

La salida que se obtuvo en el análisis realizado con el programa Bilog con leyenda "porcent", se refiere al índice de dificultad.

Análisis bajo la Teoría de Respuesta al Item (TRI)

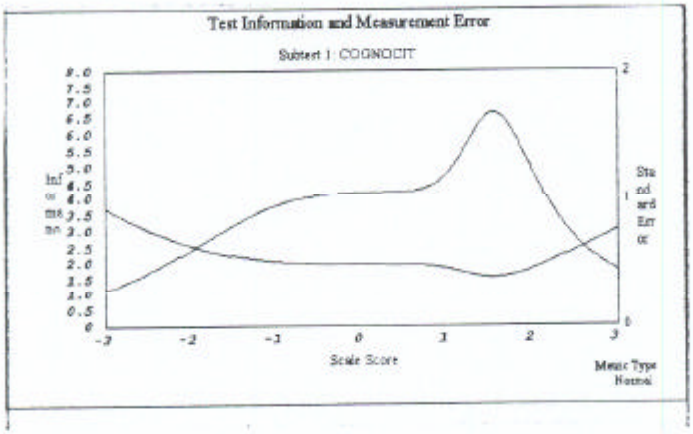
Se utilizó un modelo logístico de tres parámetros, en este modelo, se puede diferenciar los índices de dificultad y discriminación y la probabilidad de contestar al azar por ser una prueba cognoscitiva.

El análisis de los resultados bajo la TRI constituyó dos etapas. Primero, se presenta un análisis de la función de información de la prueba, el análisis individual de los parámetros y de la función de información para cada ítem. Segundo, un análisis de posible sesgo en los ítemes.

Interpretación de la función de información

Gráfico N 2

Curva Característica de la Prueba



Se refiere a la suma de las funciones de información de todos los ítemes. El gráfico nº 2 con la Curva característica del Test (CCT) muestra una curva poco usual y diferente a la curva normal que es la que se presenta comúnmente.

La curva de esta prueba se encuentra ligeramente orientada hacia la derecha, el rango donde se maximiza la función de información es reducido, brinda mayor información en niveles altos. Ante esto, lo que sería deseable es que se lograra un mayor nivel de información en los niveles intermedios. O sea que pocos ítemes discriminen en niveles bajos y pocos también en los altos.

Análisis de ítemes (TRI):

1. Los ítemes que deben ser mejorados pues no ofrecen suficiente información y cuentan con poco poder discriminatorio son:

02 -	03 -	05 -	06 -	07 -	12 -	13 -	14 -	15 -	16 -	17 -
18 -	19 -	20 -	21 -	22 -	23 -	24 -	25 -	26 -	27 -	28 -
30 -	31 -	32 -	33 -	34 -	36 -	37 -	38 -	39 -	40 -	42 -
43 -	44 -	45 -	46 -	47 -	48 -	49 -	51 -	52 -	53 -	55 -

Estos ítemes representan el 80% del total en la parte de selección de la prueba de español o sea, 44 ítemes de 55.

A continuación se presentan el comportamiento de los 55 ítemes en cada uno de los parámetros, a saber, parámetro *a* es el índice de discriminación, *b* el índice de dificultad y *c* la probabilidad de responder a azar el ítem.

TABLA 5

ITEM	PARÁMETRO: *a*	PARÁMETRO: *b*	PARÁMETRO: *c*
1	1,00	1,48	0,41
2	0,68	0,82	0,37
4	0,92	1,33	0,31
5	0,49	0,32	0,31
6	0,65	0,06	0,28
7	0,59	-2,30	0,30
8	0.85	-0.08	0,27
9	1,68	1,56	0,27
10	0,86	0,52	0,39
11	2,26	1,47	0,35
12	0,67	2,65	0,25
14	0,89	0,78	0,49
15	0,64	-0,41	0,22
16	0,66	-0,42	0,46
17	0,64	-2,39	0,31
18	0,50	0,71	0,29
19	0,46	-1,40	0,29
20	0,40	-0,57	0,28
21	0,57	-0,10	0,39
22	0,31	-2,54	0,25
23	0,38	-1,09	0,23

24	0,16	-2,31	0,35
25	0,30	-1,94	0,28
26	0,41	-2,40	0,30
28	0,23	-1,74	0,30
29	0,81	-1,15	0,20
30	0,64	-2,49	0,32
31	0,75	-1,77	0,31
32	0,59	-0,60	0,32
33	0,29	-2,12	0,27
34	0,38	0,94	0,26
35	0,71	0,87	0,20
36	0,60	0,15	0,12
37	0,41	-0,39	0,25
38	0,49	-0,01	0,39
40	0,29	-0,15	0,24
41	0,80	-1,01	0,32
42	0,23	0,50	0,27
43	0,36	1,58	0,40
44	0,49	1,53	0,19
45	0,52	2,30	0,42
46	0,47	0,08	0,23
47	0,21	-0,91	0,33
48	0,43	0,44	0,45
49	0,15	-2,47	0,36
50	0,79	-1,96	0,27
51	0,43	0,32	0,17
52	0,30	-0,14	0,25
53	0,60	-0,72	0,17
54	0,97	-1,13	0,24
55	0,59	1,34	0,38

Para realizar el análisis de cada ítem, se definieron los siguientes estándares, considerando la longitud de la escala de –3 a 3, como se muestra:

-3	-2	-1	0	1	2	3
Muy fácil	Fácil	Ligeramente fácil	Ligeramente difícil	Difícil	Muy difícil	

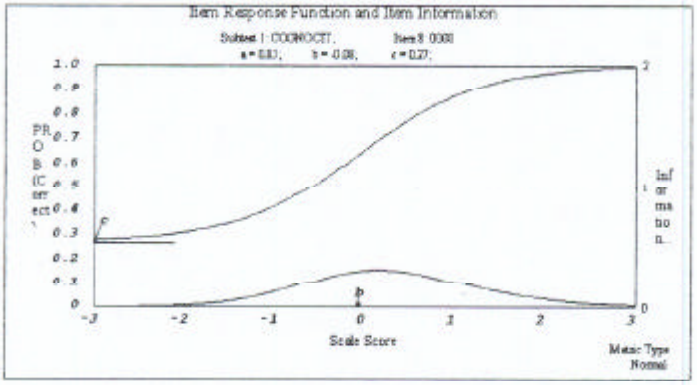
Ítemes de aceptable calidad técnica

Los ítemes que muestran una aceptable función de información y que discriminan en niveles bajos con una dificultad:

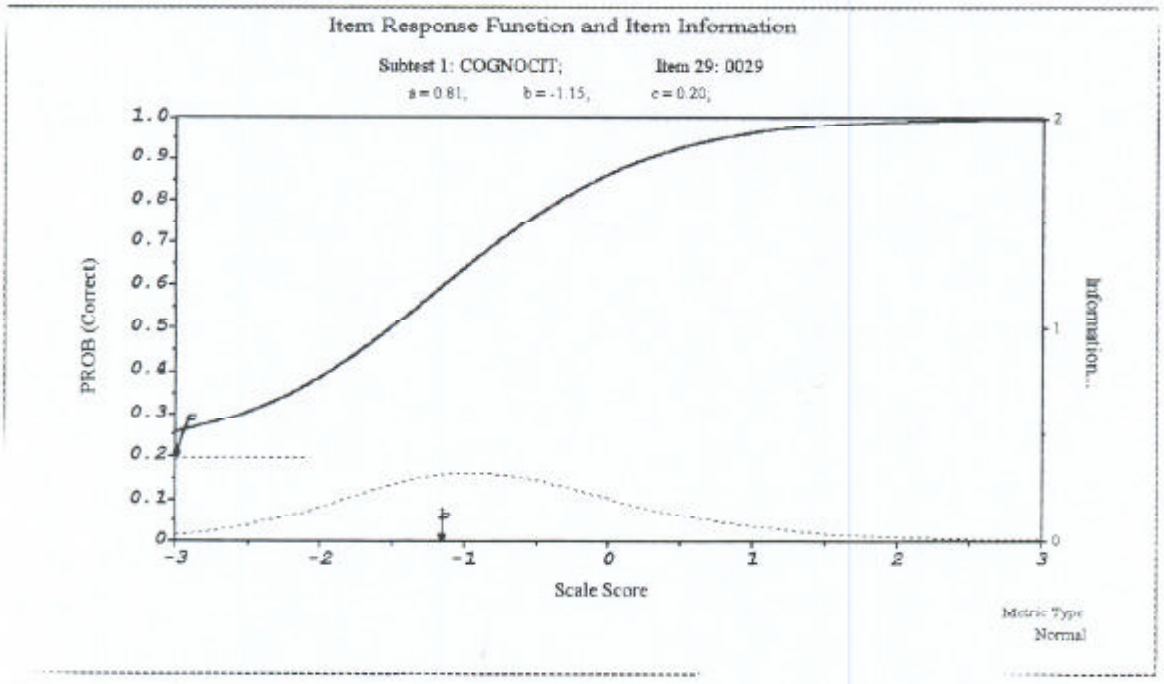
Ligeramente Fácil:	Fácil:
08	29, 41, 50, 54

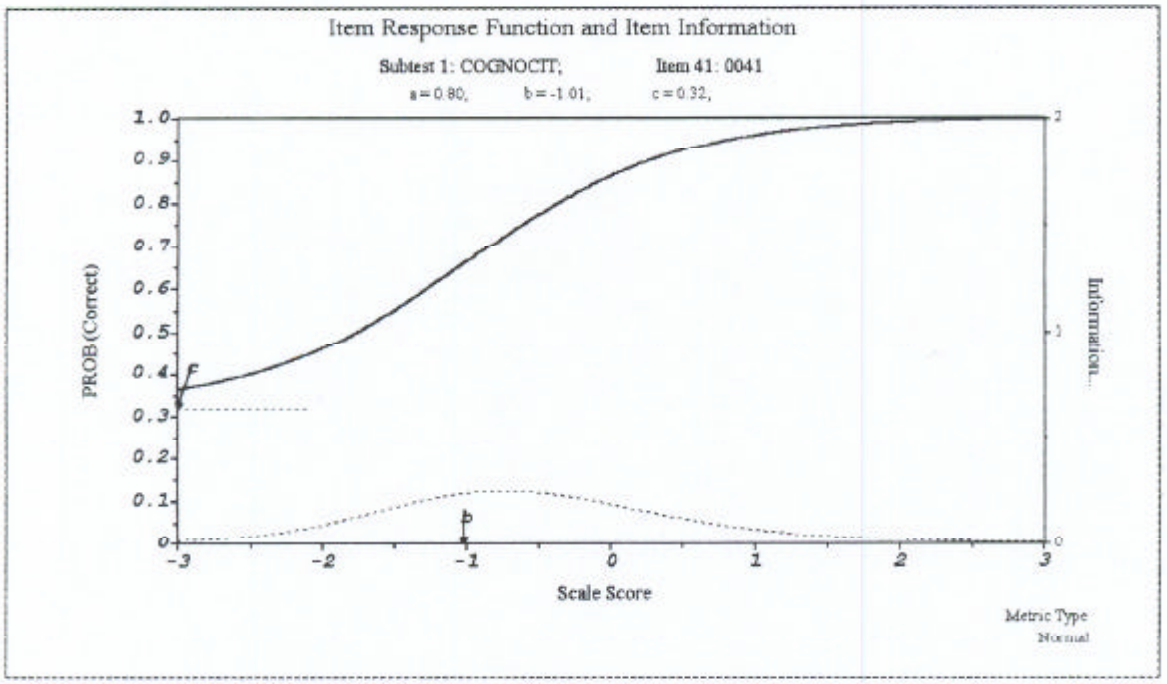
Ligeramente fácil:

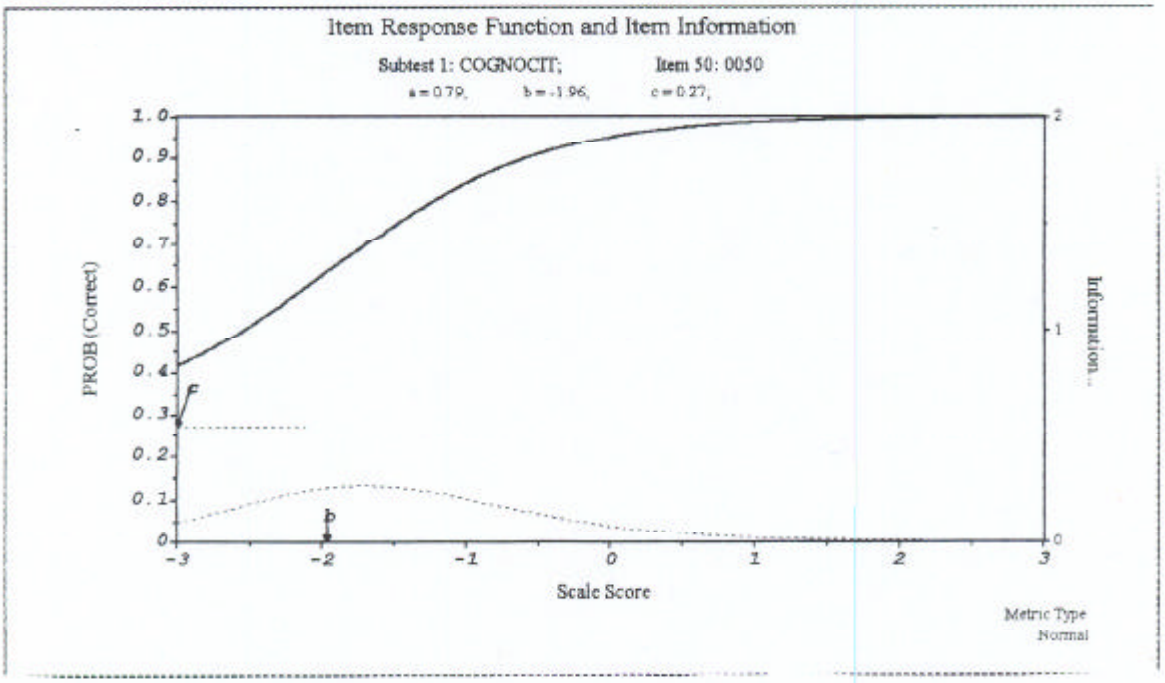
Item 08

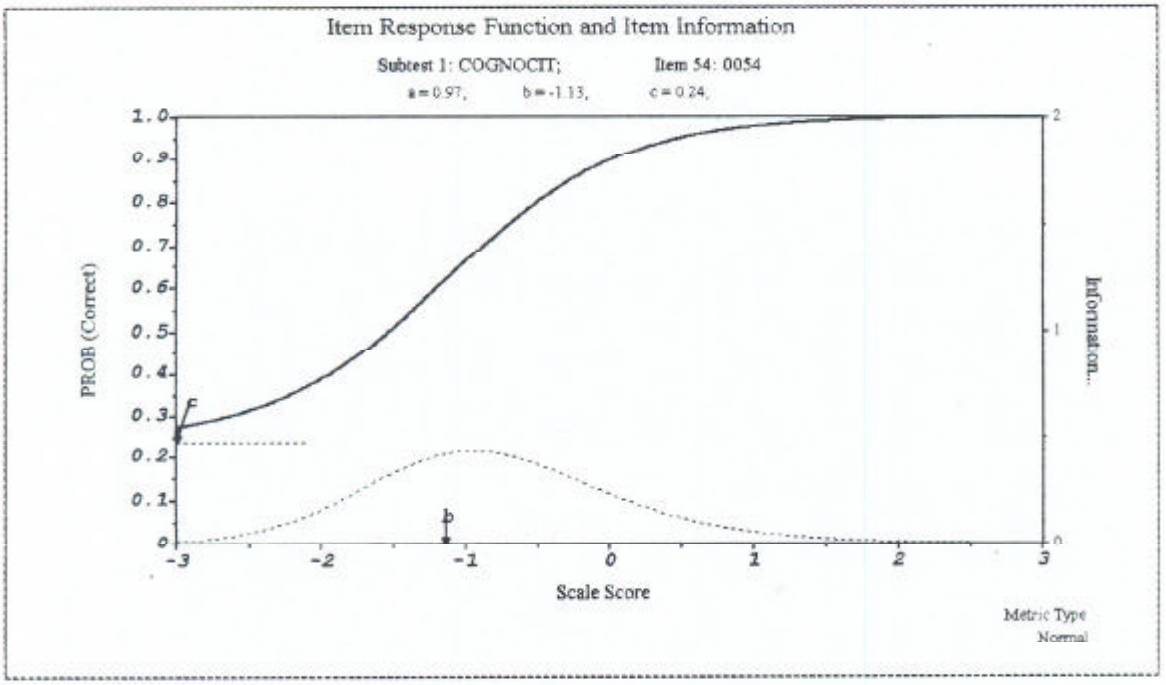


Fácil:




Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 41: 0041
a = 0.80,
b = -1.01,
c = 0.32,
PROB (Correct)
Information
Scale Score
Metric Type
Normal


Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 50: 0050
a = 0.79,
b = -1.96,
c = 0.27,
PROB (Correct)
Information
Scale Score
Metric Type
Normal

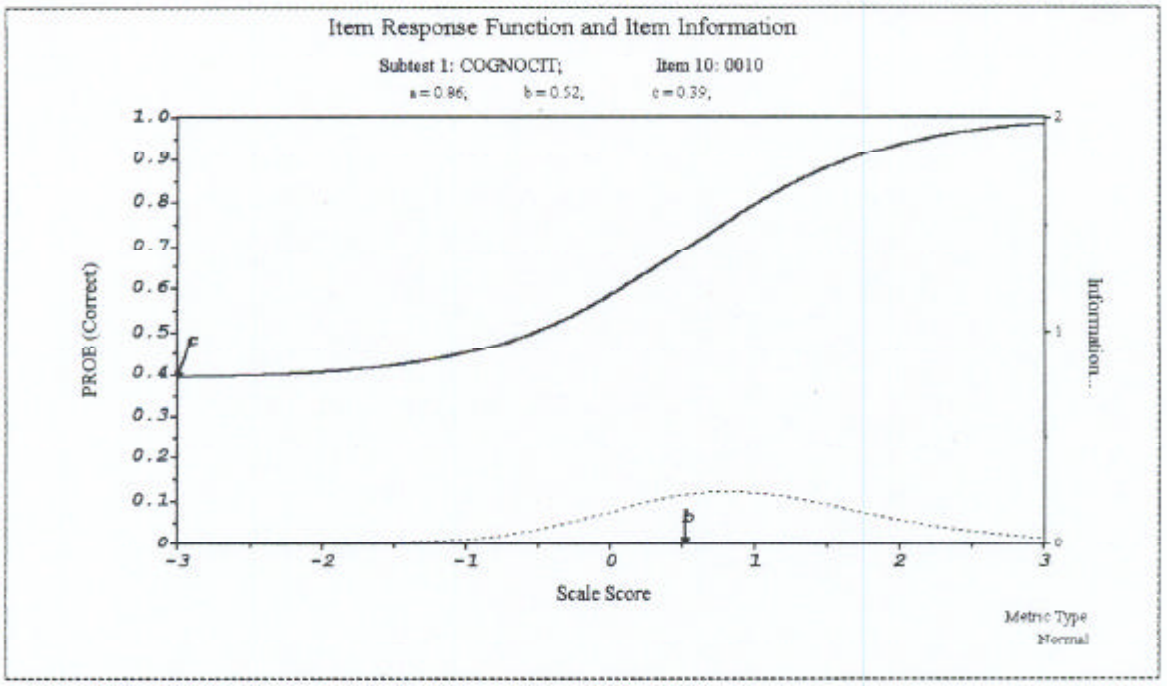


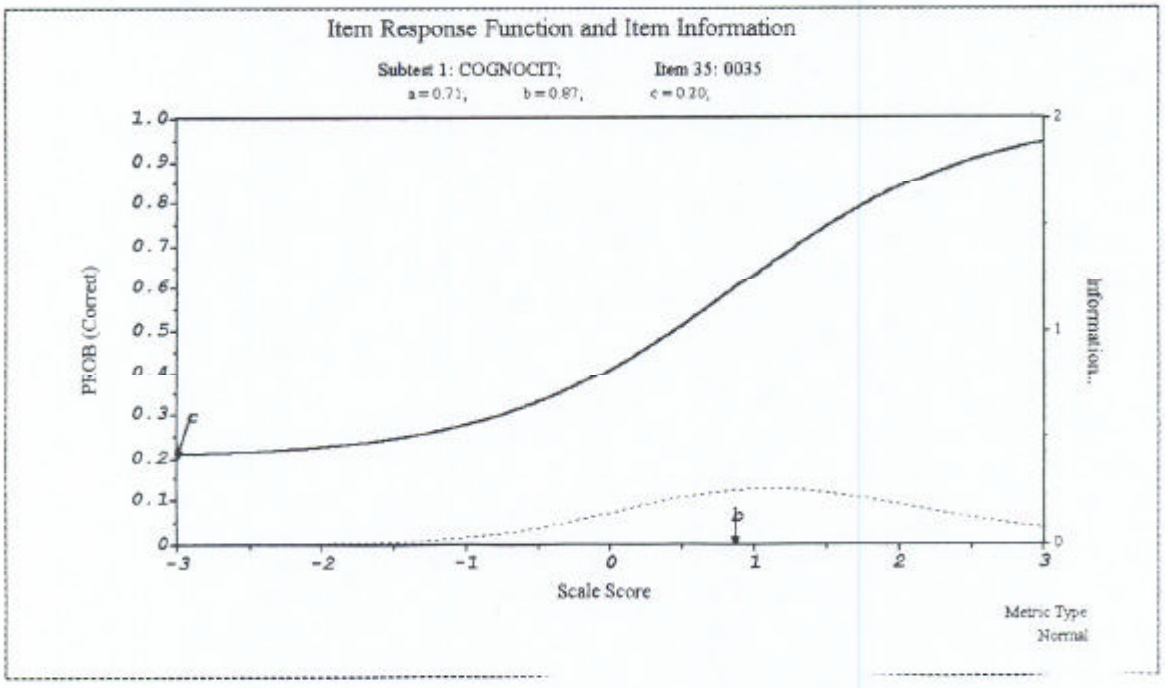
Los ítemes que muestran una aceptable función de información y que discriminan en niveles altos con una dificultad:

Ligeramente Difícil:	Difícil
10, 35	01, 04

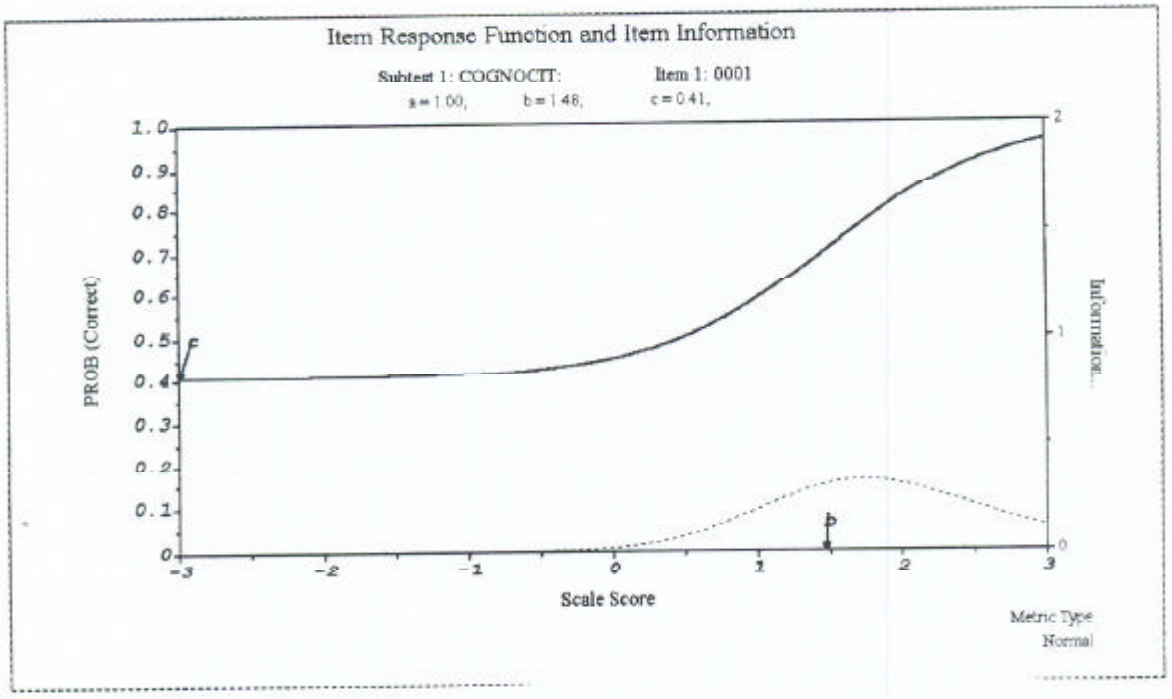
Ligeramente difícil:

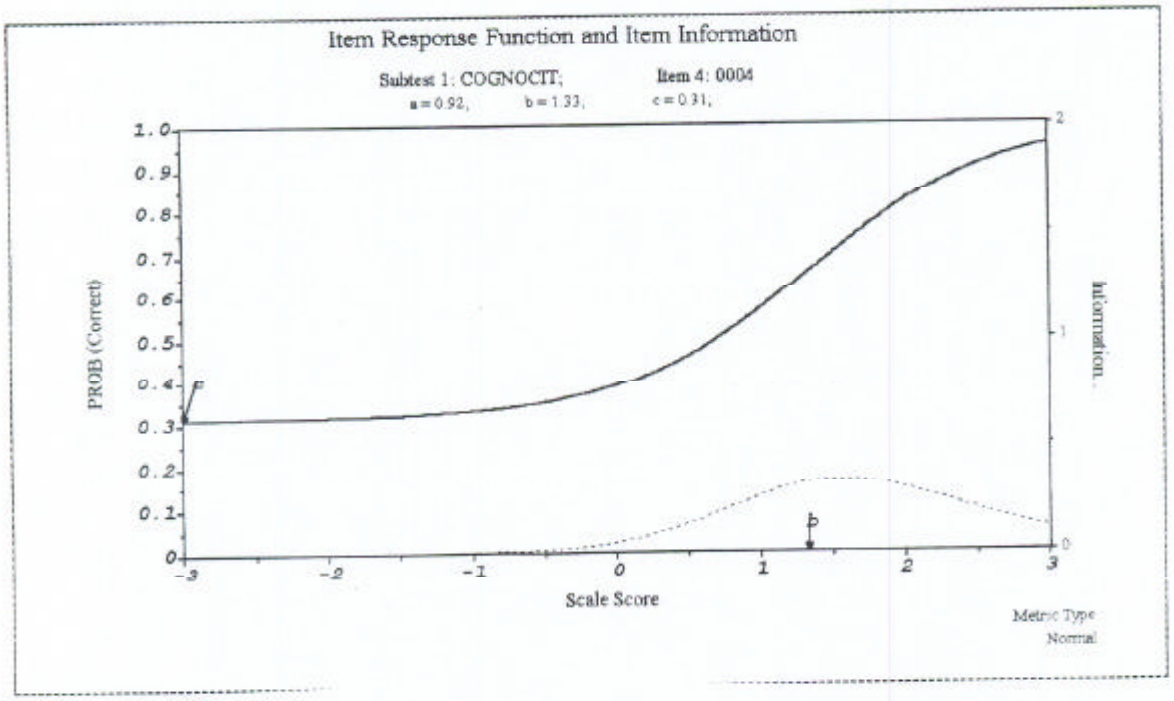
Item 10


Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 10: 0010
a = 0.86,
b = 0.52,
c = 0.39,
PROB (Correct)
Information
Scale Score
Metric Type
Normal


Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 35: 0035
a = 0.71,
b = 0.87,
c = 0.20,
PROB (Correct)
Information
Scale Score
Metric Type
Normal

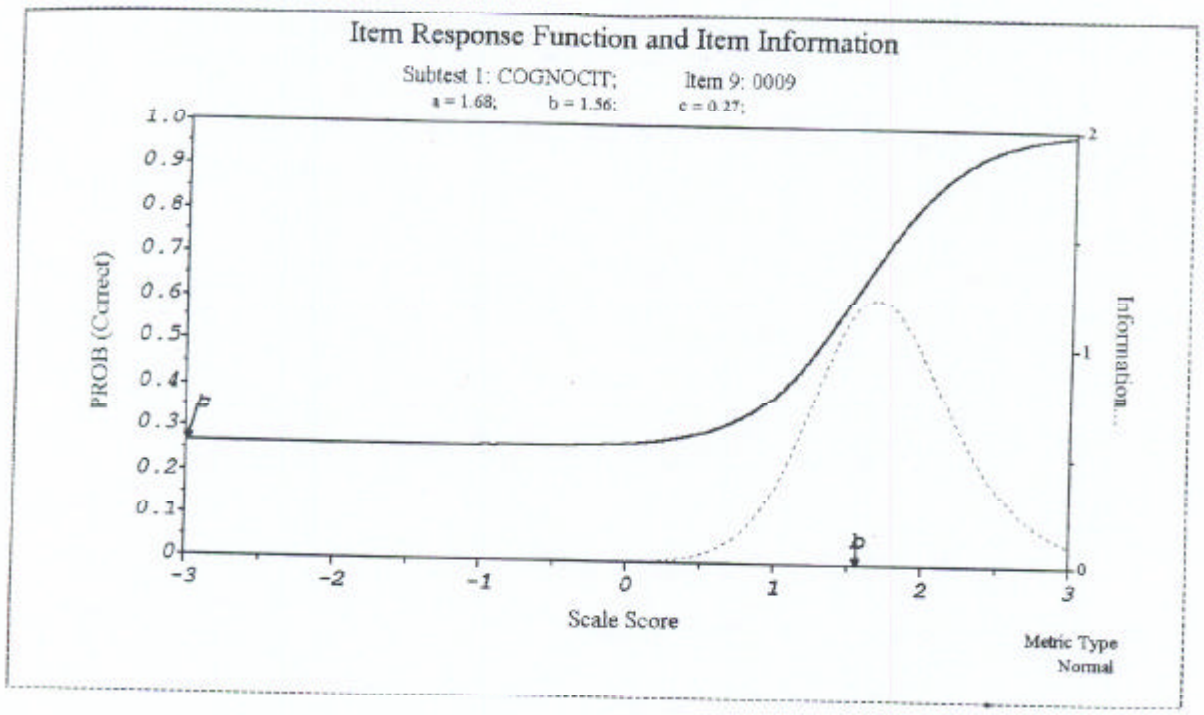
Dificil:

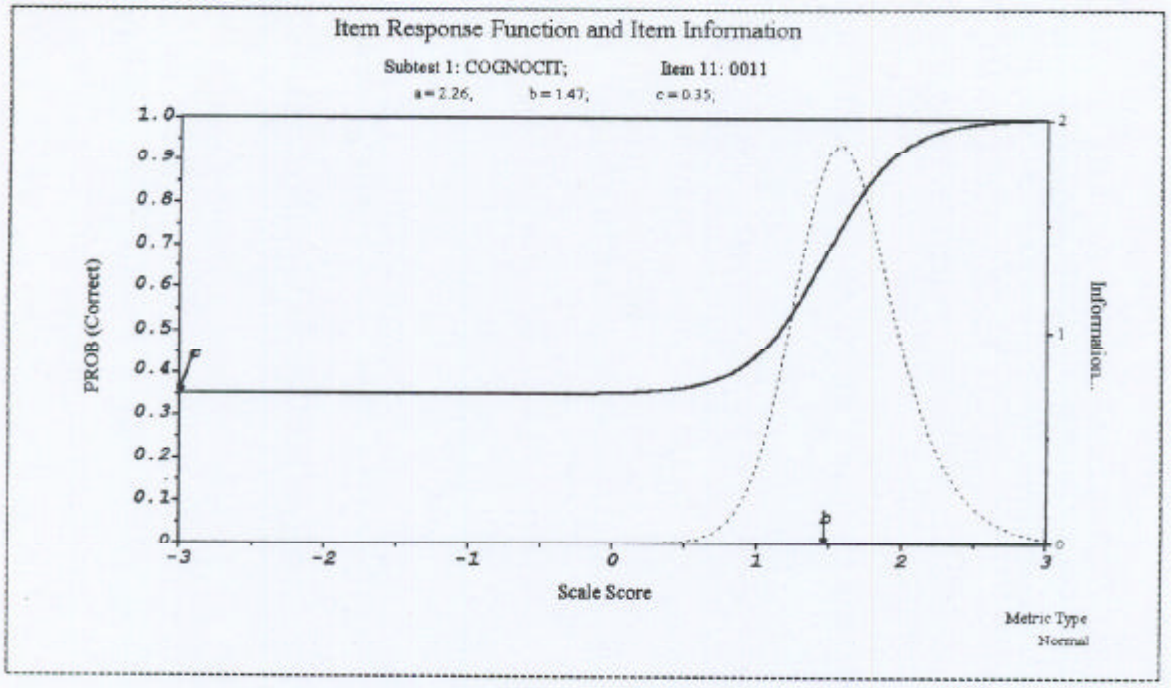

Item Response Function and Item Information
Subtest 1: COGNOCIT:
Item 1: 0001
a = 1.00,
b = 1.48,
c = 0.41,
PROB (Correct)
Information
Scale Score
Metric Type
Normal


Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 4: 0004
a = 0.92,
b = 1.33,
c = 0.31,
PROB (Correct)
Information
Scale Score
Metric Type
Normal

Los ítemes que muestran una alta calidad técnica, con suficiente función de información que discrimina en niveles altos con una dificultad:

Difícil





Prbabilidad de Azar

Los ítemes con parámetro c mayor que 0,25, se observa en siguiente tabla:

TABLA 6

>= 0,25	> 0,25 < 0,30	>= 0,30 < 0,40	>= 0,40
12	3	2	**1**
15	6	**4**	14
22	**8**	**5**	16
23	9	7	43
29	18	10	45
35	**19**	**11**	48
36	20	13	
37	25	17	
40	27	21	
44 46	33 34	24 26	
51	39	28	
52	42	30	
53	**50**	31	
54		32	
		38	
		41	
		47	
		49	
		55	

* Los ítemes que están resaltados con negrilla tienen aceptable y alta calidad técnica.

A continuación se presenta los porcentajes de los ítemes con respecto a la probabilidad de contestar al azar los ítemes:

TABLA 7

	Cantidad de Ítemes	Porcentaje %
Ítemes con un parámetro de azar menor o igual a 0,25	15	27,27 %
Ítemes con un parámetro de azar mayor a 0,25	40	72,73%
Total	55	100

El cuadro anterior indica que el 27,27% de los ítemes de la parte de la selección tienen baja probabilidad de contestar al azar y el 72,73% tiene probabilidad de ser respondido al azar.

ANALISIS DE SESGO

El análisis de sesgo de la Prueba, se realizó desde tres ámbitos: sexo, tipo de colegio (público-privado) y zona en la que se ubica el mismo (rural, urbano)

La prueba comtempla 56 ítemes, de los cuales el 55 fue eliminado por el Ministerio de Educación, por ello el ítem 56 pasa a tomar el lugar del ítem 55.

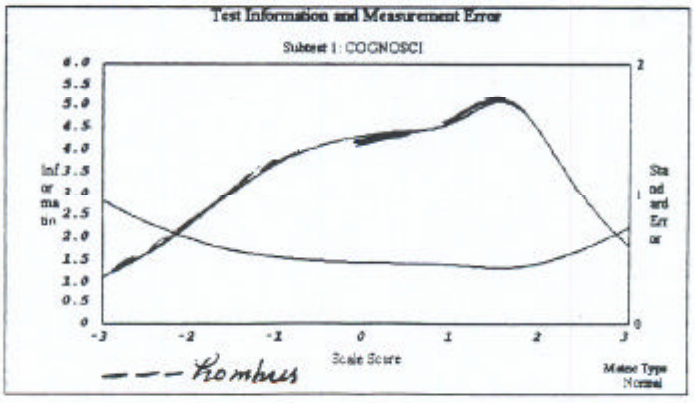
Análisis de Sesgo según Sexo

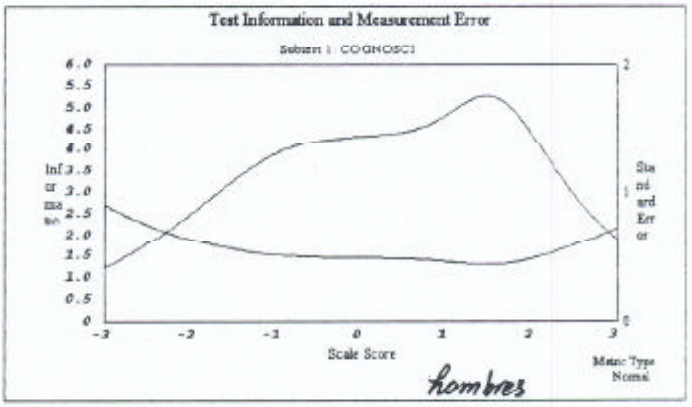
Análisis de la función de información (Mujeres – Hombres)

Se refiere a la suma de las funciones de información de todos los ítemes. Los gráficos muestran una curva no bien definida, cargadas a la derecha, niveles difíciles; en ambas poblaciones las curvas son muy parecidas.

Las curvas muestran mayores niveles de información en niveles intermedios y altos como puede verse en ambos gráficos:

Mujeres





La tabla 8 presenta los resultados en relación con la discriminación y dificultad según las salidas del programa Bilog.

A partir de los datos consignados en esta tabla se construyeron los gráficos de dispersión, como se pueden observar a continuación.

10.01571

Instituto de Investigación para el Mejoramiento de la Educación Costarricense (IIMEC) Facultad de Educación

18-0901
946

Tabla 8

c:\programs\spsswin\karla.sav

	dischom	discmuje	difhom	difmuje
1	.59	.74	.81	1.61
2	1.14	.71	1.35	.77
3	.33	.44	-5.28	-4.10
4	.76	.91	1.22	1.40
5	.50	.48	.50	.14
6	.59	.71	.59	.06
7	.54	.63	.54	-2.21
8	.79	.97	-.27	.15
9	1.37	1.23	1.59	1.63
10	1.04	.66	.53	.37
11	1.52	1.66	1.45	1.55
12	.81	.50	2.56	2.84
13	.24	.27	-6.52	-4.91
14	.68	.96	.82	.64
15	.64	.68	-.49	-.24
16	.85	.47	-.14	-1.22
17	.61	.65	-2.33	-2.50
18	.41	.62	-.76	-.57
19	.47	.45	-1.21	-1.53
20	.44	.38	-.45	-.55
21	.55	.53	.00	-.44
22	.33	.32	-2.48	-2.28
23	.42	.37	-1.19	-.71
24	.19	.18	-2.09	-2.18
25	.26	.37	-1.98	-1.66
26	.46	.39	-2.52	-2.14
27	.27	.26	-4.00	-4.81
28	.27	.23	-1.92	-1.23
29	.83	.83	-1.14	-1.02
30	.67	.59	-2.38	-2.70
31	.76	.73	-1.71	-1.86
32	.58	.59	-.41	-.83
33	.28	.31	-2.17	-1.82
34	.38	.41	.93	1.07

c:\programs\spsswin\karla.sav

	dischom	discmuje	difhom	difimuje
35	.75	.71	.91	.86
36	.59	.66	.22	.23
37	.41	.44	-.24	-.35
38	.51	.46	-.07	-.14
39	.43	.38	-3.98	-3.93
40	.30	.31	.00	.03
41	.71	.88	-1.11	-1.00
42	.26	.25	.33	1.06
43	.40	.33	1.46	1.55
44	.46	.55	1.73	1.42
45	.51	.47	2.35	2.25
46	.43	.57	.20	.16
47	.24	.22	-1.16	-.40
48	.41	.34	.09	.08
49	2.16	.17	-2.60	-1.96
50	.79	.81	-2.13	-1.77
51	.46	.45	.57	.34
52	.29	.35	.22	-.14
53	.59	.67	-.78	-.43
54	1.06	.91	-1.08	-1.15
55	.51	.75	.64	1.72

GRÁFICO A

Gráfico de dispersión de la **dificultad**
De hombres (y) y mujeres(x) de la Prueba de Bachillerato
En Español. 1997.

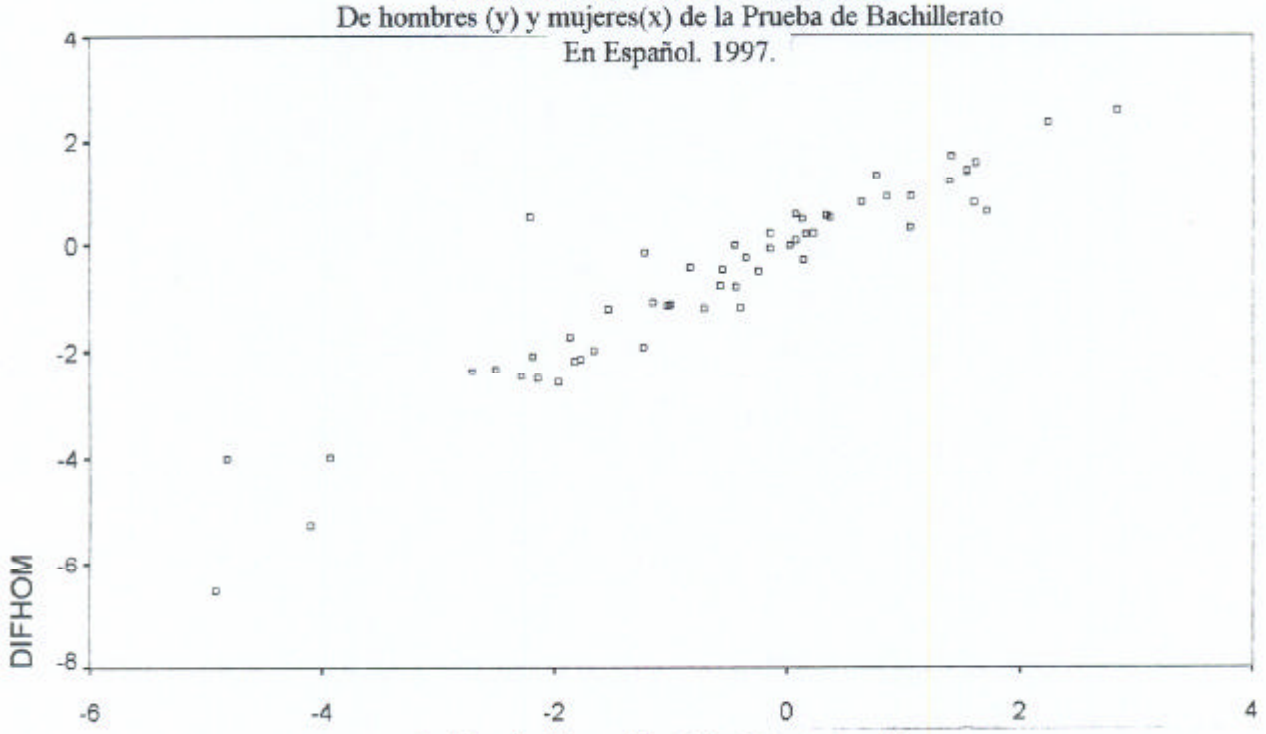
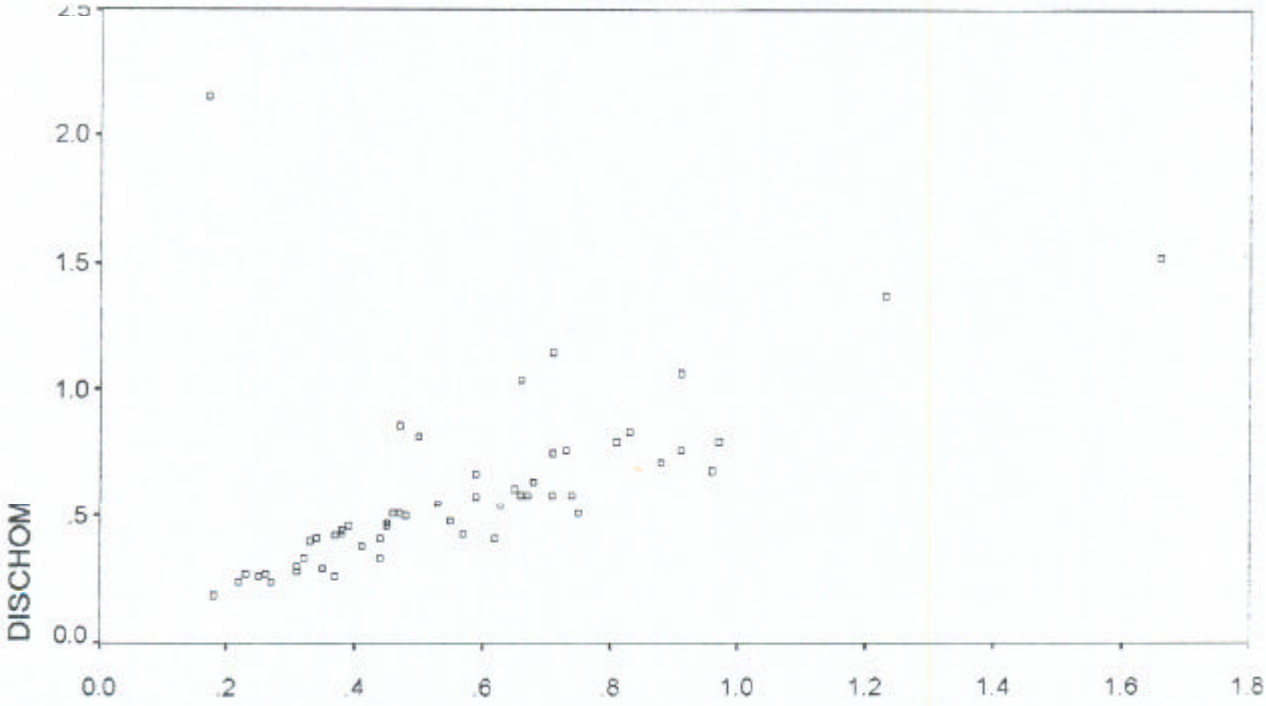


Gráfico de dispersión de la **diiscriminación**
De hombres (y) y mujeres(x) de la Prueba de Bachillerato
En Español. 1997.



El gráfico "A" de dificultad muestra que los puntos de cada uno de los ítemes tienen una tendencia de agrupamiento. Debido a la dificultad que presenta el gráfico para desde él definir cuáles son los ítemes más alejados, y que por lo tanto presentan sesgo, se procedió a complementar el análisis, con los resultados de la aplicación de la regresión múltiple (outliers), para poder determinar los ítemes que más se alejan y por ende, que presentan posibilidad de sesgo.

De estos análisis se concluye que el ítem que presenta sesgo es el número 7.

Salidas outliers

```
16 Dec 98 SPSS for MS WINDOWS Release 6.1

          * * * *   M U L T I P L E   R E G R E S S I O N   * * * *

Equation Number 1     Dependent Variable..   DIFHOM

Casewise Plot of Standardized Residual

Outliers = 3.     *: Selected   M: Missing

          -6.      -3.  3.      6.
  Case #   0:.......:  :.......:0     DIFHOM       *PRED       *RESID
       7   .           ..    *    .        .54     -2.2465      2.7865

      1 Outliers found.
```

* *

El gráfico "B" corresponde a la dispersión de la discriminación, muestra que los puntos de cada uno de los ítemes tienden a agruparse en una recta. Se requirió complementar el análisis de estos resultados con los de la regresión simple (out liers) para indicar que sólo el item 49 presenta sesgo

Salidas outliers

```
          * * * *   M U L T I P L E   R E G R E S S I O N   * * * *

Equation Number 1    Dependent Variable..   DISCHOM

Casewise Plot of Standardized Residual

Outliers = 3.    *: Selected   M: Missing

          -6.      -3.  3.      6.
  Case #   O:.......:  :.......:O    DISCHOM       *PRED       *RESID
     49    .          ..          *       2.16       .3205       1.8395

      1 Outliers found.
```

* *

De acuerdo con la dificultad:

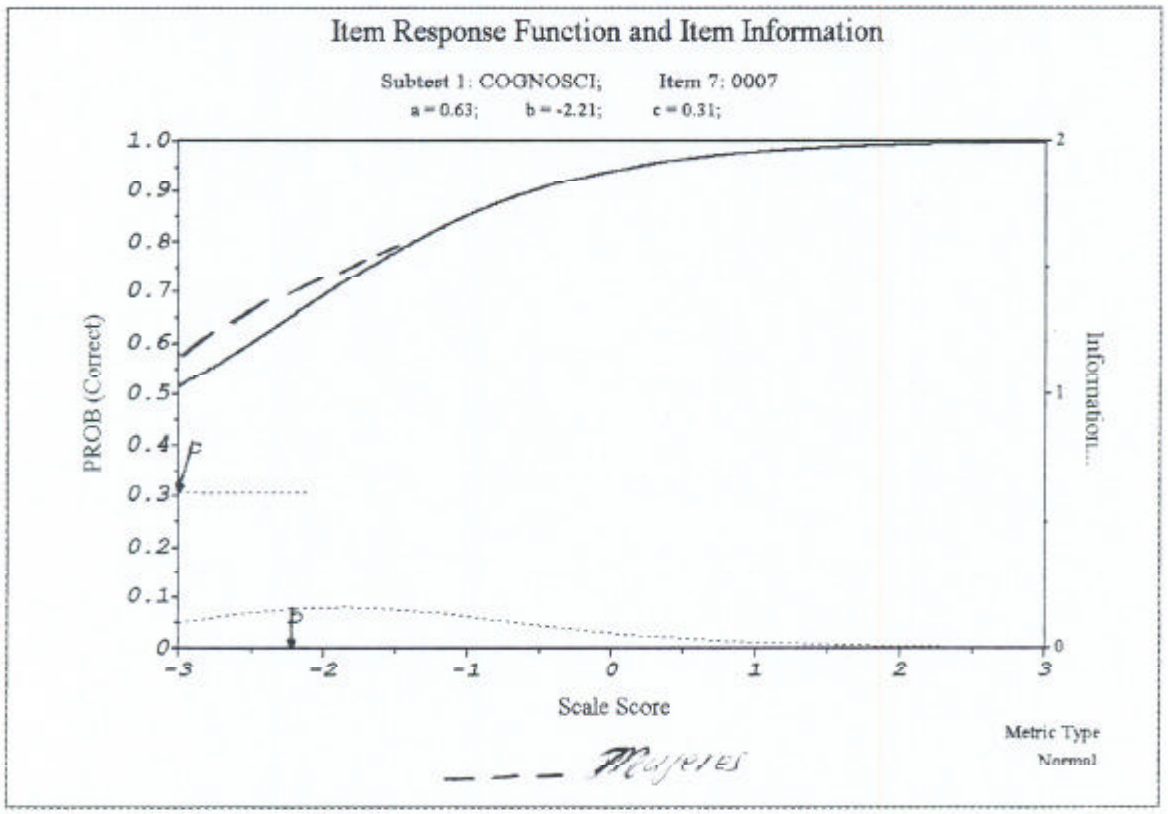
Item 7

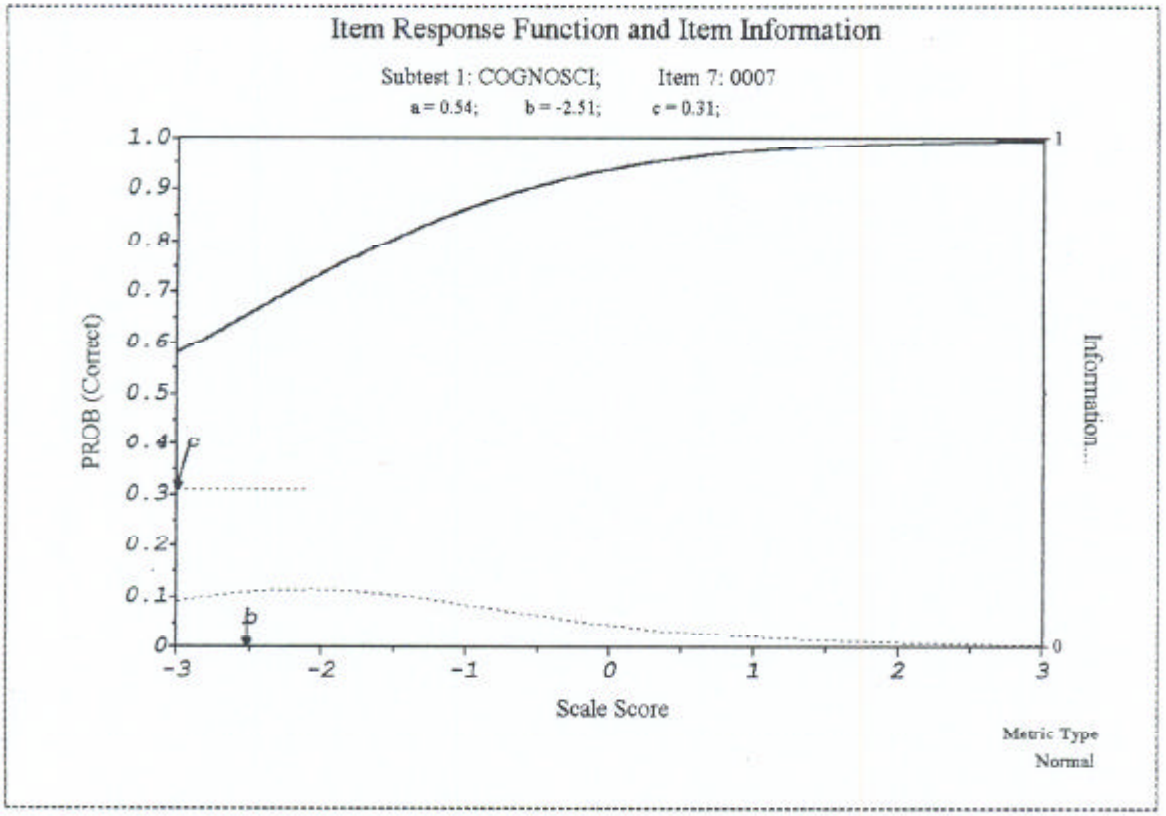
Presenta sesgo a favor de las mujeres en niveles bajos de habilidad. Es un item que no da suficiente información y no presenta suficiente poder discriminatorio en el nivel general. Para las dos subpoblaciones es un item muy fácil con un azar de 0.31 en ambos casos.

No se pueden precisar las causas del sesgo porque el item, en las dos subpoblaciones, tampoco brinda suficiente información y no presenta suficiente poder discriminatorio.

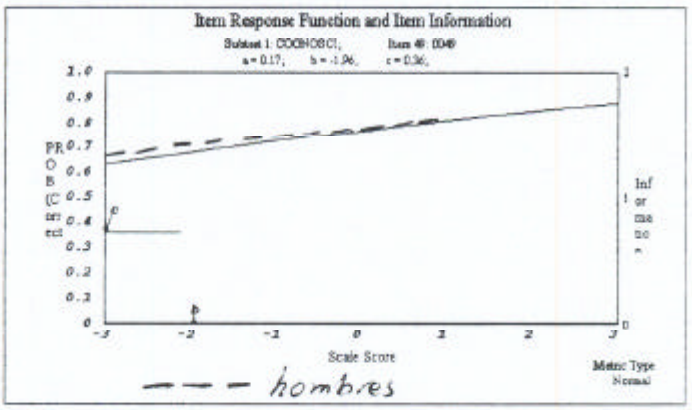
De acuerdo con la discriminación :

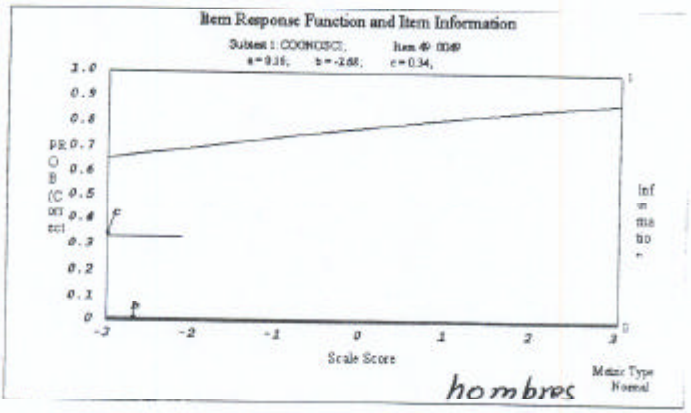
Item 49


Item Response Function and Item Information
Subtest 1: COGNOSCI; Item 7: 0007
a = 0.63; b = -2.21; c = 0.31;
PROB (Correct)
Information
Scale Score
Metric Type
Normal


Item Response Function and Item Information
Subtest 1: COGNOSCI; Item 7: 0007
a = 0.54; b = -2.51; c = 0.31;
PROB (Correct)
Information
Scale Score
Metric Type
Normal
c
b
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
0
-3
-2
-1
0
1
2
3

Mujeres


Item Response Function and Item Information
Subtest 1: COGNOSCI,
a = 0.17; b = -1.96; c = 0.36;
PROB (Correct)
Information
Scale Score
Metric Type
Normal
hombres


Item Response Function and Item Information
Subtest 1: COGNOSCI,
a = 0.16; b = -2.68; c = 0.34;
PROB (Correct)
Information
Scale Score
Metric Type
Normal
hombres

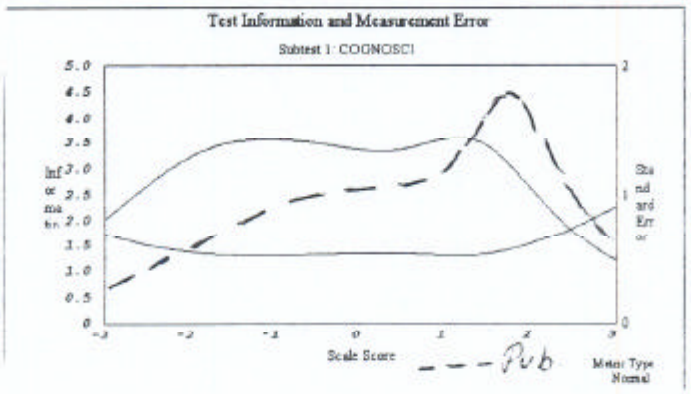
Es un item que brinda muy poca información y tiene muy poco poder discriminatorio en el nivel general. Es un item fácil para las dos subpoblaciones presenta sesgo en los niveles bajos e intermedios a favor de los hombres.

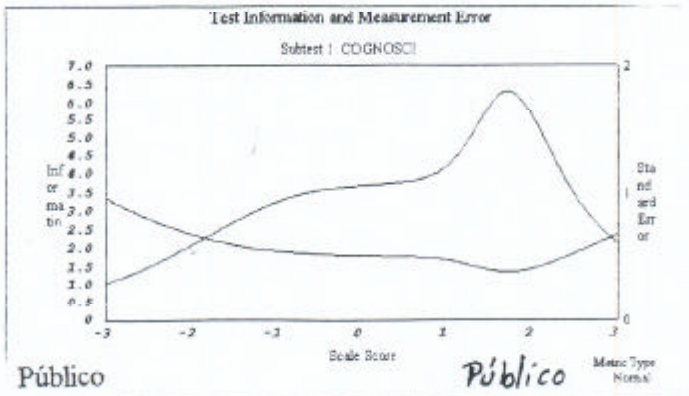
Sesgo según tipo de colegio

Análisis de la función de información

Se refiere a la suma de las funciones de información de todos los ítemes . los gráficos muestran curvas no bien definidas, a lo esperado semejantes a una curva normal.

En el caso de los colegios púbicos hay una mayor información en los niveles de habilidad altos, situación que se evidencia por la ubicación de la parte más empinada de la curva. La mayor información se concentra en los niveles más difíciles de habilidad para los colegios públicos. En el caso de los colegios privados la información se concentra en niveles intermedios teniendo mas relevancia los niveles fáciles y difíciles de habilidad.





Público

Tabla 9

c:\programs\spsswin\karla.sav

	discpub	discpriv	difpub	difpriv
1	.88	.96	1.25	1.62
2	.40	.90	-.47	1.19
3	.31	.39	-6.42	-4.44
4	.89	.87	1.15	1.48
5	.49	.41	-.40	.51
6	.58	.62	-.66	.29
7	.59	.53	-3.24	-2.28
8	.64	.87	-.90	.17
9	.96	1.68	1.20	.73
10	.59	.74	-.54	.76
11	1.34	2.04	1.24	.60
12	.49	.70	2.36	2.85
13	.32	.20	-5.20	-6.66
14	.48	.86	-1.15	1.30
15	.56	.57	-1.34	-.18
16	.67	.45	-1.67	-.64
17	.55	.58	-3.78	-2.37
18	.36	.51	-1.81	-.44
19	.34	.45	-2.62	-1.23
20	.35	.38	-1.40	-.31
21	.49	.47	-1.29	.00
22	.33	.30	-2.82	-2.46
23	.38	.35	-1.63	-.93
24	.24	.15	-2.15	-2.42
25	.34	.28	-2.21	-1.89
26	.47	.36	-2.80	-2.45
27	.16	.30	-7.97	-3.80
28	.22	.22	-2.58	-1.47
29	.69	.76	-2.06	-.98
30	.71	.61	-2.74	-2.50
31	.65	.75	-2.53	-1.65
32	.58	.51	-1.33	-.57
33	.24	.27	-3.52	-1.93
34	.33	.46	.86	1.09

c:\programs\spsswin\karla.sav

	discpub	discpriv	difpub	difpriv
35	.67	.68	.28	1.15
36	.68	.59	-.09	.31
37	.45	.37	-.88	-.21
38	.49	.46	-.85	.13
39	.26	.42	-6.82	-3.64
40	.29	.28	-.34	.01
41	.75	.70	-1.93	-.96
42	.23	.22	-.22	1.12
43	.37	.30	-.90	1.62
44	.51	.46	-1.16	1.84
45	.34	.61	-1.33	2.47
46	.50	.45	-.52	.36
47	.19	.21	-2.00	-.57
48	.35	.37	-.92	.50
49	.19	.15	-2.64	-2.40
50	.73	.75	-2.62	-1.89
51	.35	.46	-.03	.62
52	.27	.31	-.62	.15
53	.53	.59	-1.14	-.56
54	.79	.96	-2.01	-.95
55	.44	.58	-.56	1.59

En la tabla 9 se presentan los resultados que corresponden a la discriminación y la dificultad, de acuerdo con las salidas del programa Bilog. A partir de los datos allí contemplados se construyeron los gráficos de dispersión de la discriminación y la dificultad de acuerdo con el tipo de colegio.

GRÁFICO C

Gráfico de dispersión de la **dificultad**
privados (y) y públicos(x) de la Prueba de Bachillerato
En Español. 1997.

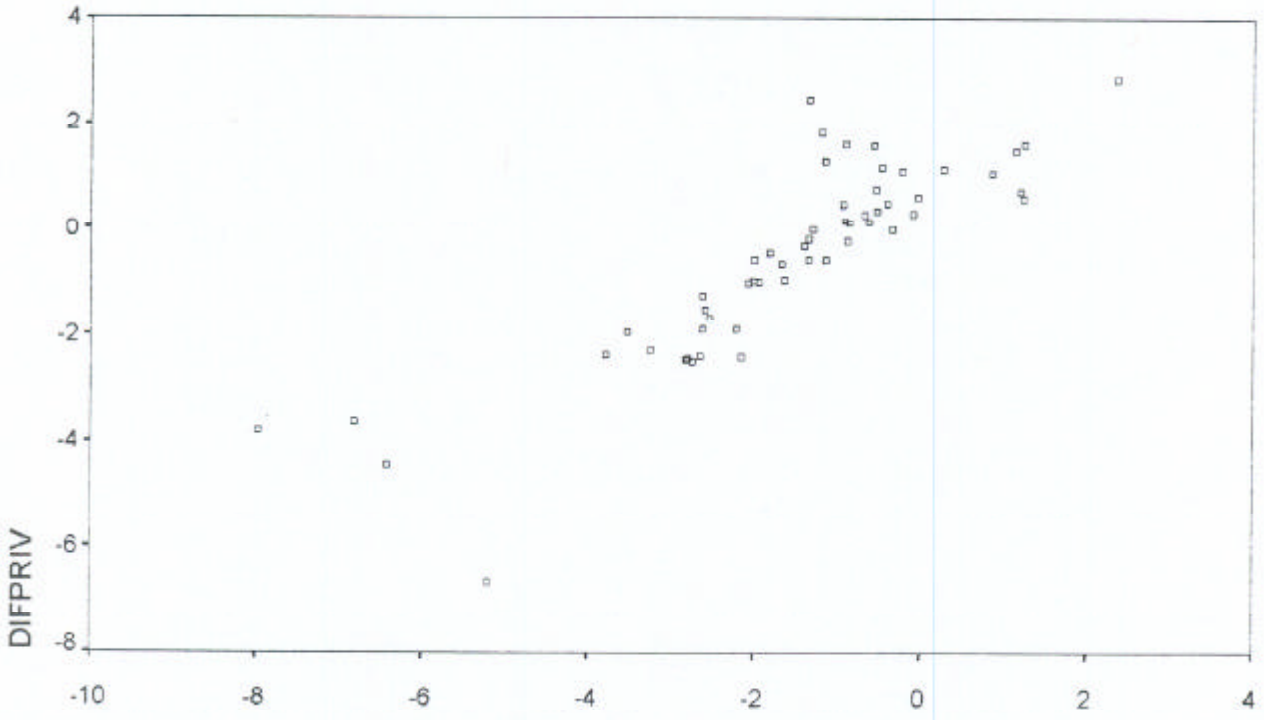
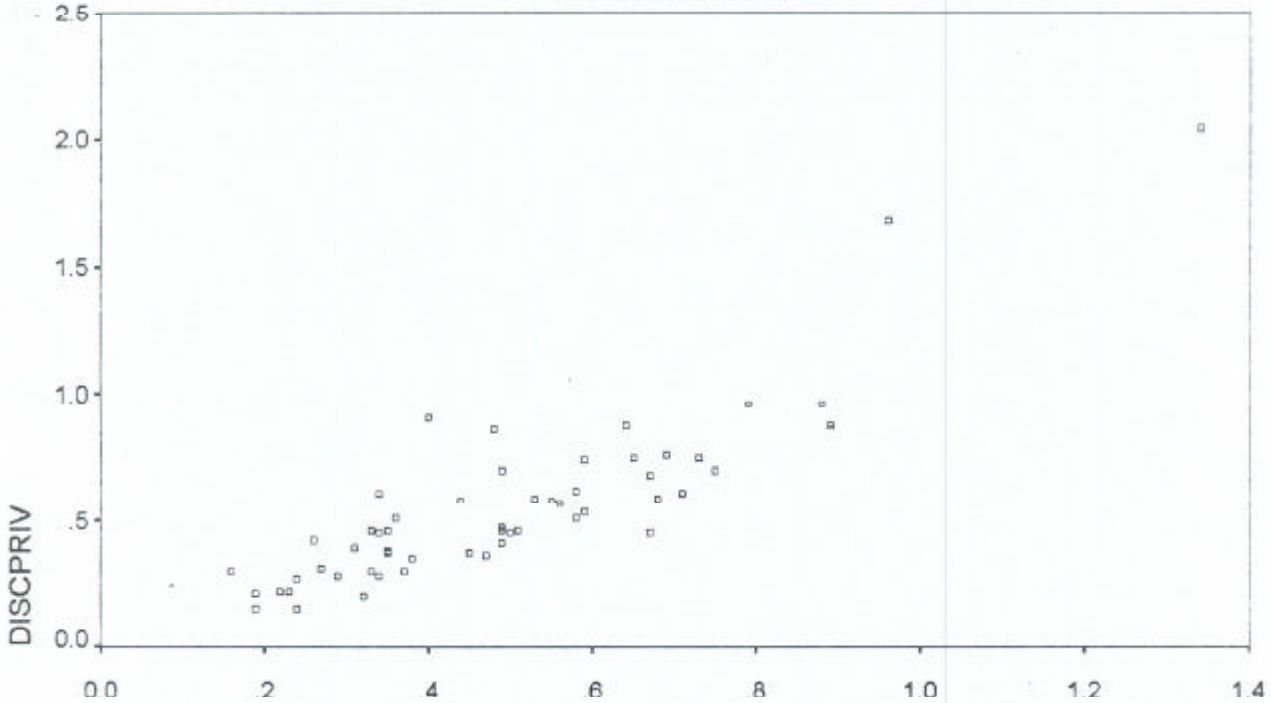


GRÁFICO D

Gráfico de dispersión de la discriminación
De privados (y) y públicos(x) de la Prueba de Bachillerato
En Español. 1997.



El gráfico C de dificultad demuestra la tendencia de agrupamiento de los puntos de cada item. Debido a la dificultad que presenta el gráfico para definir cuáles son los ítemes más alejados y que por lo tanto evidencian sesgo, se procedió a complementar el análisis con los resultados de la aplicación de la regresión simple (outliers) para poder establecer cuáles itemes podrían presentar evidencia de sesgo.

De estos análisis se concluye que el único item que muestra evidencia de sesgo es el item 9.

Salidas outliers

```
          * * * *   M U L T I P L E   R E G R E S S I O N   * * * *

Equation Number 1    Dependent Variable..   DISCPRIV

Casewise Plot of Standardized Residual

Outliers = 3.    *: Selected   M: Missing

          -6.      -3.  3.      6.
  Case #   O:.......:  :.......:O   DISCPRIV       *PRED       *RESID
       9   .          ..*         .       1.68     1.1570       .5230

      1 Outliers found.


          * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Residuals Statistics:

              Min       Max      Mean   Std Dev   N

*PRED       .1271    1.6461    .5531     .2903  55
*RESID     -.3336     .5230    .0000     .1664  55
*ZPRED    -1.4676    3.7657    .0000    1.0000  55
*ZRESID   -1.9862    3.1137    .0000     .9907  55

Total Cases =       55
```

El gráfico D que da cuenta de la dispersión de la discriminación muestra que los puntos tienden a agruparse en una recta. Al complementar esta información con los resultados de la regresión simple (outliers) se concluye que sólo el item 13 presenta sesgo.

Salidas outliers.

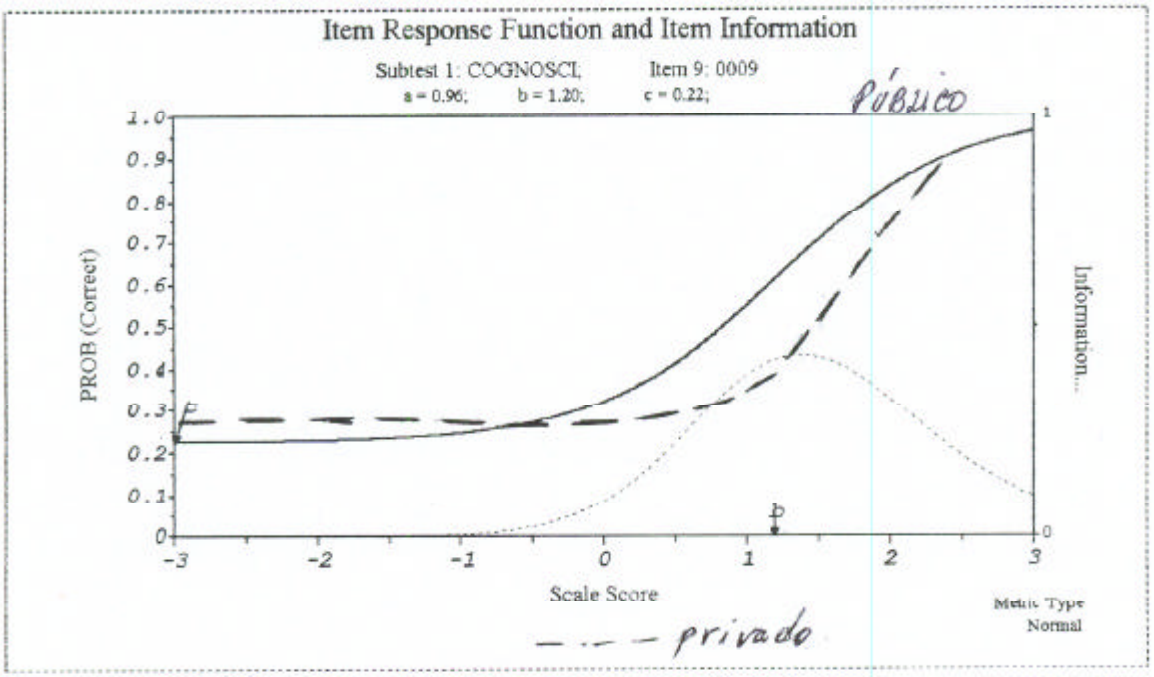
```
          * * * *   M U L T I P L E   R E G R E S S I O N   * * * *

Equation Number 1    Dependent Variable..   DIFPRIV

Casewise Plot of Standardized Residual

Outliers = 3.    *: Selected   M: Missing

          -6.     -3.  3.     6.
  Case #   O:.......:  :.......:O    DIFPRIV       *PRED      *RESID
      13   .        * ..         .      -6.66     -3.4444     -3.2156

       1 Outliers found.


          * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Residuals Statistics:

              Min      Max     Mean  Std Dev   N

*PRED     -5.6818   2.6618   -.5145   1.5629  55
*RESID    -3.2156   2.7886    .0000    .9220  55
*ZPRED    -3.3062   2.0323    .0000   1.0000  55
*ZRESID   -3.4550   2.9963    .0000    .9907  55

Total Cases =       55
```

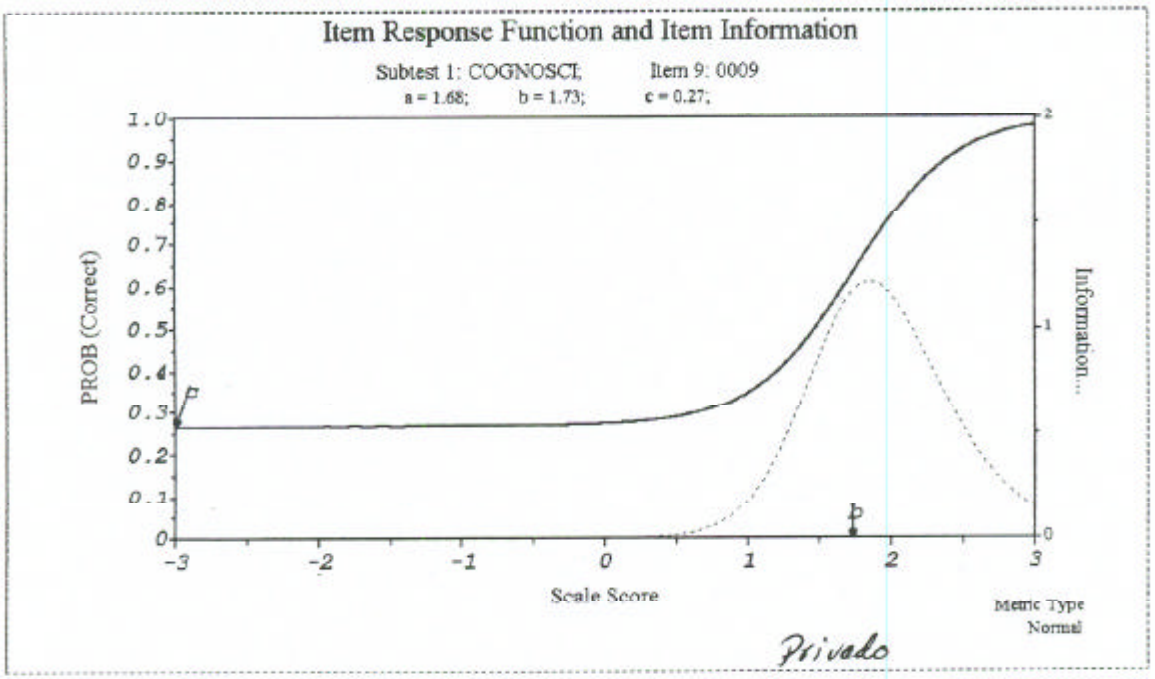
Análisis de la presencia de sesgo en los ítemes.

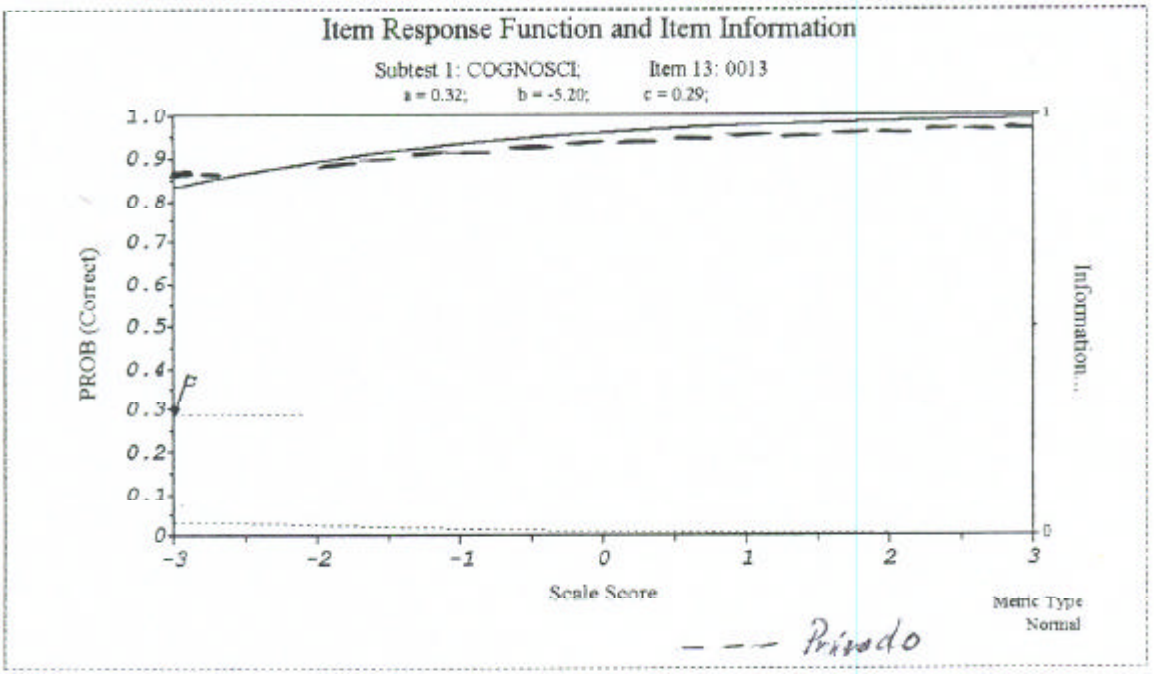
De acuerdo con la discriminación.

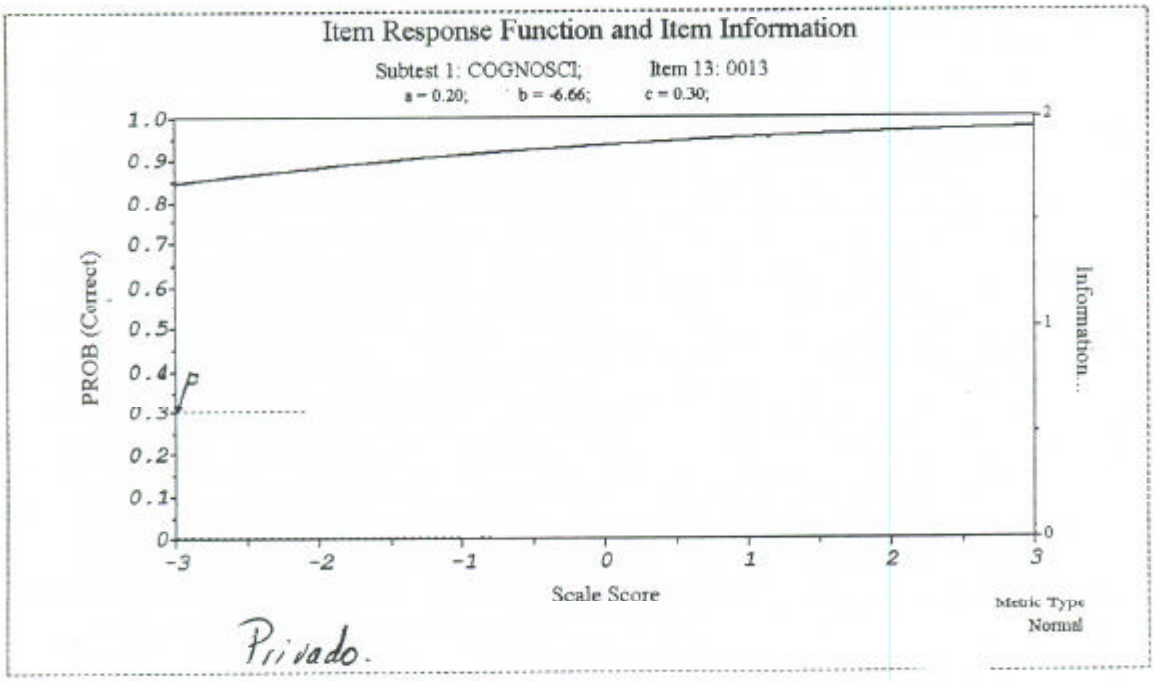
Item 9

Es de alta calidad técnica general. Presenta un sesgo que favorece a los colegios privados en los niveles bajos de habilidad. En los niveles ligeramente difícil y difícil presenta un sesgo que favorece a los colegios públicos. Para ambas subpoblaciones es un ítem de alta calidad técnica. Es un item difícil especialmente para los colegios privados.


Item Response Function and Item Information
Subtest 1: COGNOSCI;
Item 9: 0009
a = 0.96;
b = 1.20;
c = 0.22;
público
PROB (Correct)
Information...
Scale Score
Metric Type
Normal
privado


Item Response Function and Item Information
Subtest 1: COGNOSCI;
Item 9: 0009
a = 1.68;
b = 1.73;
c = 0.27;
PROB (Correct)
Information...
Scale Score
Metric Type
Normal
Privado


Item Response Function and Item Information
Subtest 1: COGNOSCI;
Item 13: 0013
a = 0.32;
b = -5.20;
c = 0.29;
PROB (Correct)
Information
Scale Score
Metric Type
Normal
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-3
-2
-1
0
1
2
3
P
Privado


Item Response Function and Item Information
Subtest 1: COGNOSCI;
Item 13: 0013
a = 0.20;
b = -6.66;
c = 0.30;
PROB (Correct)
Information
Scale Score
Metric Type
Normal
Privado.

De acuerdo con la dificultad.

Item 13

Este item no da suficiente información y no presenta suficiente poder discriminatorio. Situación que se evidencia también en la calidad técnica de ambas subpoblaciones. Es un item muy fácil. Presenta un sesgo que favorece levemente a los colegios públicos. No se pueden inferir posibles causas del comportamiento diferenciado por ser un item que no da suficiente información y no presentar suficiente poder discriminatorio.

Sesgo según zona en la que se ubican los colegios

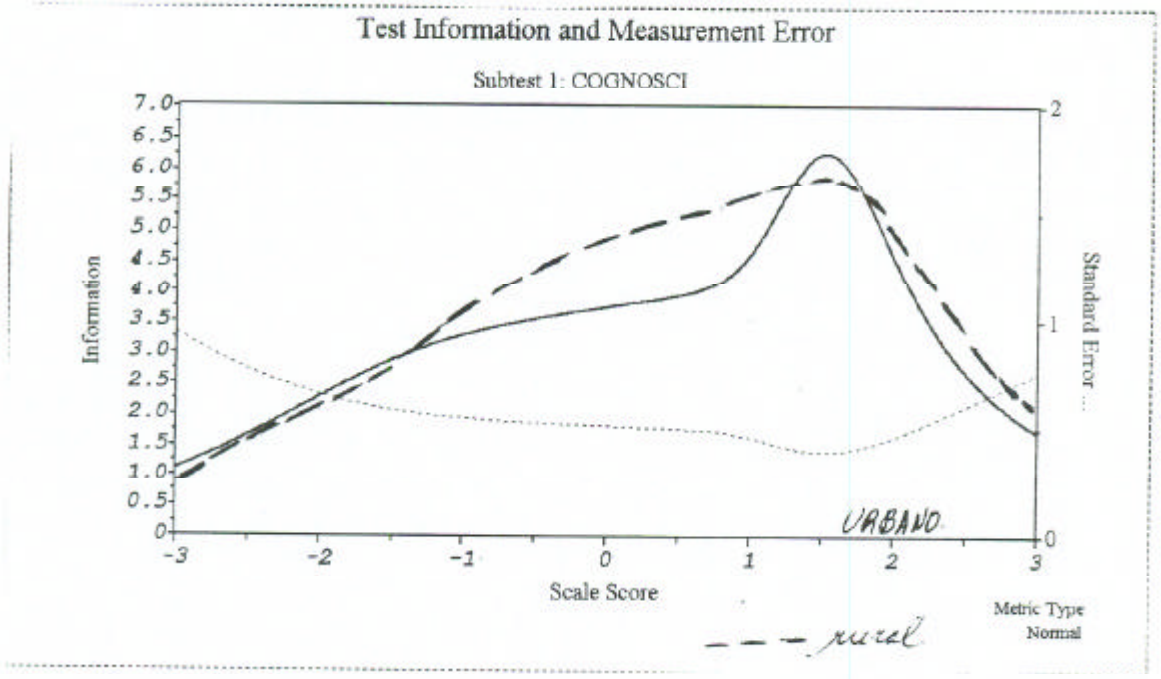
(rural/urbano)

Análisis de la función de información

Se refiere a la suma de las funciones de información de todos los ítemes. Los gráficos muestran curvas no bien definidas, a lo esperado, semejantes a una curva normal.

En el caso de los colegios urbanos hay una mayor información en los niveles de habilidad altos, debido a que está orientada hacia la derecha. El item once en los colegios urbanos es el que hace que una parte de la curva presente un mayor empinamiento, justamente en los niveles altos, brindando mayor información. Este mismo ítem tiene una forma más achatada para el caso de los colegios rurales, lo que conlleva a que esa curva presente un menor empinamiento, o sea más achatada en ese nivel de habilidad, con un menor nivel de información en el mismo.

En general los ítemes tienen un nivel de discriminación en niveles intermedios.


Test Information and Measurement Error
Subtest 1: COGNOSCI
Information
Standard Error
Scale Score
Metric Type
Normal
URBANO
rural

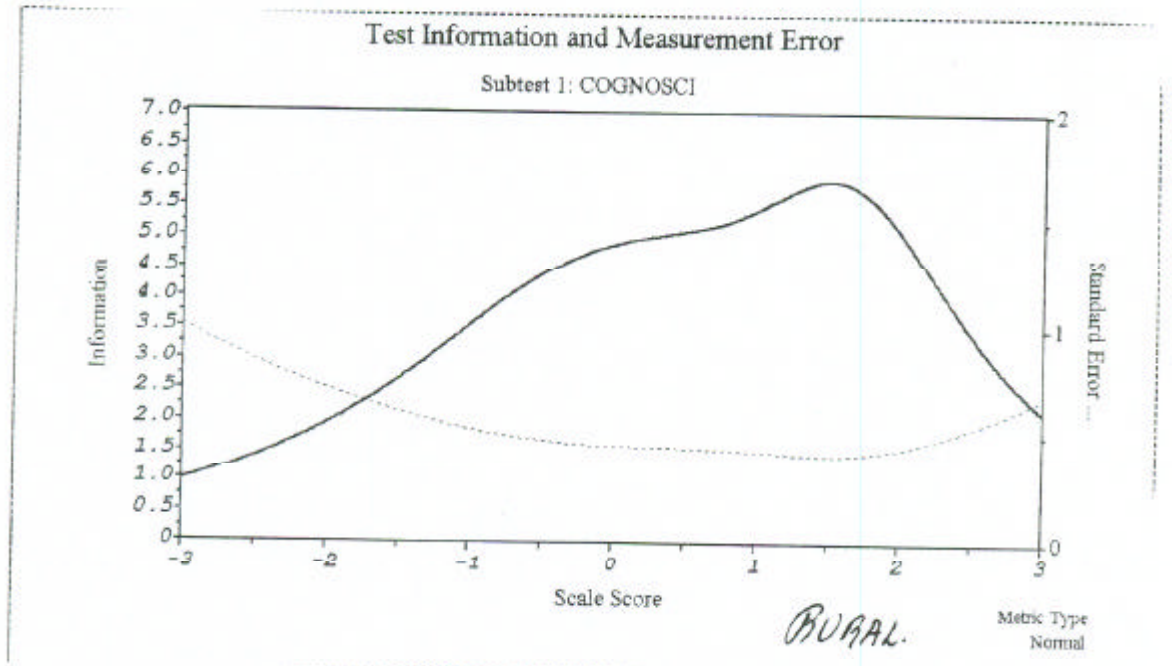

Test Information and Measurement Error
Subtest 1: COGNOSCI
Information
Standard Error
Scale Score
Metric Type
Normal
RURAL.

Tabla 10

c:\programs\spsswin\karia.sav

	discurba	difurb	discrura	difrural
1	.79	1.40	1.36	1.47
2	.55	.55	1.02	1.07
3	.28	-6.65	.44	-3.37
4	1.20	1.36	.64	1.28
5	.45	.46	.70	.74
6	.81	.36	.67	.06
7	.60	-2.49	.48	-2.14
8	.88	-.05	.92	.07
9	1.54	1.54	1.41	1.66
10	.92	.49	.66	.93
11	2.21	1.44	.24	1.58
12	.73	2.51	.43	2.53
13	.20	-6.86	.34	-4.04
14	.89	.64	.52	1.20
15	.68	-.34	.60	-.09
16	.66	-.54	.47	-.55
17	.56	-2.77	.72	-1.82
18	.44	-.80	.69	-.14
19	.41	-1.68	.52	-.71
20	.39	-.77	.42	.17
21	.62	-.14	.38	.27
22	.26	-2.96	.41	-1.54
23	.37	-1.00	.40	-.54
24	.17	-2.16	.18	-1.67
25	.29	-1.88	.34	-1.46
26	.43	-2.35	.37	-2.33
27	.16	-7.49	.40	-2.43
28	.21	-1.80	.29	-1.00
29	.80	-1.26	.76	-.59
30	.72	-2.44	.45	-2.84
31	.73	-2.02	.69	-1.36
32	.55	-.85	.62	-.03
33	.28	-2.15	.30	-1.53
34	.40	1.10	.43	1.07

c:\programs\spsswin\karla.sav

	discurba	difurb	discrura	difrural
35	.68	.76	.81	1.36
36	.59	.09	.63	.78
37	.38	-.46	.52	.26
38	.45	-.33	.48	.34
39	.28	-5.75	.58	-2.54
40	.28	.22	.35	.38
41	.75	-1.27	.76	-.53
42	.22	1.12	.33	.67
43	.52	1.91	.36	.70
44	.52	1.47	.54	2.10
45	.62	2.25	.39	2.79
46	.45	-.05	.60	.64
47	.21	-.99	.24	-.02
48	.40	-.06	.26	.52
49	.17	-1.87	.18	-2.53
50	.78	-2.11	.72	-1.60
51	.40	.52	.64	.51
52	.29	-.05	.36	.55
53	.56	-.73	.77	-.22
54	.92	-1.32	1.01	-.65
55	.50	1.01	.54	1.69

La tabla 10 presenta los resultados de la discriminación y la dificultad según las salidas del programa Bilog. Con esos datos se construyeron los gráficos de dispersión.

GRÁFICO E

Gráfico de dispersión de la **dificultad**
rurales (y) y urbanos(x) de la Prueba de Bachillerato
En Español. 1997.

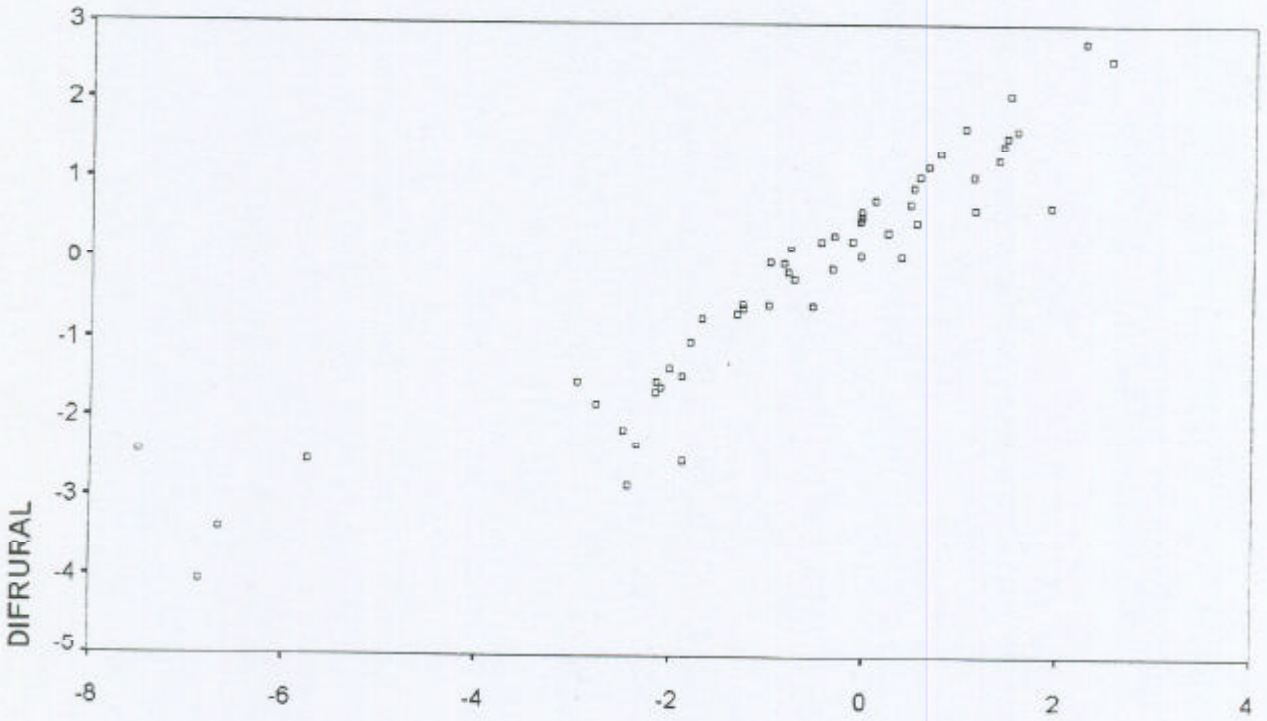
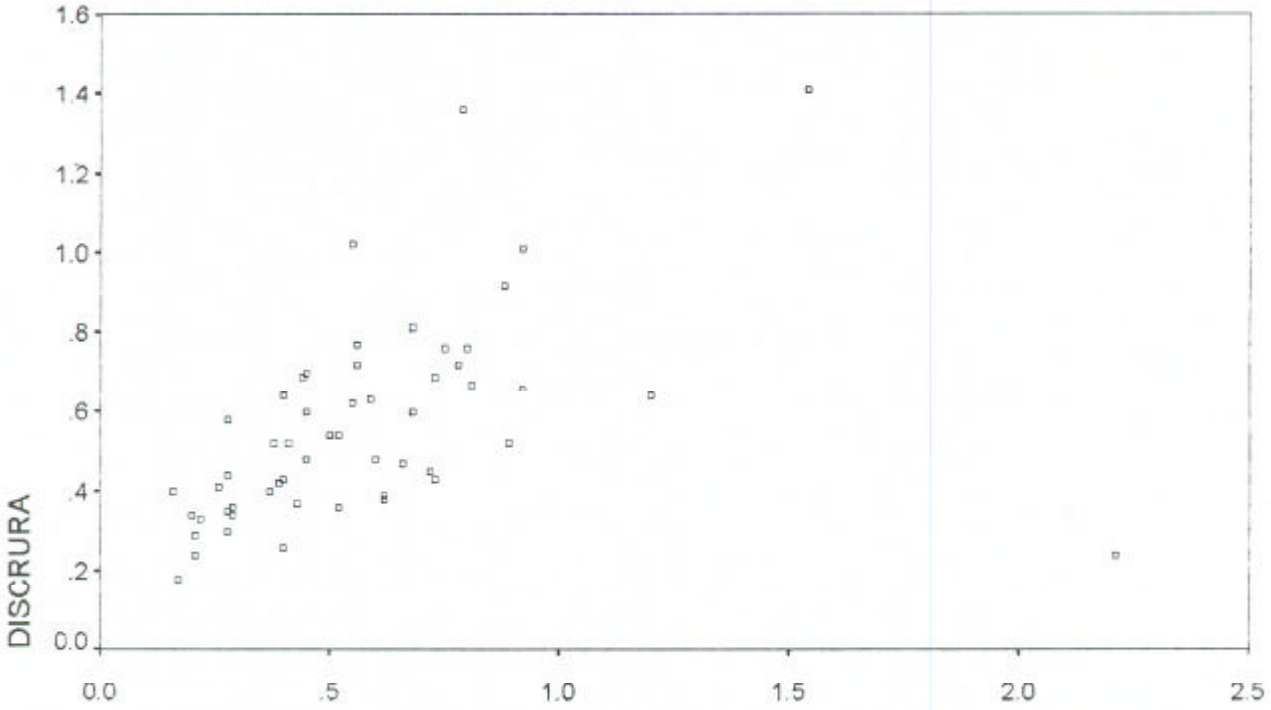


GRÁFICO F

Gráfico de dispersión de la **discriminación**
De rurales (y) y urbanos(x) de la Prueba de Bachillerato
En Español. 1997.



El gráfico E de dificultad muestra que los puntos de cada item presentan una tendencia de agrupamiento. Debido a la dificultad que presenta el gráfico para definir cuales son los ítemes más alejados, con probabilidades de sesgo, se procedió a complementar la información con los resultados de la aplicación de la regresión simple (outliers), para poder determinar los ítemes que evidencian sesgo.

De este análisis se obtiene que el item que presenta sesgo es el **27**.

```
Equation Number 1    Dependent Variable..   DIFRURAL

Casewise Plot of Standardized Residual

Outliers = 3.    *: Selected   M: Missing

          -6.      -3.  3.      6.
  Case #   O:........:  :........:O   DIFRURAL       *PRED       *RESID
     27    .           .. *        .      -2.43     -4.4919      2.0619

       1 Outliers found.
```

* *

```
Residuals Statistics:

              Min       Max      Mean   Std Dev    N

*PRED      -4.4919    1.9485   -.1978    1.3845   55
*RESID     -1.6576    2.0619    .0000     .6110   55
*ZPRED     -3.1016    1.5503    .0000    1.0000   55
*ZRESID    -2.6878    3.3435    .0000     .9907   55

Total Cases =       55
```

El gráfico F brinda información de la dispersión de los datos en el análisis de discriminación, muestra que los puntos de cada item tienden a agruparse. Para establecer cuales itemes presentan sesgo se procedió a aplicar la regresión simple (outliers) de cuyo análisis se precisa que sólo los ítemes 1 y 11 presentan sesgo.

Salidas de outliers.

```
16 Dec 98 SPSS for MS WINDOWS Release 6.1                              Page 12

          * * * *   M U L T I P L E   R E G R E S S I O N   * * * *

Equation Number 1    Dependent Variable..   DISCRURA

Casewise Plot of Standardized Residual

Outliers = 3.    *: Selected   M: Missing

          -6.      -3.  3.      6.
  Case #   O:........:  :........:O   DISCRURA       *PRED       *RESID
      1    .          ..*         .      1.36       .6030       .7370
     11    .        * ..          .       .24      1.0762      -.8362

       2 Outliers found.
```

\- - - - - - - - - - - - - -

Residuals Statistics:

	Min	Max	Mean	Std Dev	N
*PRED	.4219	1.0762	.5516	.1126	55
*RESID	-.8362	.7370	.0000	.2290	55
*ZPRED	-1.1515	4.6574	.0000	1.0000	55
*ZRESID	-3.6184	3.1889	.0000	.9907	55

Total Cases = 55

Análisis de la presencia de sesgo en los ítemes

De acuerdo con la dificultad:

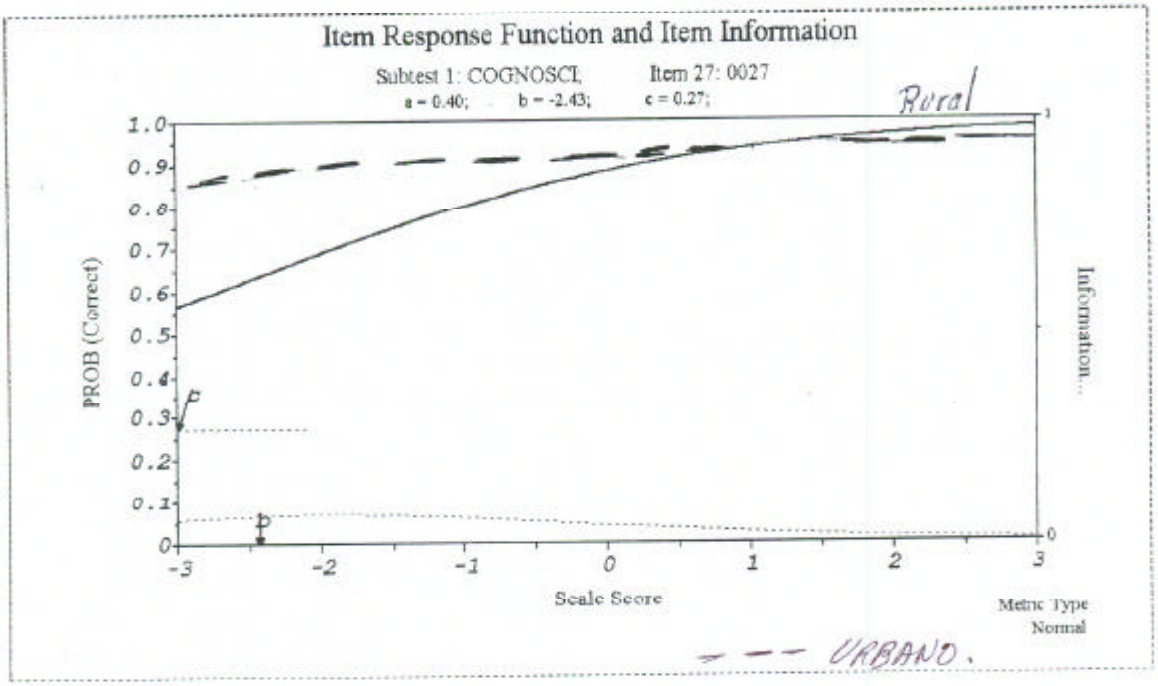
Item 27

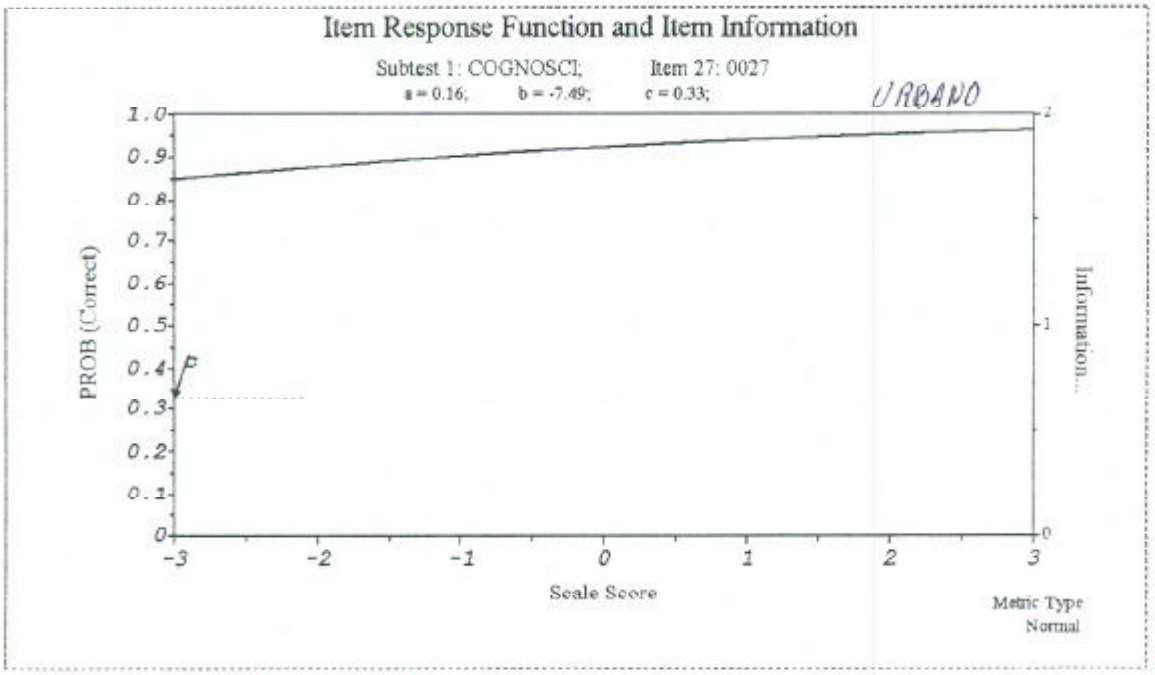
Es un ítem que no da suficiente información y no presenta suficiente poder discriminatorio en general; en las dos subpoblaciones también presenta el mismo comportamiento. Es un item muy fácil sobretodo para los y las estudiantes de colegios urbanos. Presenta un sesgo que favorece a los colegios urbanos en los niveles bajos y un sesgo muy leve que favorece a los estudiantes de colegios rurales en los niveles altos de habilidad.

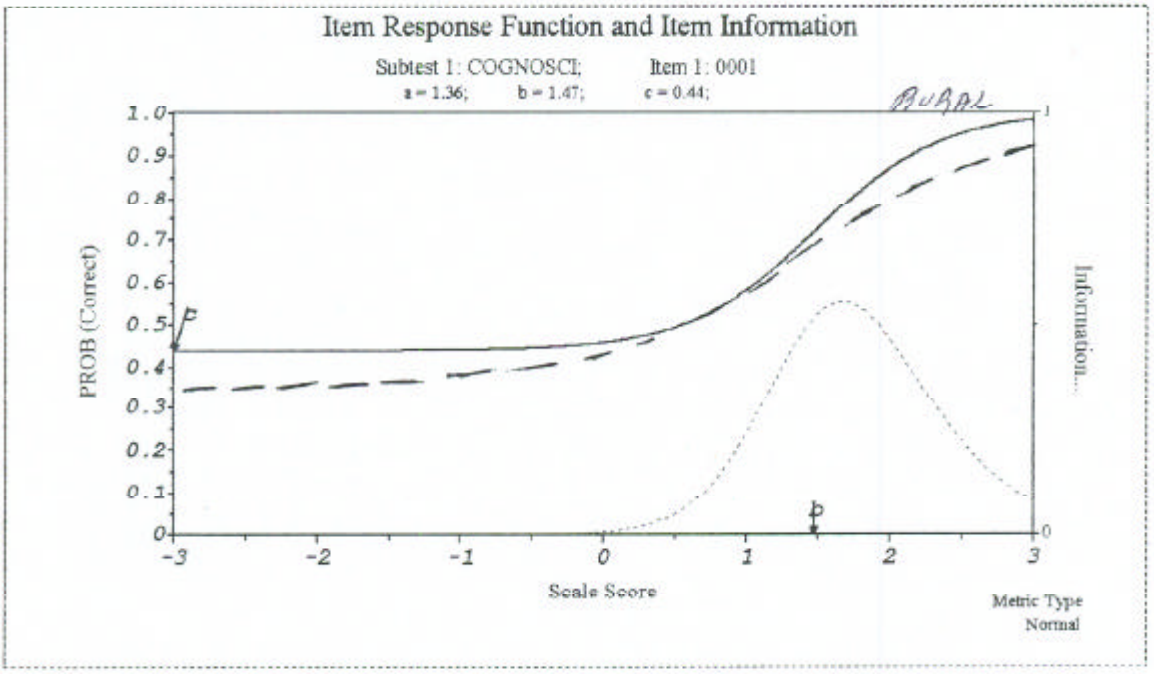
De acuerdo con la discriminación:

Item 1

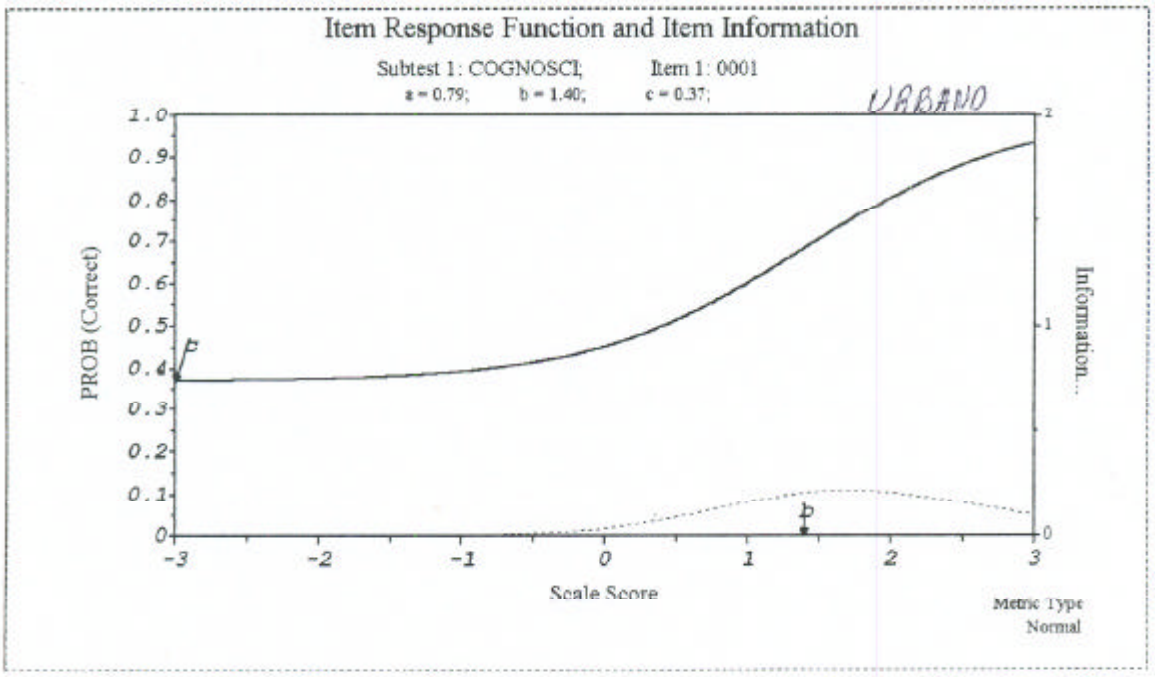
Presenta un sesgo que favorece a los colegios rurales en niveles bajos y altos de habilidad. Es un ítem que a nivel general brinda una aceptable información y presenta un aceptable poder discriminatorio. Evidencia que es un ítem más difícil para estudiantes de colegios urbanos, acompañado de un alto azar.en ambas subpoblaciones es un item que discrimina en niveles altos.

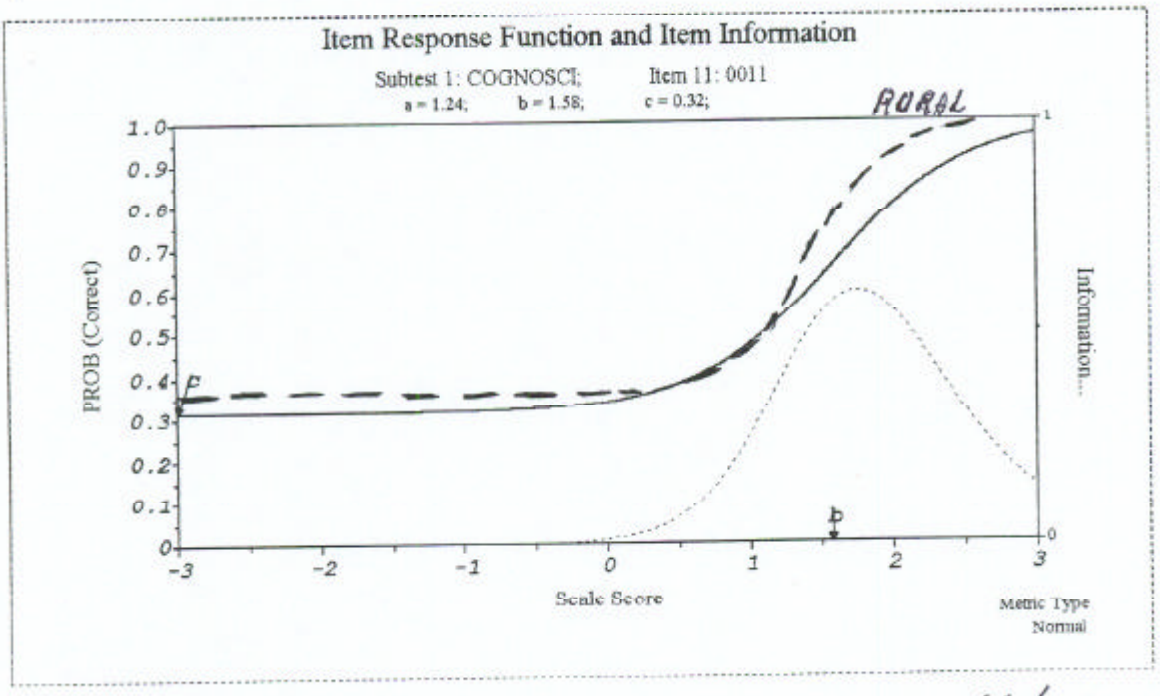

Item Response Function and Item Information
Subtest 1: COGNOSCI,
Item 27: 0027
a = 0.40;
b = -2.43;
c = 0.27;
Rural
PROB (Correct)
Information
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
c
b
-3
-2
-1
0
1
2
3
Scale Score
Metric Type
Normal
URBANO.


Item Response Function and Item Information
Subtest 1: COGNOSCI;
Item 27: 0027
a = 0.16;
b = -7.49;
c = 0.33;
URBANO
PROB (Correct)
Information
Scale Score
Metric Type
Normal

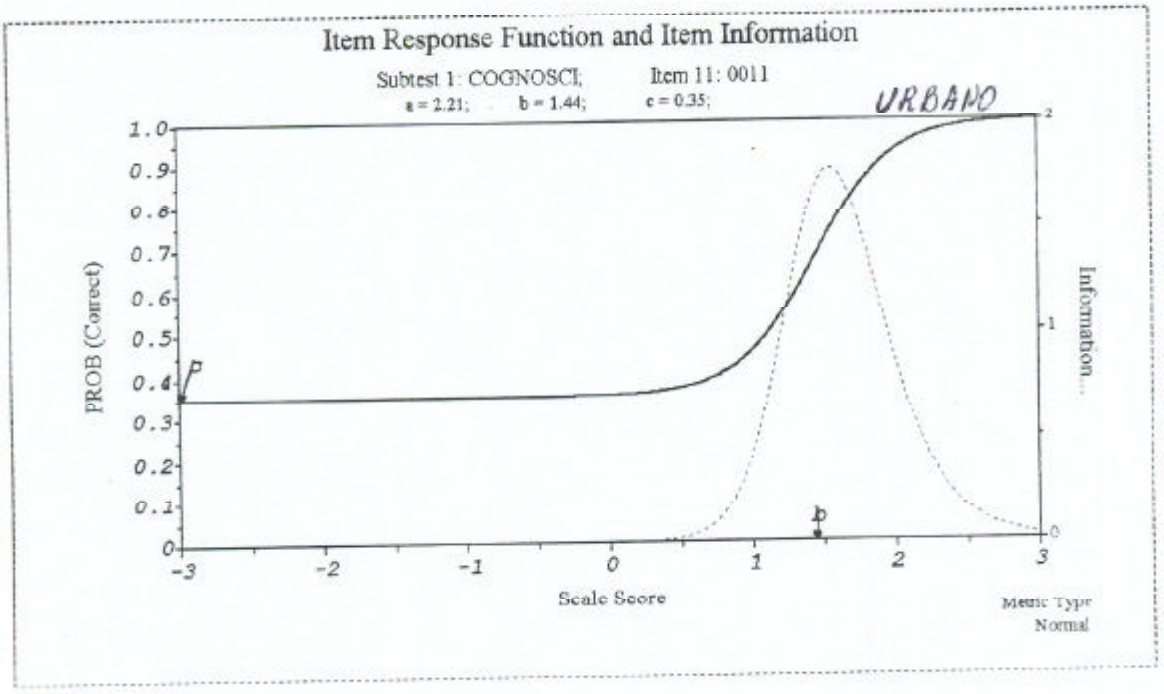

Item Response Function and Item Information
Subtest 1: COGNOSCI;
Item 1: 0001
a = 1.36;
b = 1.47;
c = 0.44;
RURAL
PROB (Correct)
Information
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
c
b
-3
-2
-1
0
1
2
3
Scale Score
Metric Type
Normal

URBANO


Item Response Function and Item Information
Subtest 1: COGNOSCI;
Item 1: 0001
a = 0.79;
b = 1.40;
c = 0.37;
URBANO
PROB (Correct)
Information
Scale Score
Metric Type
Normal


Item Response Function and Item Information
Subtest 1: COGNOSCI;
Item 11: 0011
a = 1.24;
b = 1.58;
c = 0.32;
RURAL
PROB (Correct)
Information
Scale Score
Metric Type
Normal
c
b

--- Urbano


Item Response Function and Item Information
Subtest 1: COGNOSCI;
Item 11: 0011
a = 2.21;
b = 1.44;
c = 0.35;
URBANO
PROB (Correct)
Information
Scale Score
Metric Type
Normal

Item11

Es un item de alta calidad técnica genral. Presenta sesgo a favor de los colegios urbanos en los niveles bajos y sesga a favor de los colegios urbanos en los niveles altos. Para ambas subpoblaciones es un item de alta calidad técnica. En ambos casos discrimina en niveles altos.

Este item es el que le da a la curva general de la función de información de la prueba, una forma poco usual, presente esta figura de manera parecida en las curvas de función de información de cada una de las subpoblaciones analizadas.

Los comportamientos diferenciales de los ítemes expresan que un total de siete ítemes presentan sesgo (un13 %) de los cuales ninguno evidencia sesgo en más de un ámbito, como se observa en el cuadro siguiente.

Ámbito de sesgo	Número de ítemes	Total de ítemes
Sexo	7 – **49**	2
Tipo de colegio (Público) Privado)	**9** – 13	2
Zona de ubicación del colegio (rural, urbano)	27 – **1** – **11**	3
Total		7

Los ítemes subrayados corresponden a la discriminación lo que indica que hay una diferencia de un item entre la dificultad y la discriminación.

Es importante agregar que dos ítemes que en el análisis general presentaron una alta calidad técnica, porque dan suficiente información y brindan suficiente poder discriminatorio, presentan un comportamiento diferencial. El ítem nueve lo presenta en el análisis referido a tipo de colegio, cuyo sesgo favorece a los colegios privados en niveles de habilidad bajos y a colegios públicos en niveles de habilidad altos.

El otro item es el once el cual presenta un comportamiento diferencial referido a la zona de ubicación del colegio. Es un item sesgado a favor de colegios urbanos en niveles bajos y altos.

El item uno presenta sesgo y además es un item que brinda una aceptable información y presenta un aceptable poder discriminatorio. a favor de colegios rurales

El resto de ítemes que presentan sesgo en alguno de los ámbitos estudiados (7,49,13,27) son ítemes que no dan suficiente información, y además, no presentan suficiente poder discriminatorio. Son ítemes que deben ser revisados para mejorar su calidad técnica y para controlar ese comportamiento diferencial que evidencian.

Del total de ítemes (9) que presentan una aceptable información y tienen un aceptable poder discriminatorio, a nivel general, sólo uno presentó sesgo; son ítemes que podrían utilizarse en otras pruebas por evidenciar una aceptable calidad técnica. Específicamente estos ítemes son: 4,8,10,29,35,41,50,y el 54. Estos representan el 14% de la totalidad de ítemes de la prueba.

CONCLUSIONES Y RECOMENDACIONES

Generales

El trabajo permitió un acercamiento al manejo de la TRI, por medio del cual pueden reconocer las ventajas que ofrece la TRI para la elaboración de pruebas o instrumentos. Además, en relación con la TCT, tanto para el análisis de los ítemes ya que permite realizar pruebas con un mayor nivel de precisión.

Ese nivel de precisión facilita una toma de decisión más certera y permite seleccionar ítemes en los distintos niveles de habilidad para la elaboración de pruebas relacionadas con el constructo que se investiga. Esto contribuye en la atención de la urgente necesidad de crear un banco de ítemes que sustenten las pruebas de español en el nivel de bachillerato.

Aspectos relacionados con la prueba:

La prueba de acuerdo con el análisis clásico es una prueba unidimensional , pues la carga factorial está asociada a un factor. El cual tiene una varianza explicada de 7.8%

Para determinar la confiabilidad o el grado de precisión de la prueba se utilizó el Alfa de Cronbach, cuyo resultado fue de .7539 lo que indica que la prueba podria mejorarse, tal vez recurriendo a la creación de un banco de ítemes.

El índice de discriminación de cada uno de los ítemes según el análisis de la Teoría Clásica señala que los ítemes 8, 15, 29, 36, 41, 53, 54 cumplen con los requisitos estandar establecidos. Los ítemes restantes según este análisis deberían ser revisados. Para el caso de los ítemes 8, 29, 41, y 54 son cuatro ítemes que segun el analisis de calidad técnica de la TRI son ítemes aceptables por sus niveles aceptables de información y por su poder

discriminatorio aceptable. Además son ítemes que no presentan sesgo. O sea, son los mejores ítemes de la prueba.

De acuerdo con el análisis TRI el 80% de los ítemes de la prueba deben ser revisados para mejorarlos, pues no ofrecen suficiente información y cuentan con poco poder discriminatorio. El 20 % podría ser utilizado en futuras pruebas.

Sólo el 13% del total de ítemes de la prueba presenta sesgo. De este 13% el 57% lo expresa en la discriminación (4 ítemes), y el 43% en la dificultad (3 ítemes).

Los únicos dos ítemes que evidencian alta calidad técnica en la prueba, presentaron un comportamiento diferenciado, uno en el análisis correspondiente al tipo de colegio y el otro en el análisis de la zona de ubicación del colegio. Por lo que son ítemes que deben revisarse para corregir el sesgo.

El 57% de los ítemes que evidencian sesgo (4) son ítemes que además no dan suficiente información y presentan insuficiente poder discriminatorio. Son ítemes que deben ser doblemente revisados.

Del total de ítemes que a nivel general presentan una aceptable información y tienen un aceptable poder discriminatorio, sólo uno tiene sesgo. Estos ítemes podrían utilizarse en otras pruebas, por evidenciar una aceptable calidad técnica, estos representan el 14% de la totalidad de los ítemes. De lo anterior se concluye que el 20% de los ítemes aceptables que indica el análisis general del TRI, se reduce a un 14% pues el 6% restante presenta sesgo.

BIBLIOGRAFÍA

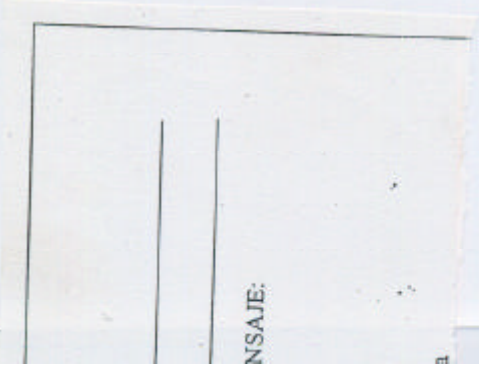
Anastasi, A. (1971). **Test Psicológicos.** Madrid, España: Aguilar Ediciones.

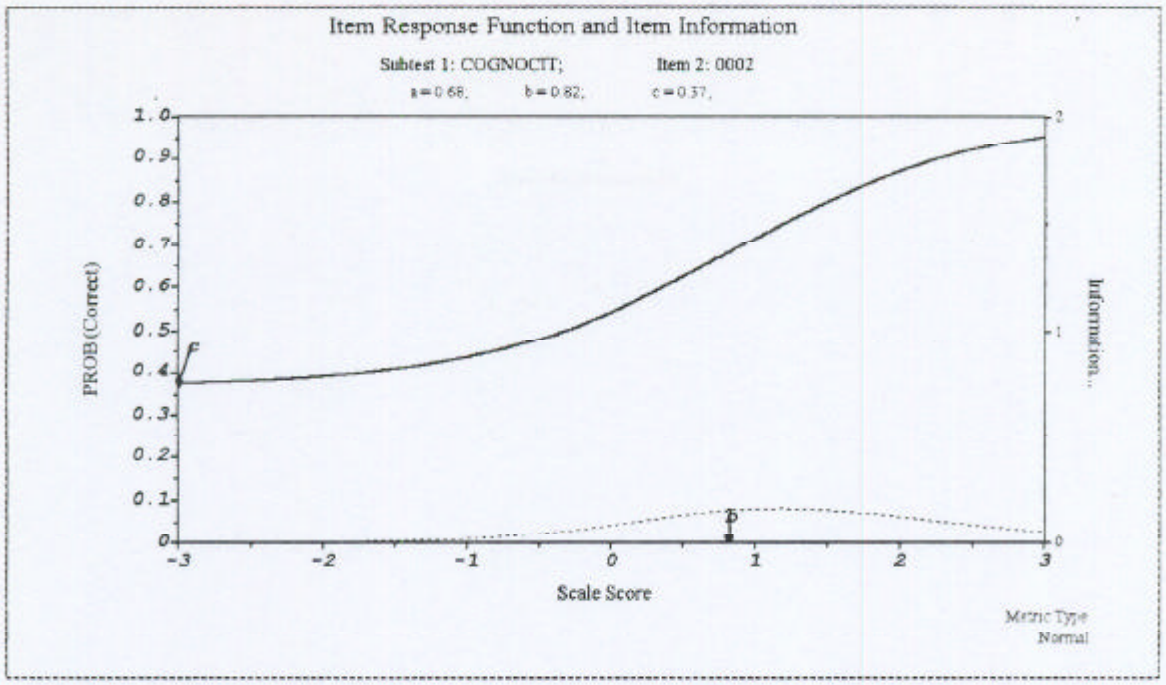
Muñiz – Fernández, J. (1990). **Teoría de Respuesta a los Items: un nuevo enfoque en evolución psicológica y educativa.** Madrid: Pirámide.

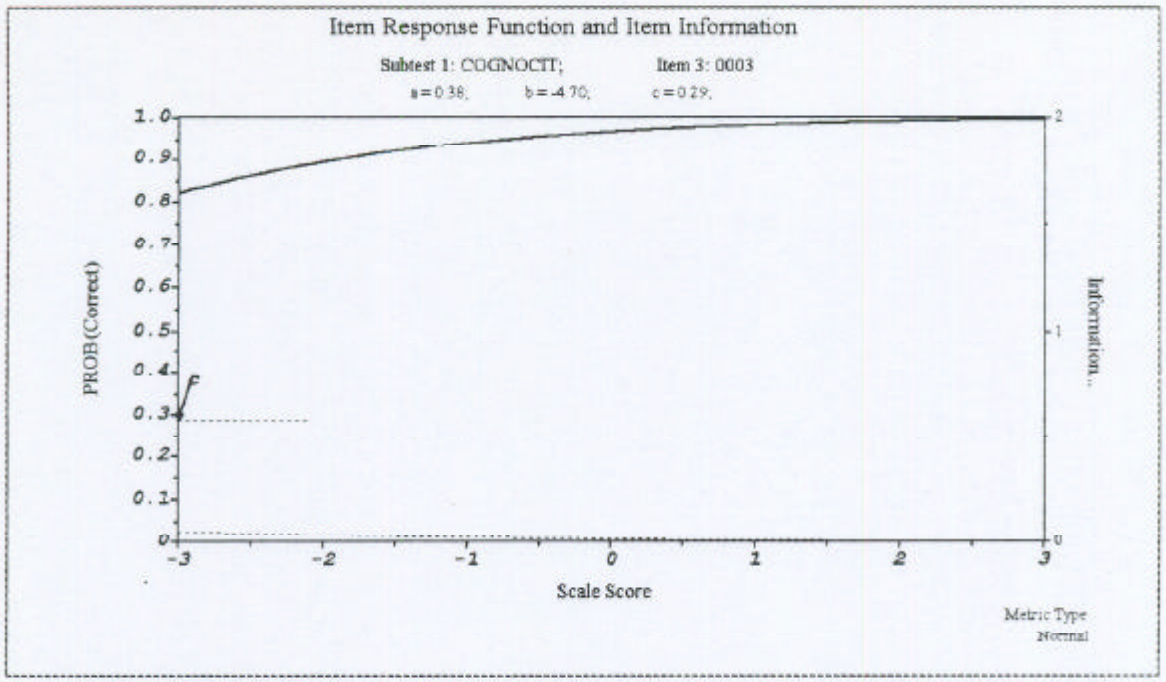
Nunnally, J. C. & Bernstein, I. J. (1995). **Teoría Psicométrica.** (3 ed.). México, D. F.: Editorial McGraw-Hill Latinoamericana.

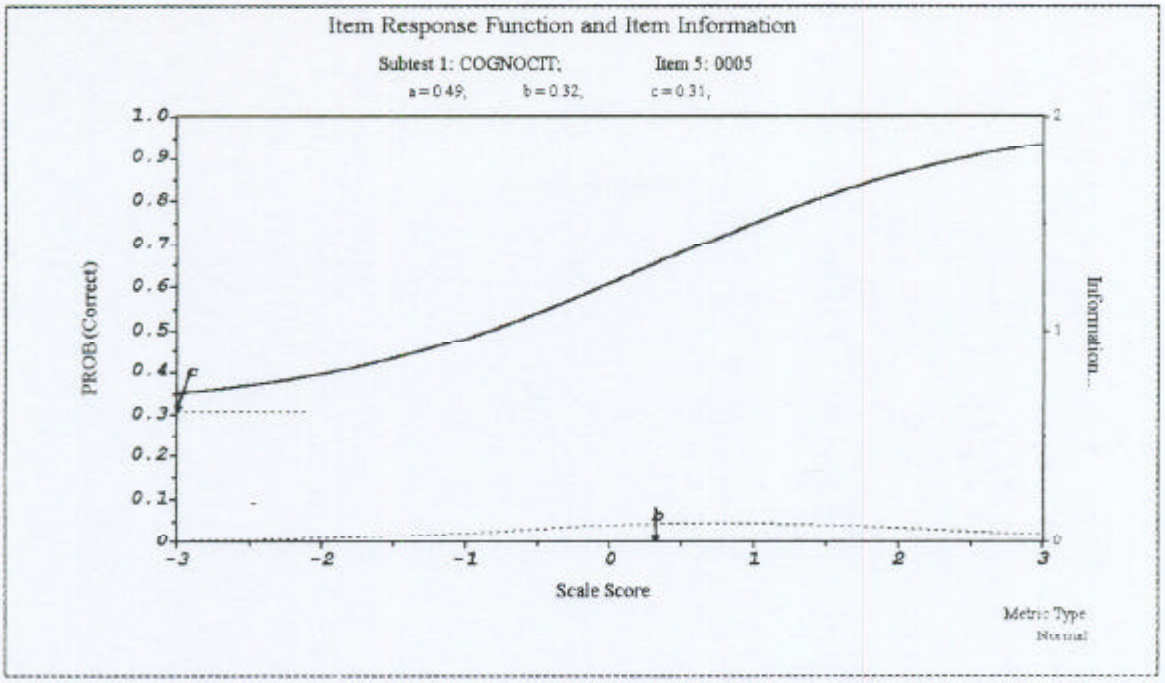
ANEXOS

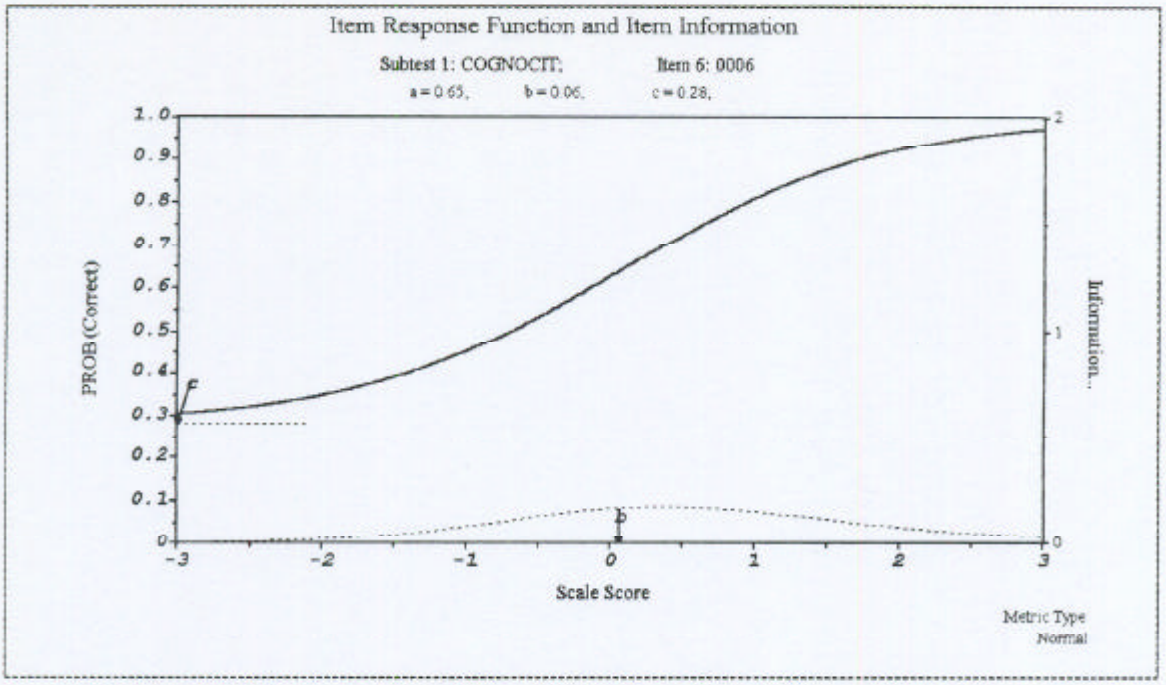
Ítems de baja calidad técnica

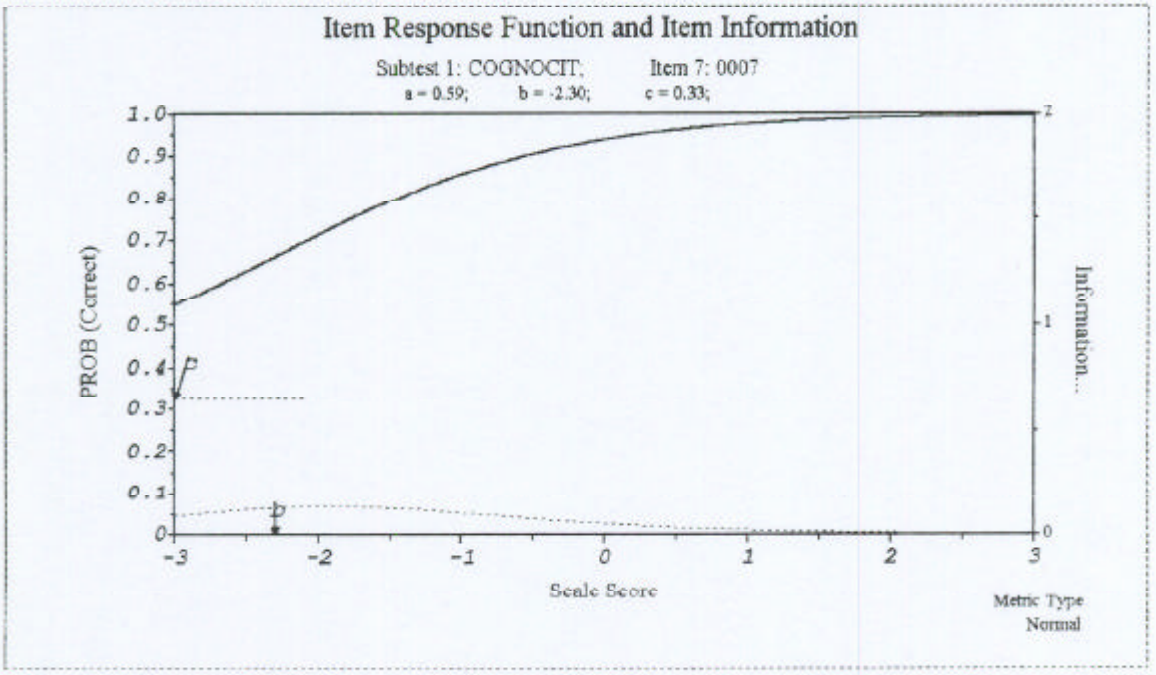


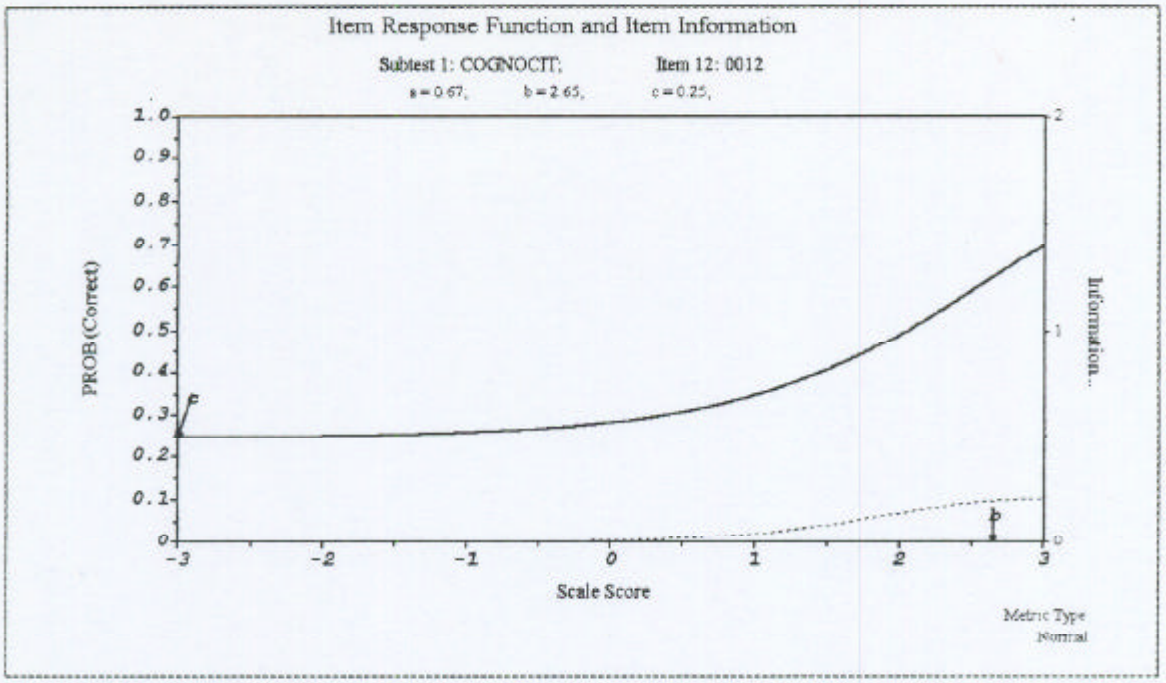

Item Response Function and Item Information
Subtest 1: COGNOCTT;
Item 2: 0002
a = 0.68;
b = 0.82;
c = 0.37;
PROB(Correct)
Information
Scale Score
Metric Type
Normal

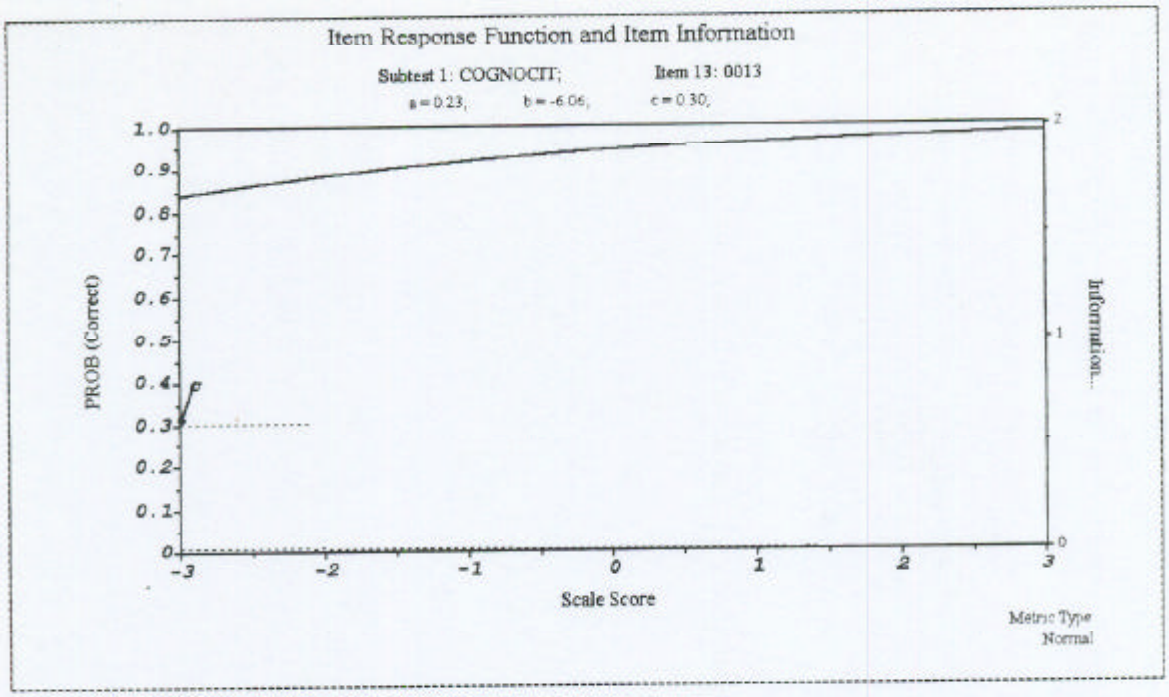

Item Response Function and Item Information
Subtest 1: COGNOCTT;
Item 3: 0003
a=0.38;
b=-4.70;
c=0.29;
PROB(Correct)
Information
Scale Score
Metric Type
Normal

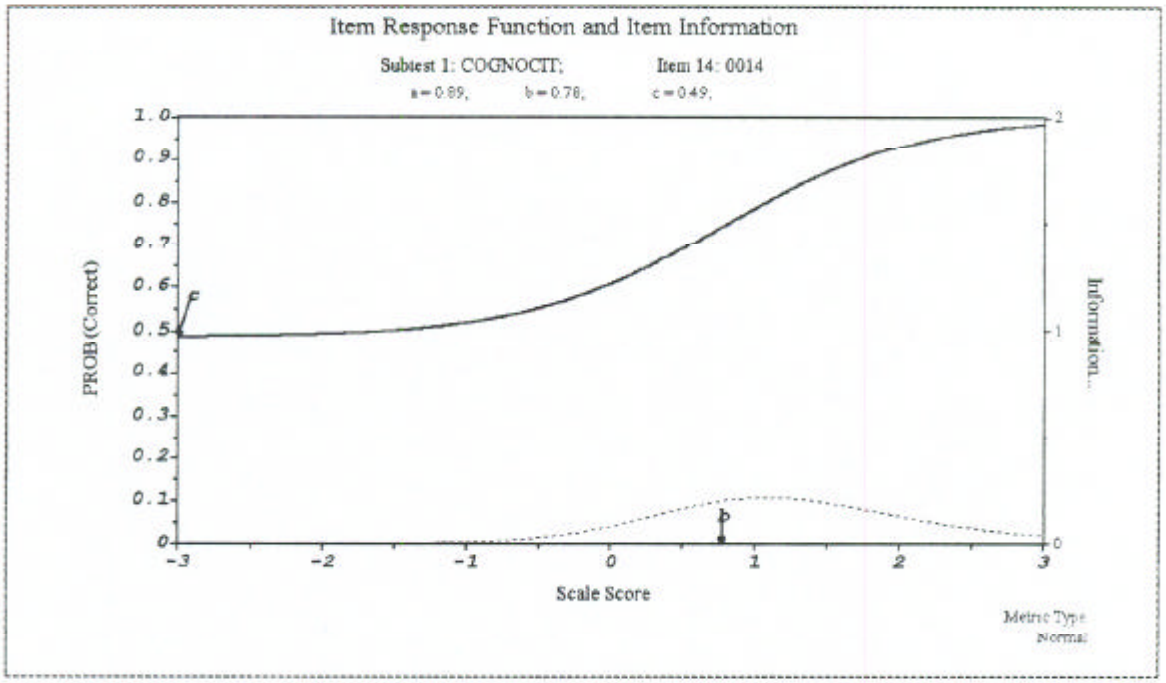

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 5: 0005
a=0.49;
b=0.32;
c=0.31;
PROB(Correct)
Information
Scale Score
Metric Type
Normal

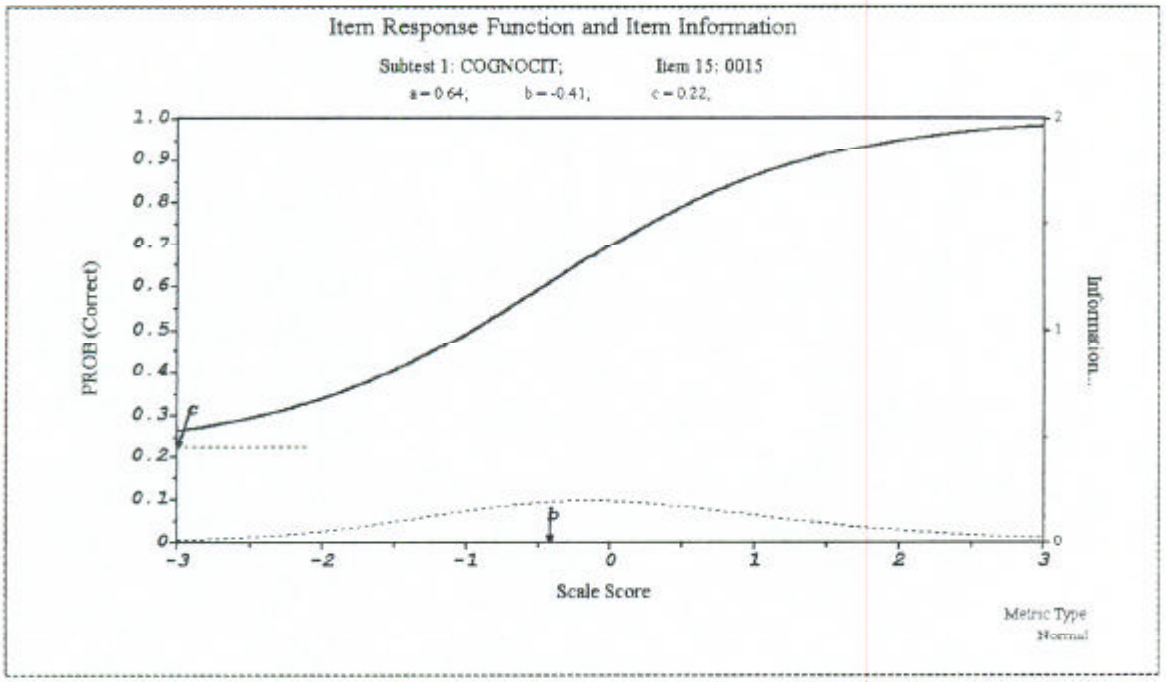

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 6: 0006
a = 0.65,
b = 0.06,
c = 0.28,
PROB (Correct)
Information
Scale Score
Metric Type
Normal

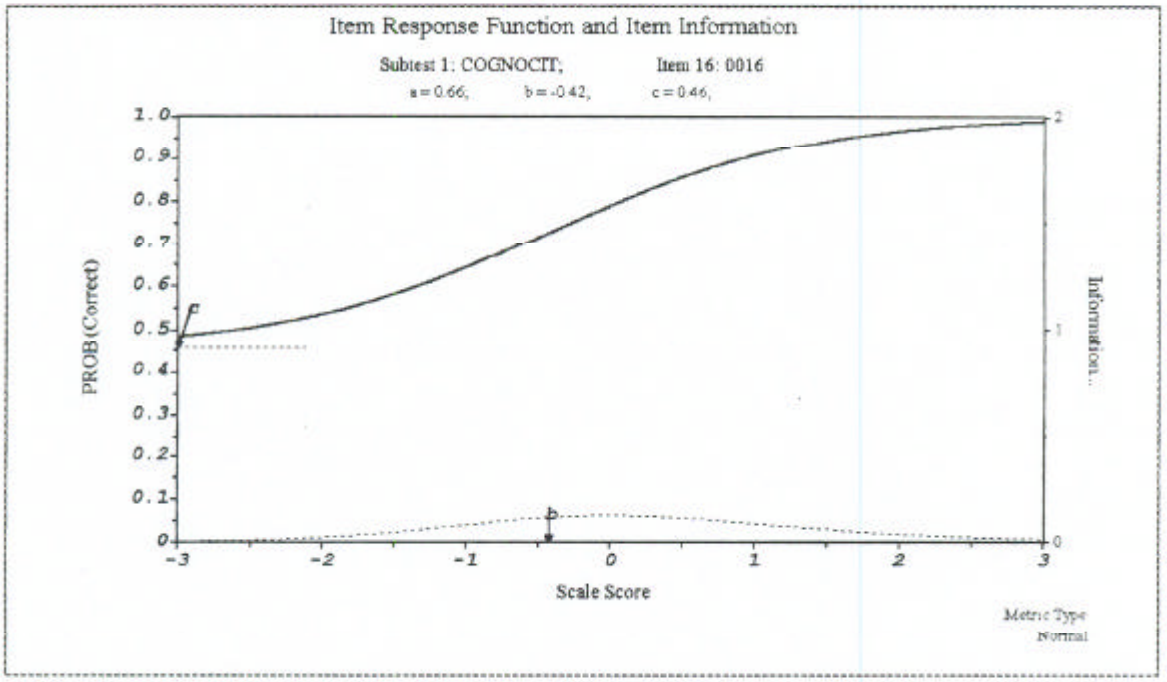

Item Response Function and Item Information
Subtest 1: COGNOCIT; Item 7: 0007
a = 0.59; b = -2.30; c = 0.33;
PROB (Correct)
Information
Scale Score
Metric Type
Normal
P
b

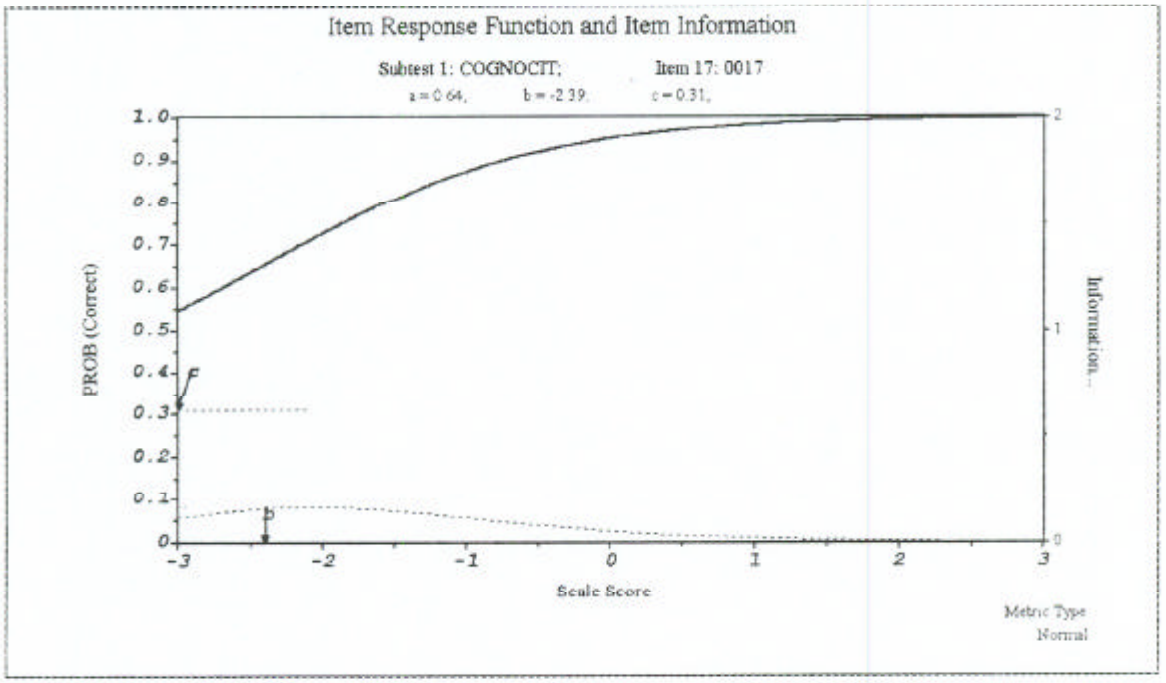

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 12: 0012
a = 0.67,
b = 2.65,
c = 0.25,
PROB (Correct)
Information
Scale Score
Metric Type
Normal

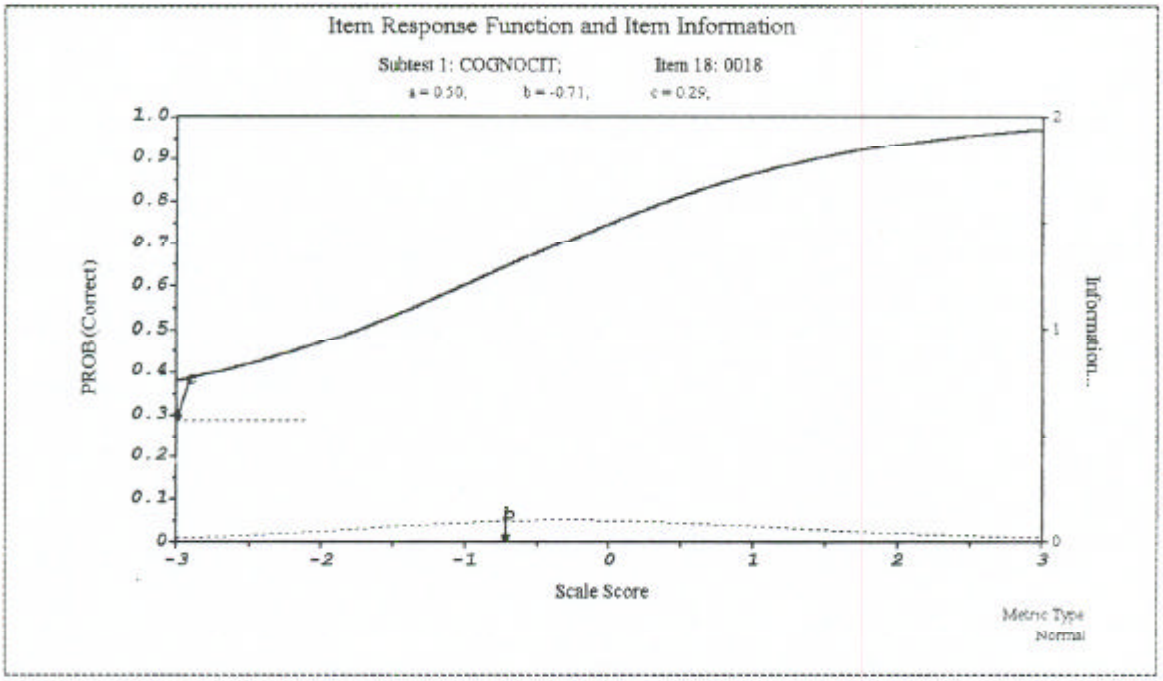

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 13: 0013
a = 0.23;
b = -6.06;
c = 0.30;
PROB (Correct)
Information
Scale Score
Metric Type
Normal

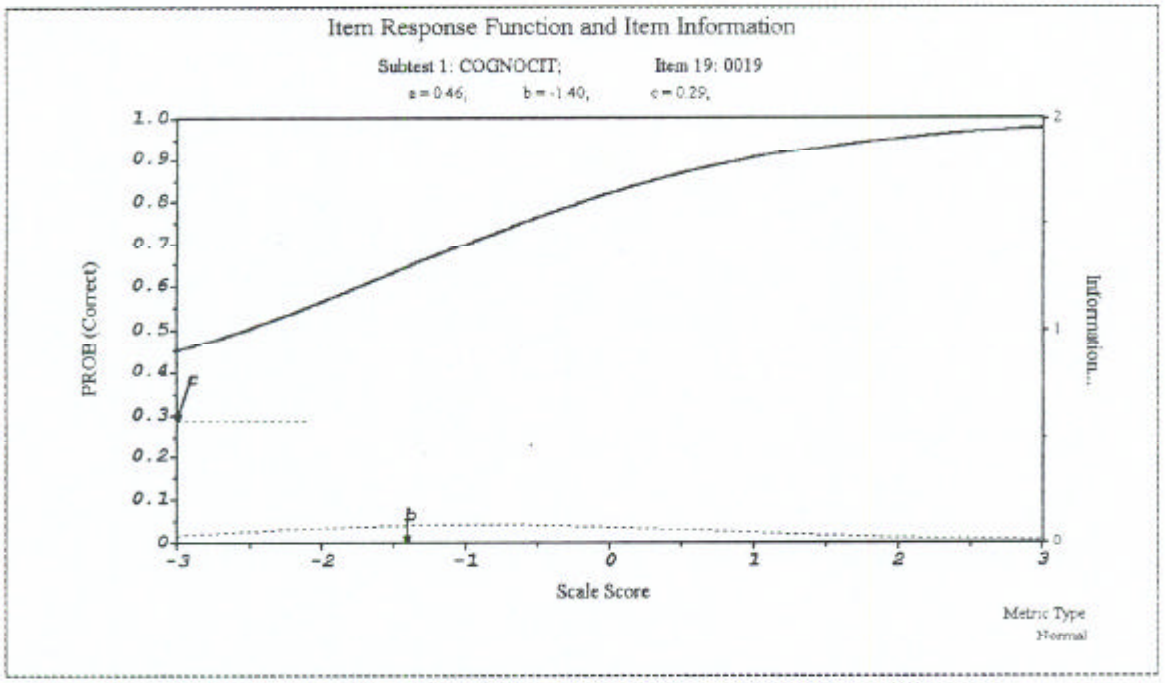

Item Response Function and Item Information
Subtest 1: COGNOCTT;
Item 14: 0014
a=0.89,
b=0.78,
c=0.49,
PROB (Correct)
Information
Scale Score
Metric Type
Normal

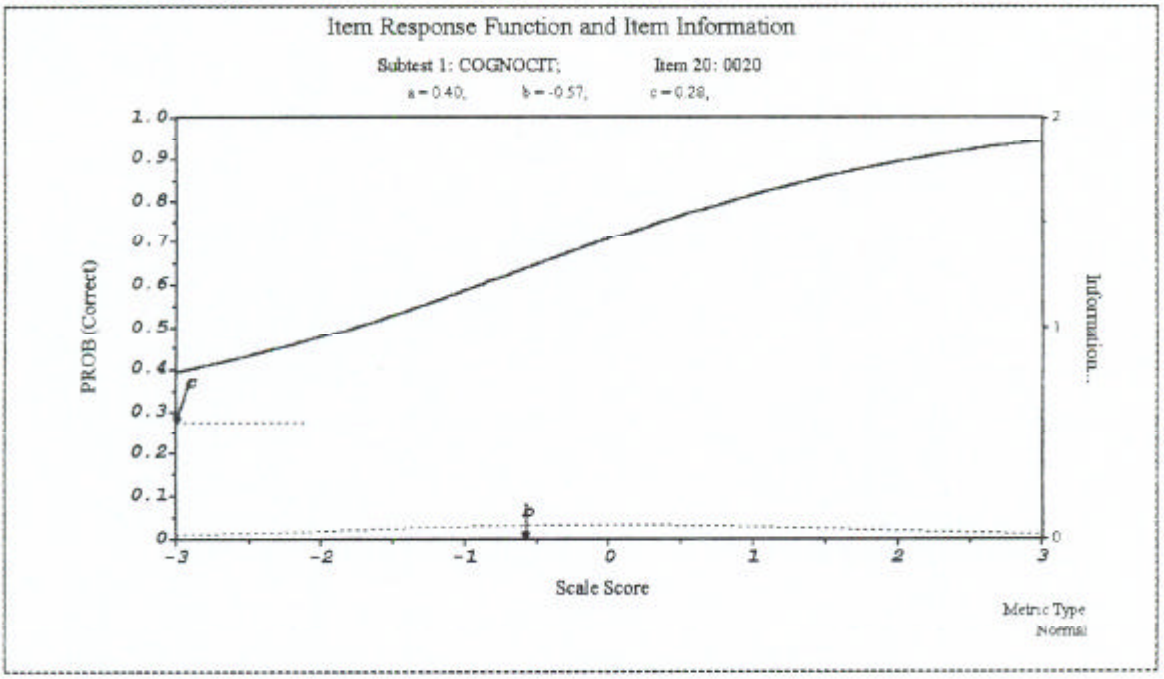

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 15: 0015
a = 0.64,
b = -0.41,
c = 0.22,
PROB (Correct)
Information
Scale Score
Metric Type
Normal

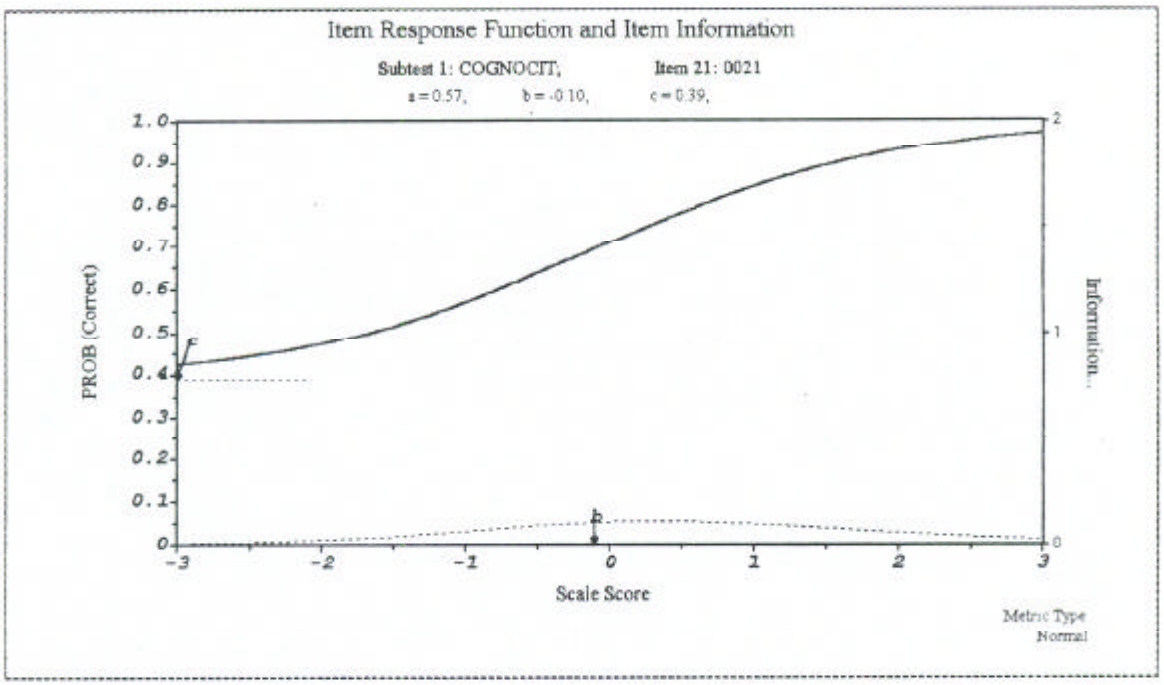

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 16: 0016
a = 0.66;
b = -0.42;
c = 0.46;
PROB (Correct)
Information
Scale Score
Metric Type
Normal

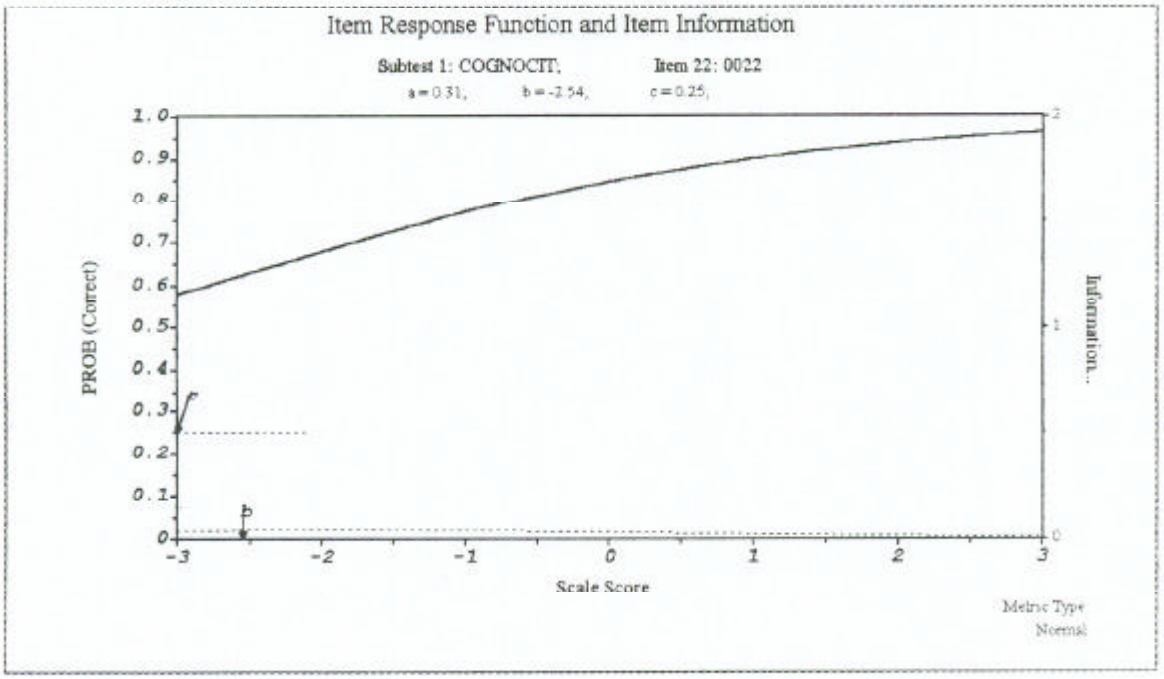

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 17: 0017
a = 0.64,
b = -2.39,
c = 0.31,
PROB (Correct)
Information
Scale Score
Metric Type
Normal

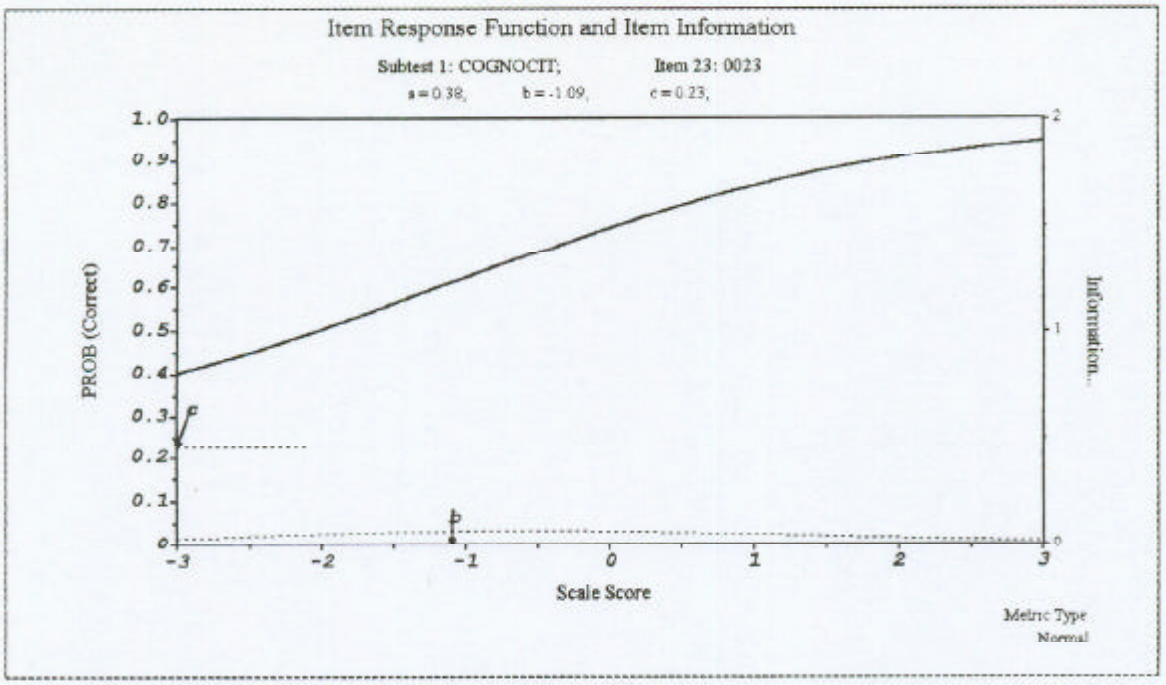

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 18: 0018
a = 0.50,
b = -0.71,
c = 0.29,
PROB (Correct)
Information...
Scale Score
Metric Type
Normal

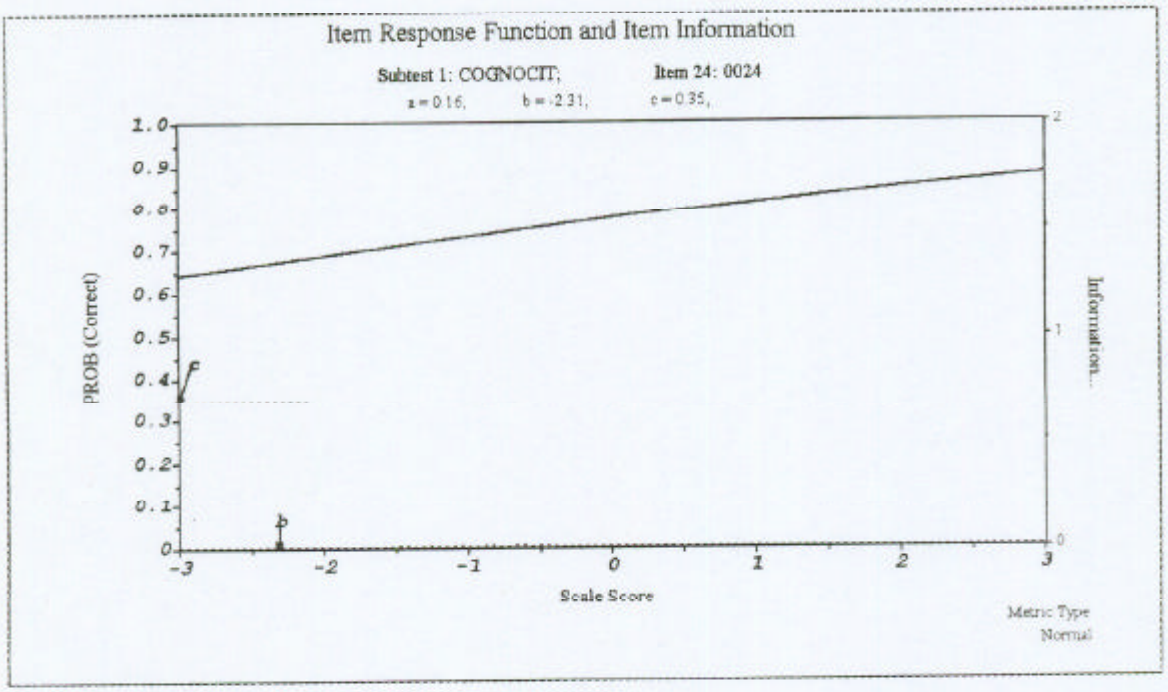

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 19: 0019
a = 0.46;
b = -1.40;
c = 0.29;
PROB (Correct)
Information
Scale Score
Metric Type
Normal

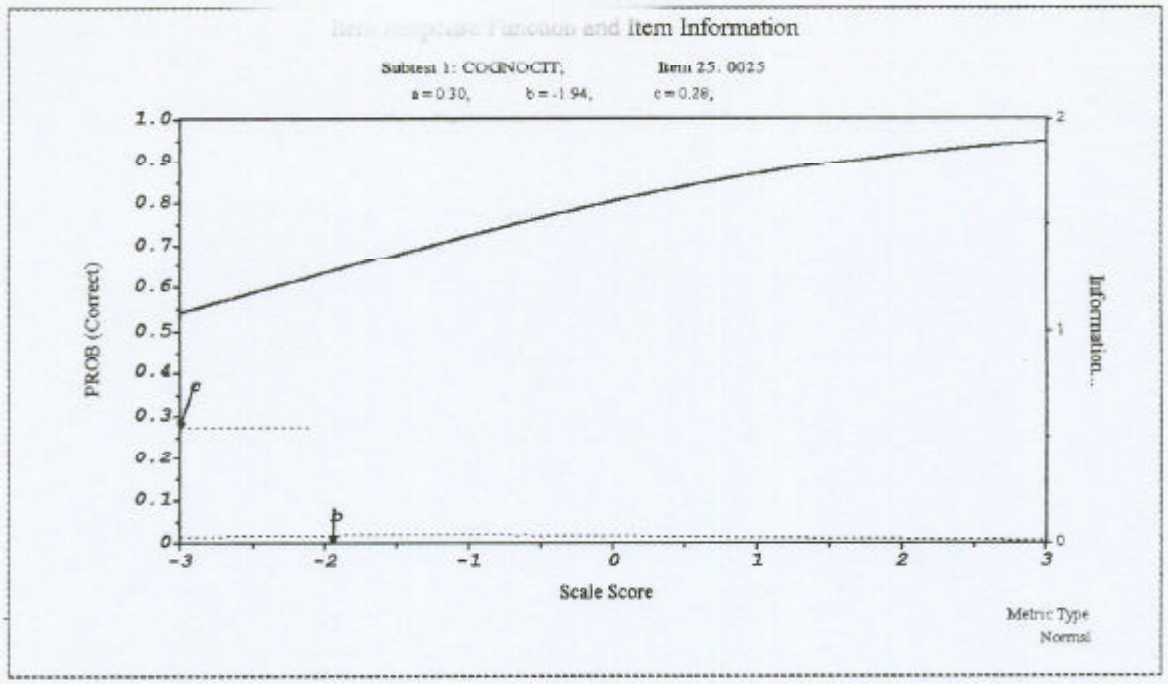

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 20: 0020
a = 0.40,
b = -0.57,
c = 0.28,
PROB (Correct)
Information
Scale Score
Metric Type
Normal

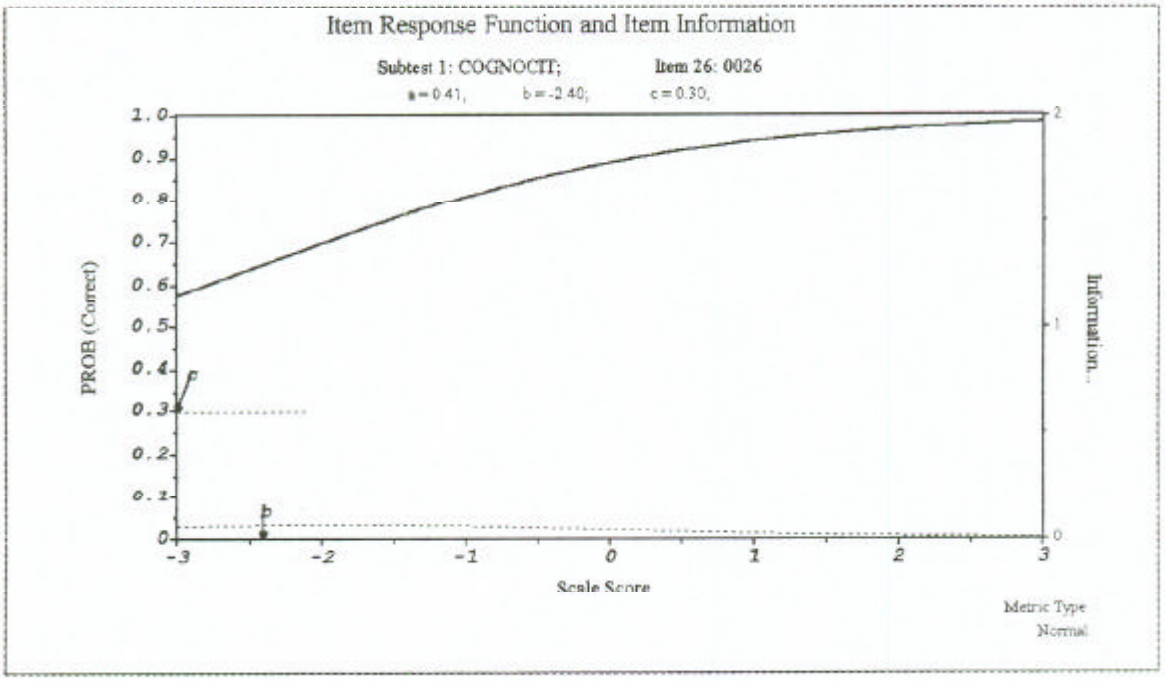

Item Response Function and Item Information
Subtest 1: COGNOCIT,
Item 21: 0021
a = 0.57,
b = -0.10,
c = 0.39,
PROB (Correct)
Information
Scale Score
Metric Type
Normal

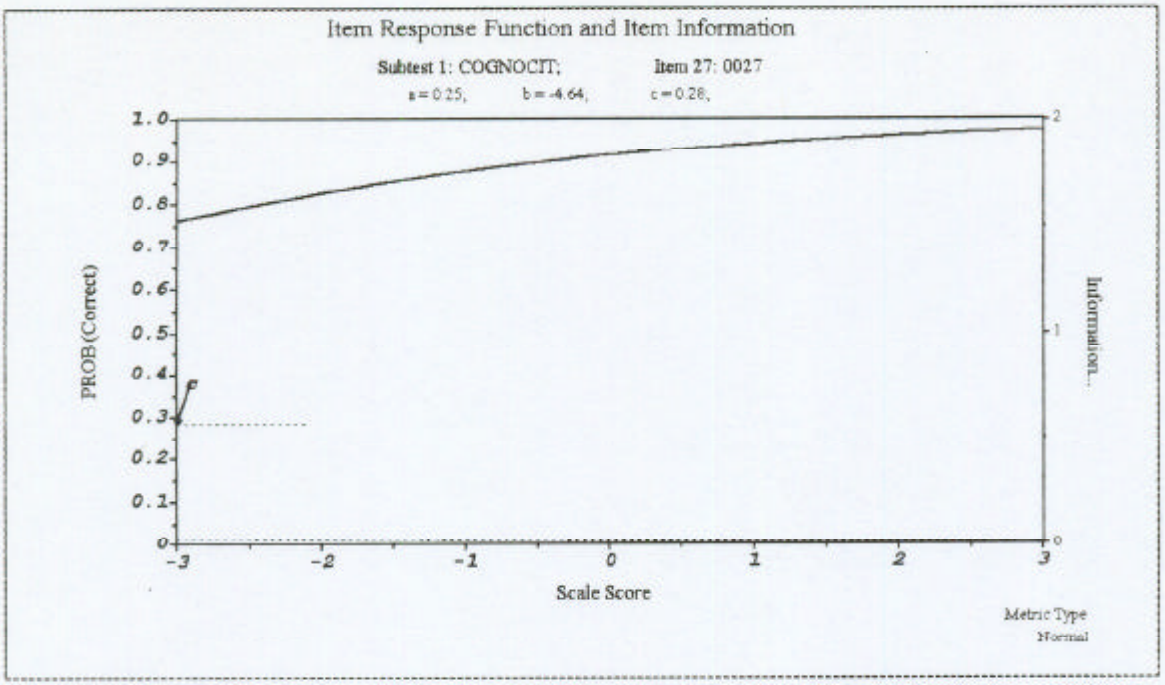

Item Response Function and Item Information
Subtest 1: COGNOCIT,
Item 22: 0022
a = 0.31,
b = -2.54,
c = 0.25,
PROB (Correct)
Information
Scale Score
Metric Type
Normal
b

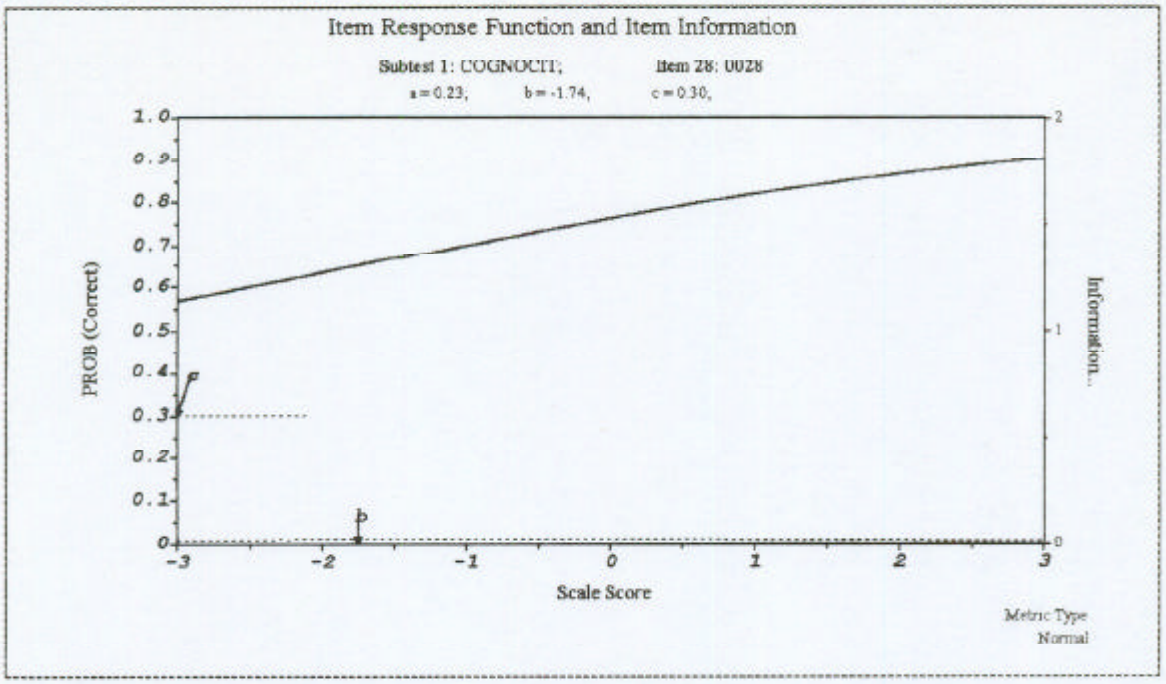

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 23: 0023
a = 0.38,
b = -1.09,
c = 0.23,
PROB (Correct)
Information
Scale Score
Metric Type
Normal

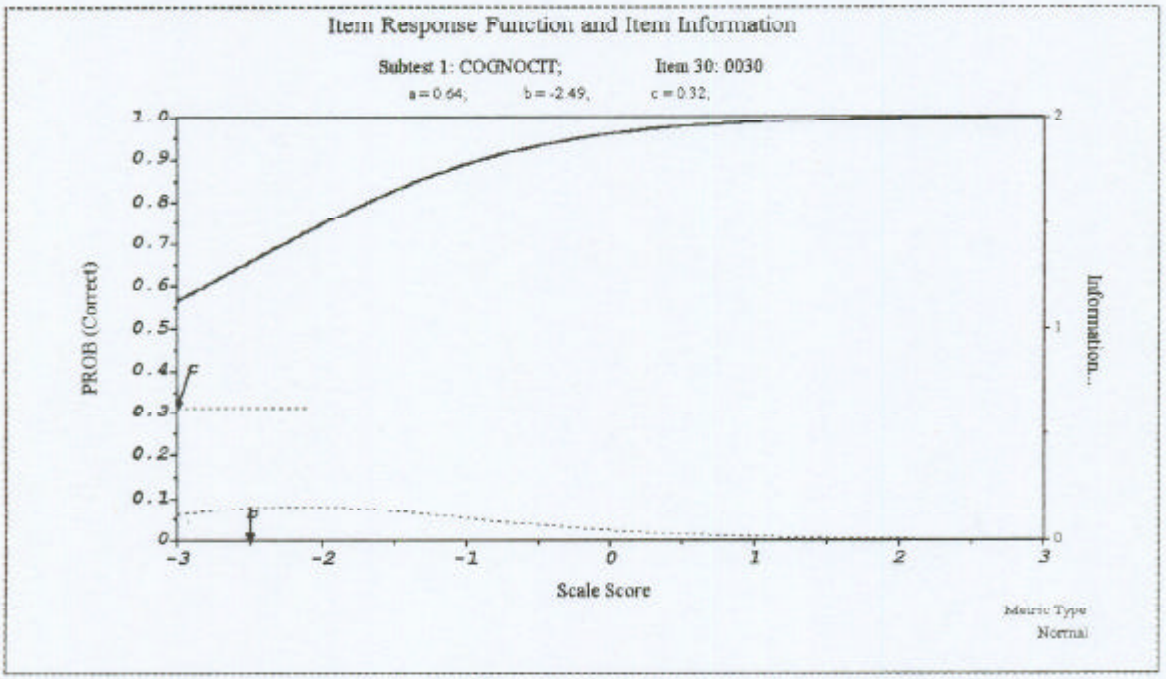

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 24: 0024
a = 0.16,
b = -2.31,
c = 0.35,
PROB (Correct)
Information...
Scale Score
Metric Type
Normal

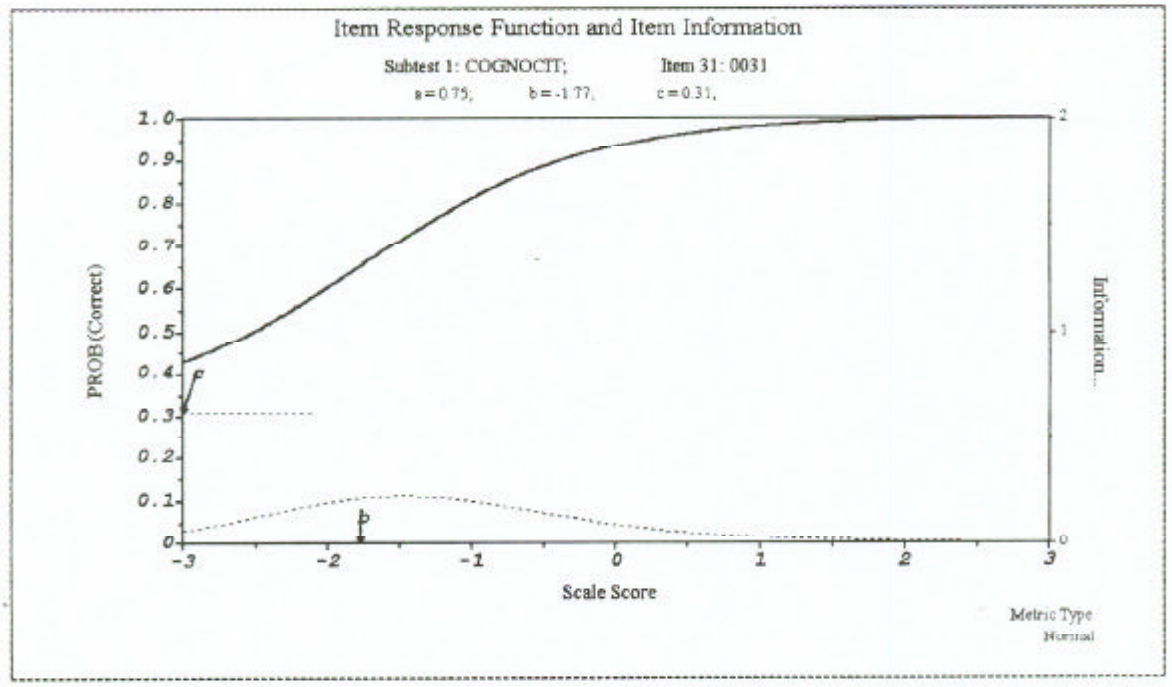

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 25. 0025
a = 0.30,
b = -1.94,
c = 0.28,
PROB (Correct)
Information...
Scale Score
Metric Type
Normal

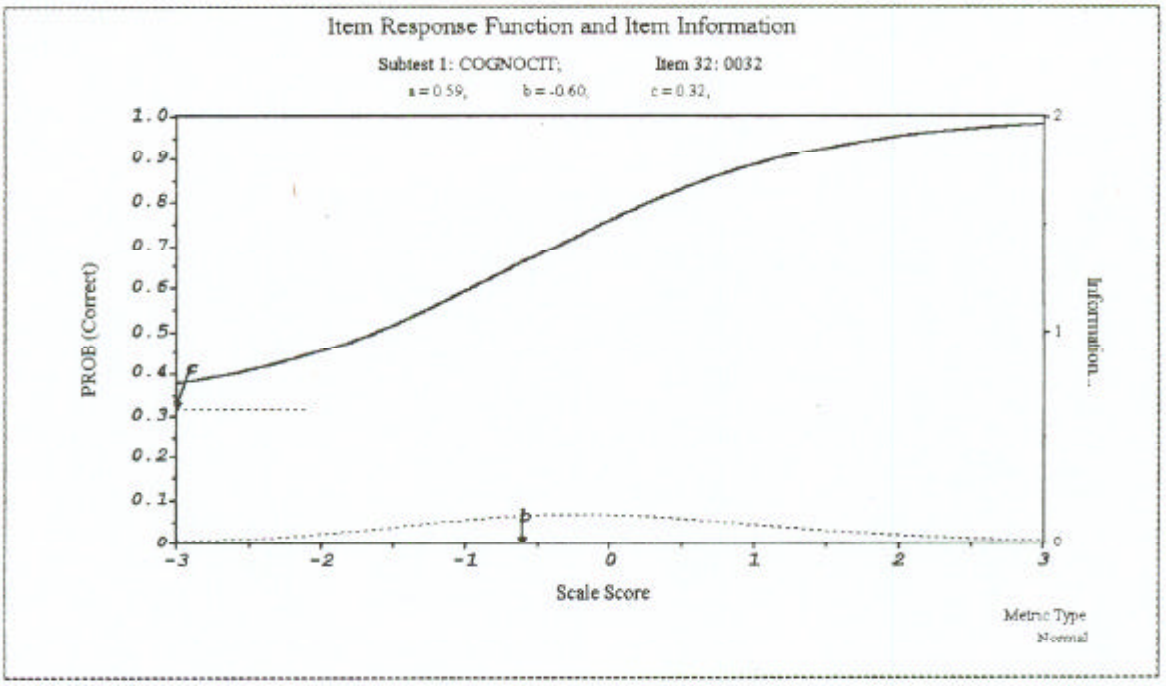

Item Response Function and Item Information
Subtest 1: COGNOCITI;
Item 26: 0026
a = 0.41;
b = -2.40;
c = 0.30;
PROB (Correct)
Information
Scale Score
Metric Type
Normal

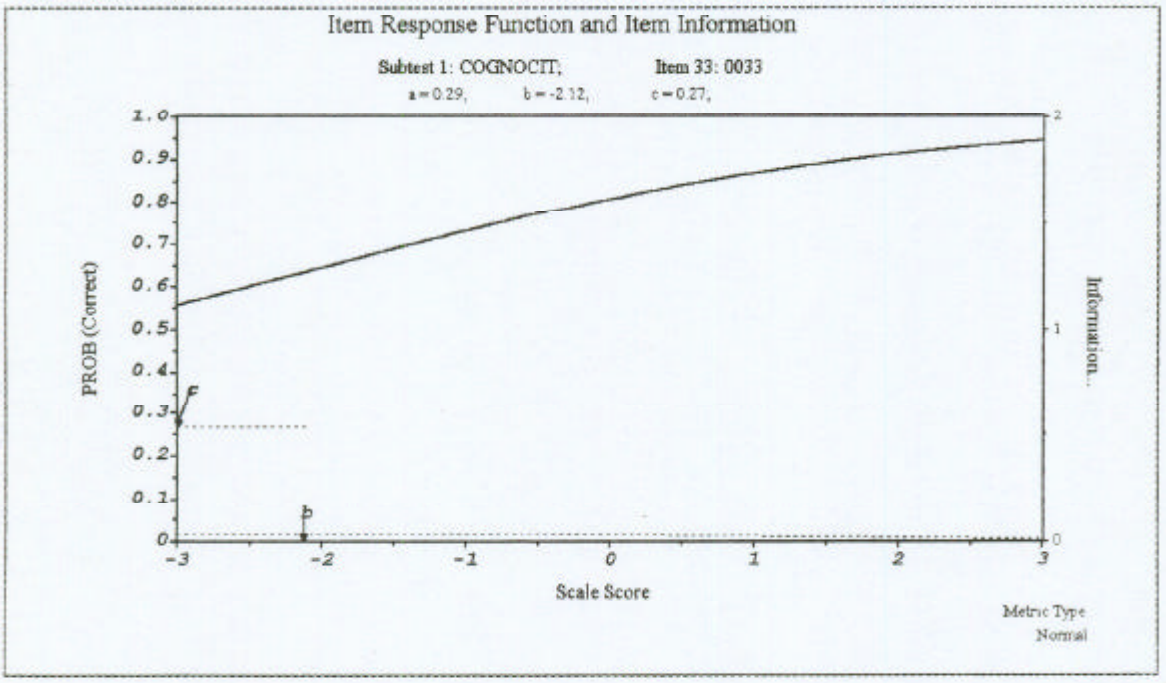

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 27: 0027
a = 0.25,
b = -4.64,
c = 0.28,
PROB (Correct)
Information
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2
1
0
-3
-2
-1
0
1
2
3
Scale Score
Metric Type
Normal

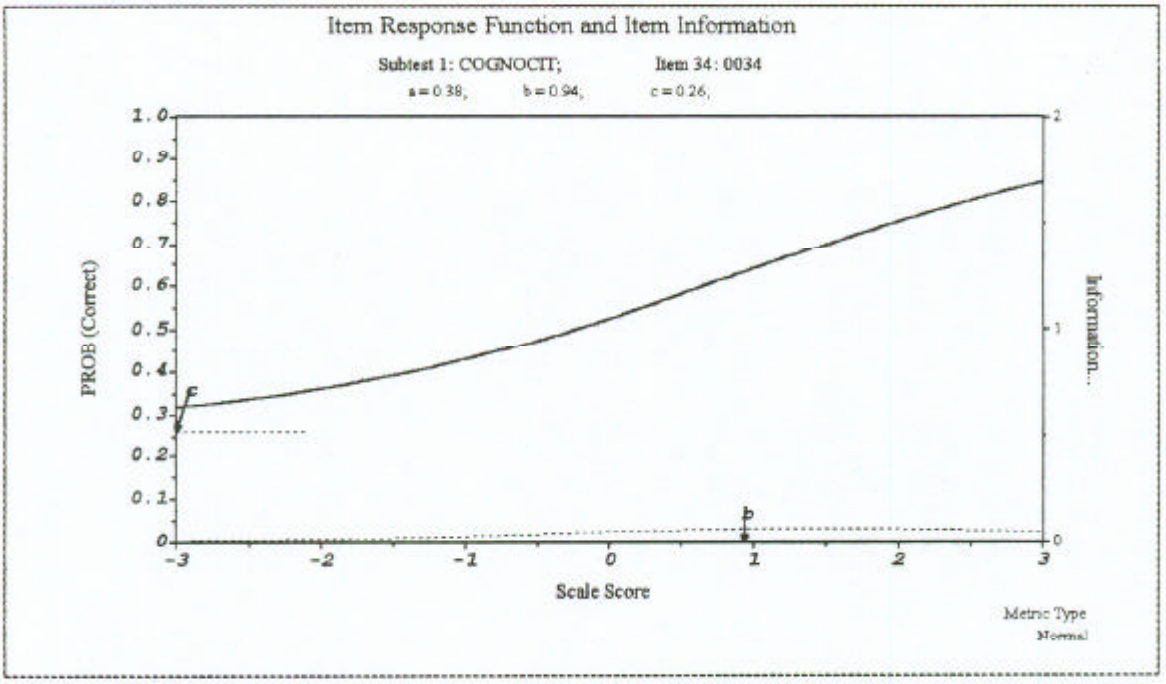

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 28: 0028
a = 0.23;
b = -1.74;
c = 0.30;
PROB (Correct)
Information
Scale Score
Metric Type
Normal

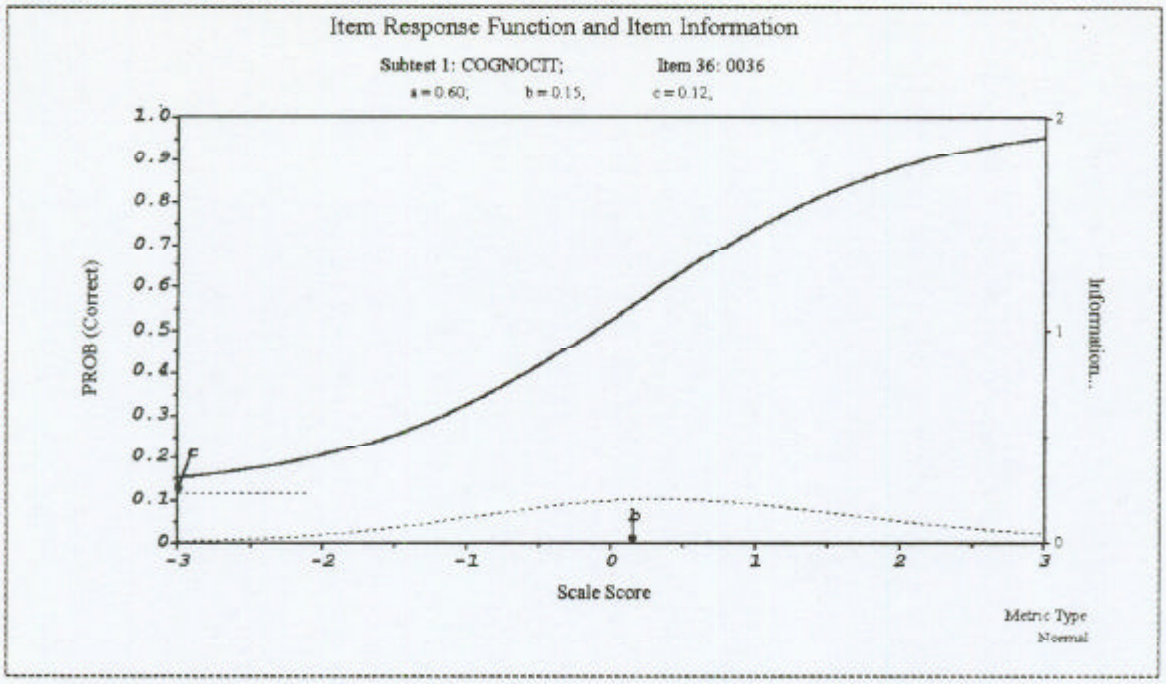

Item Response Function and Item Information
Subtest 1: COGNOCTT;
Item 30: 0030
a = 0.64;
b = -2.49;
c = 0.32;
PROB (Correct)
Information
Scale Score
Metric Type
Normal

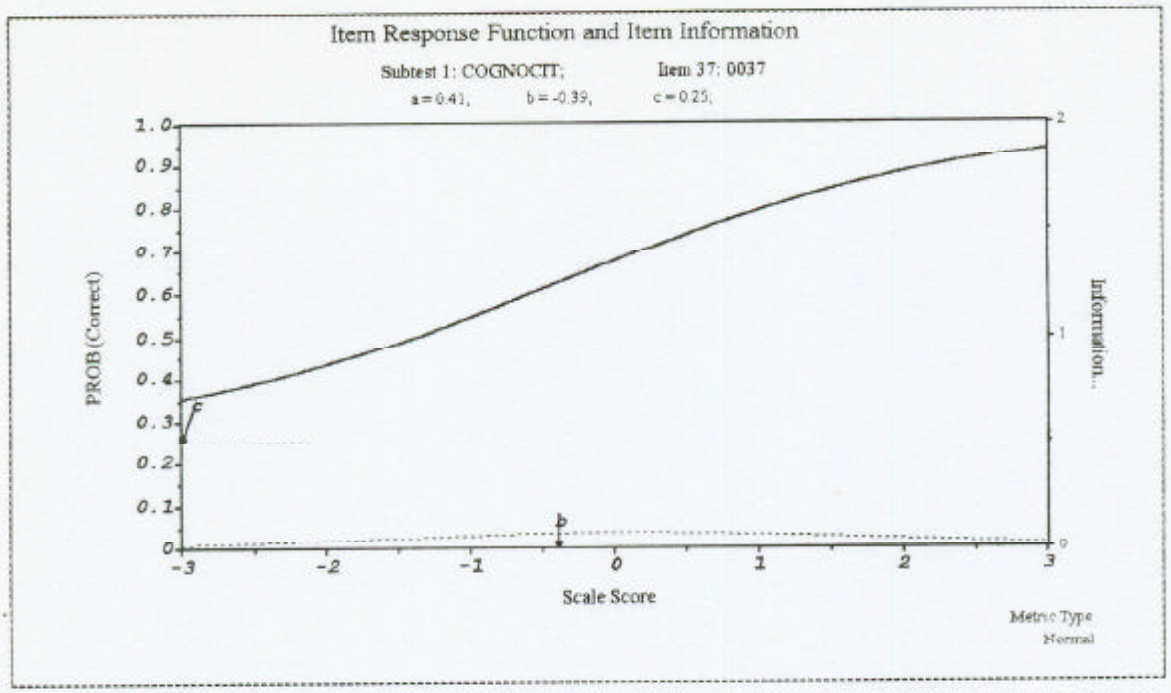

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 31: 0031
a=0.75;
b=-1.77;
c=0.31;
PROB (Correct)
Information
Scale Score
Metric Type
Normal

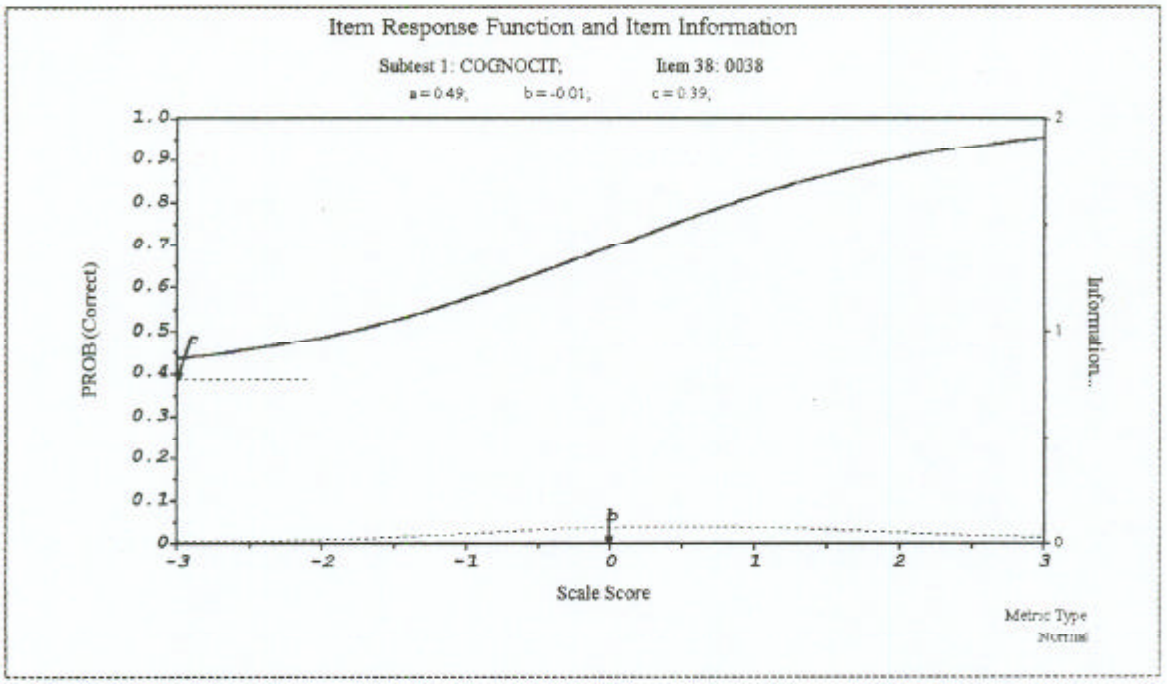

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 32: 0032
a = 0.59;
b = -0.60;
c = 0.32;
PROB (Correct)
Information
Scale Score
Metric Type
Normal

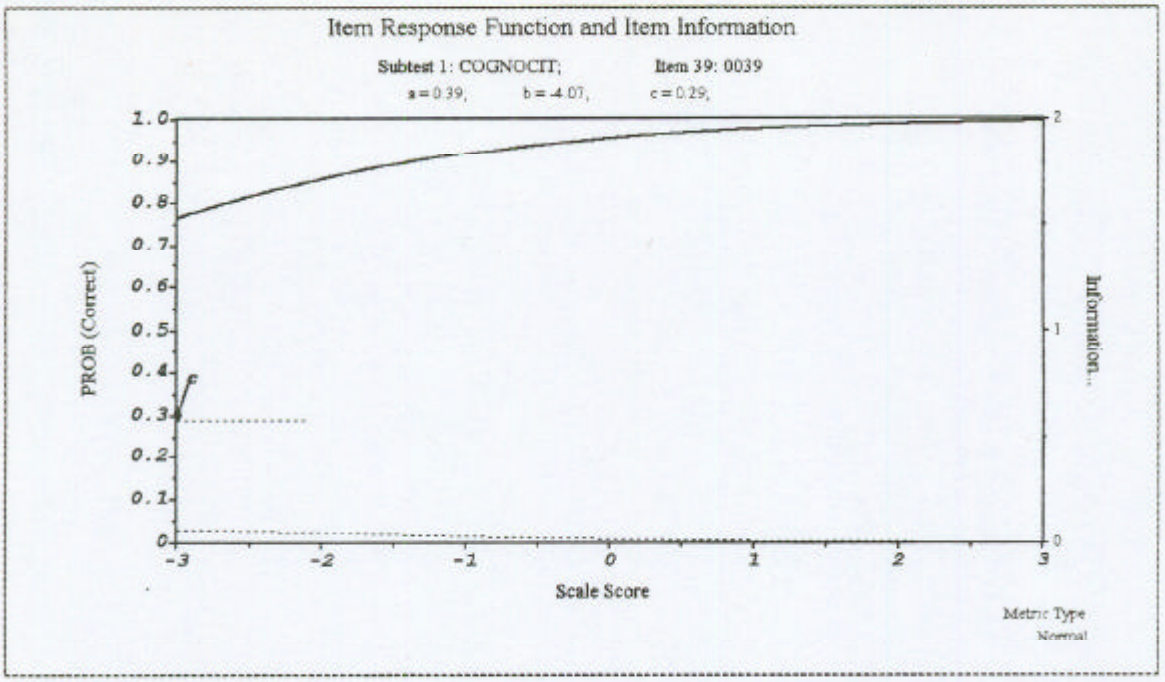

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 33: 0033
a = 0.29;
b = -2.12;
c = 0.27;
PROB (Correct)
Information
Scale Score
Metric Type
Normal
c
b
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-3
-2
-1
0
1
2
3
2
1
0

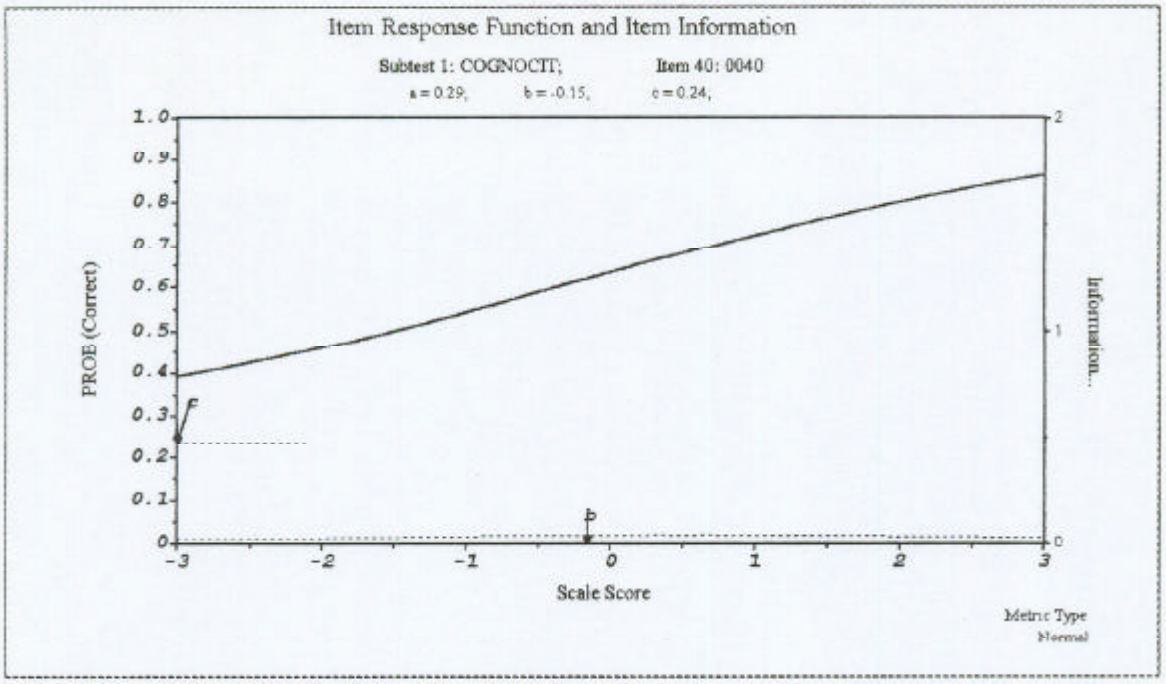

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 34 : 0034
a = 0.38,
b = 0.94,
c = 0.26,
PROB (Correct)
Information...
Scale Score
Metric Type
Normal

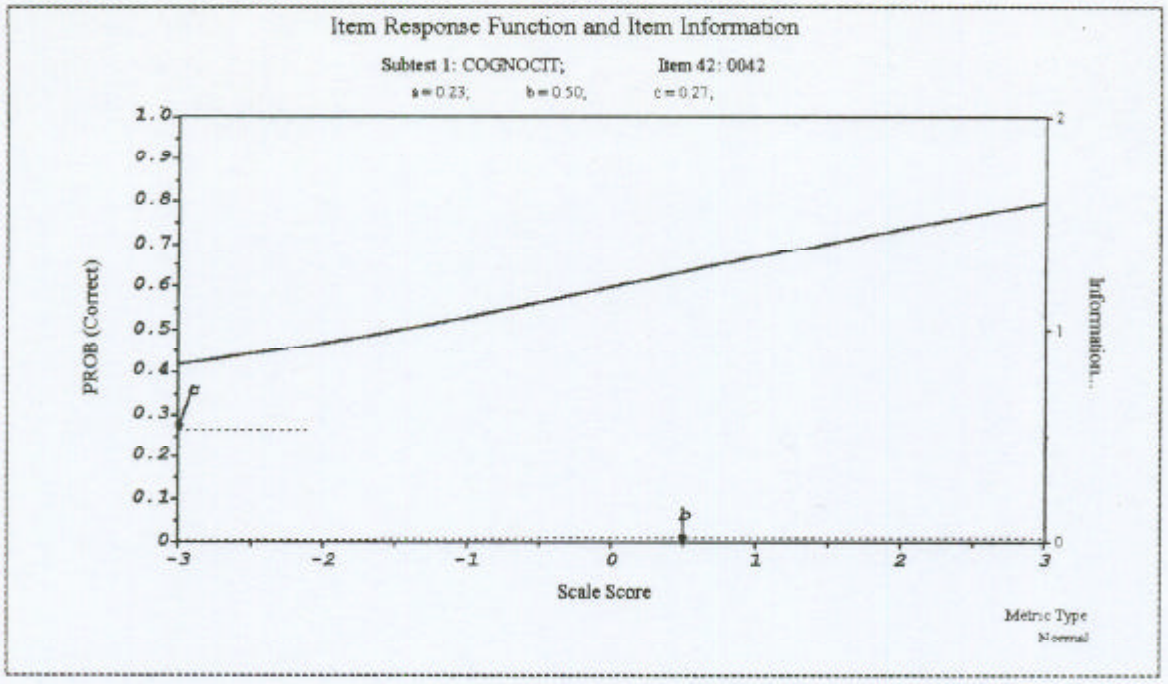

Item Response Function and Item Information
Subtest 1: COGNOCTT;
Item 36: 0036
a = 0.60;
b = 0.15;
c = 0.12;
PROB (Correct)
Information
Scale Score
Metric Type
Normal

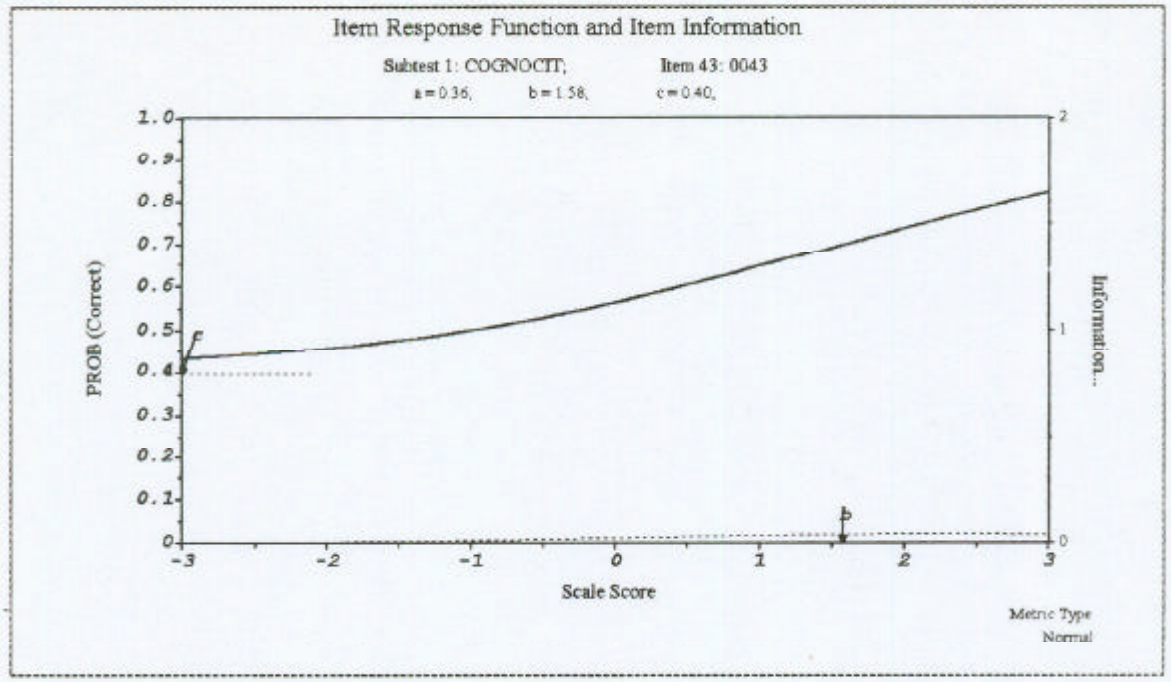

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 37: 0037
a = 0.41,
b = -0.39,
c = 0.25,
PROB (Correct)
Information
Scale Score
Metric Type
Normal

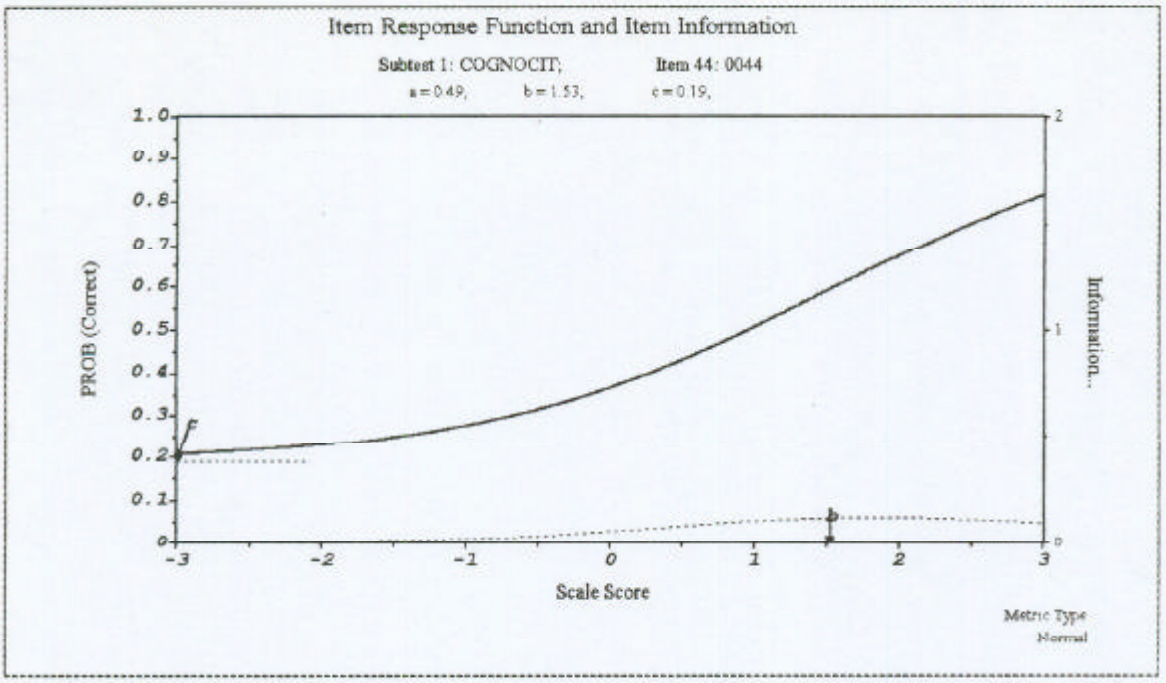

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 38: 0038
a = 0.49;
b = -0.01;
c = 0.39;
PROB (Correct)
Information
Scale Score
Metric Type

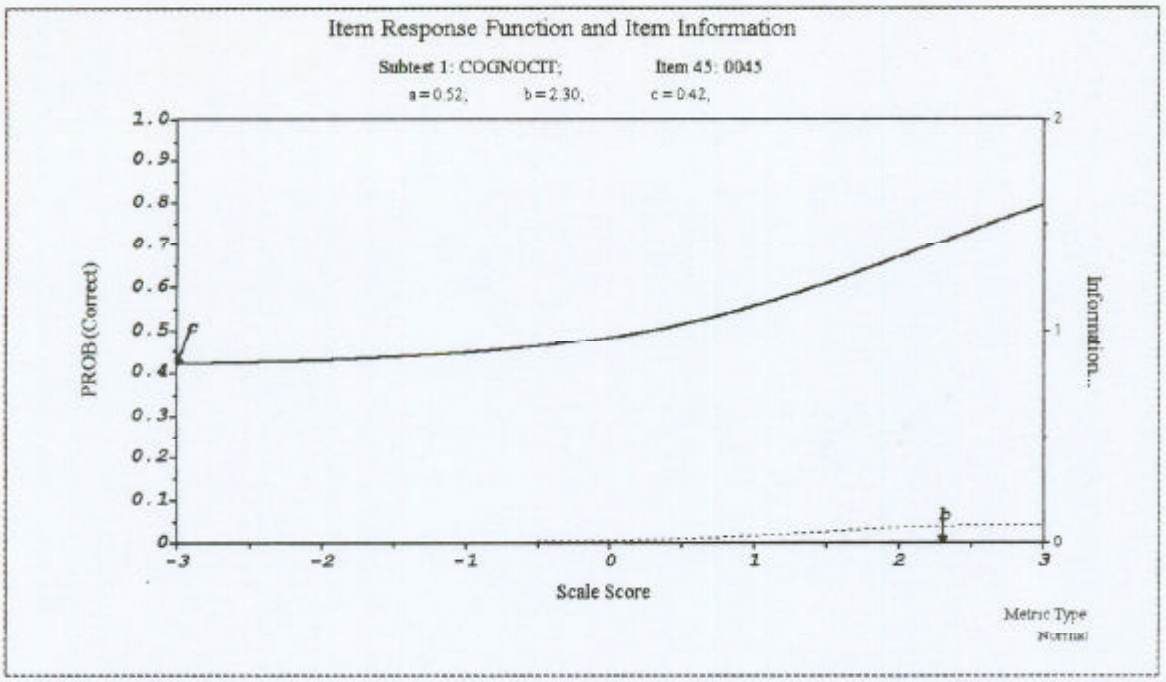

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 39: 0039
a = 0.39;
b = -4.07;
c = 0.29;
PROB (Correct)
Information
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2
1
0
-3
-2
-1
0
1
2
3
Scale Score
Metric Type
Normal

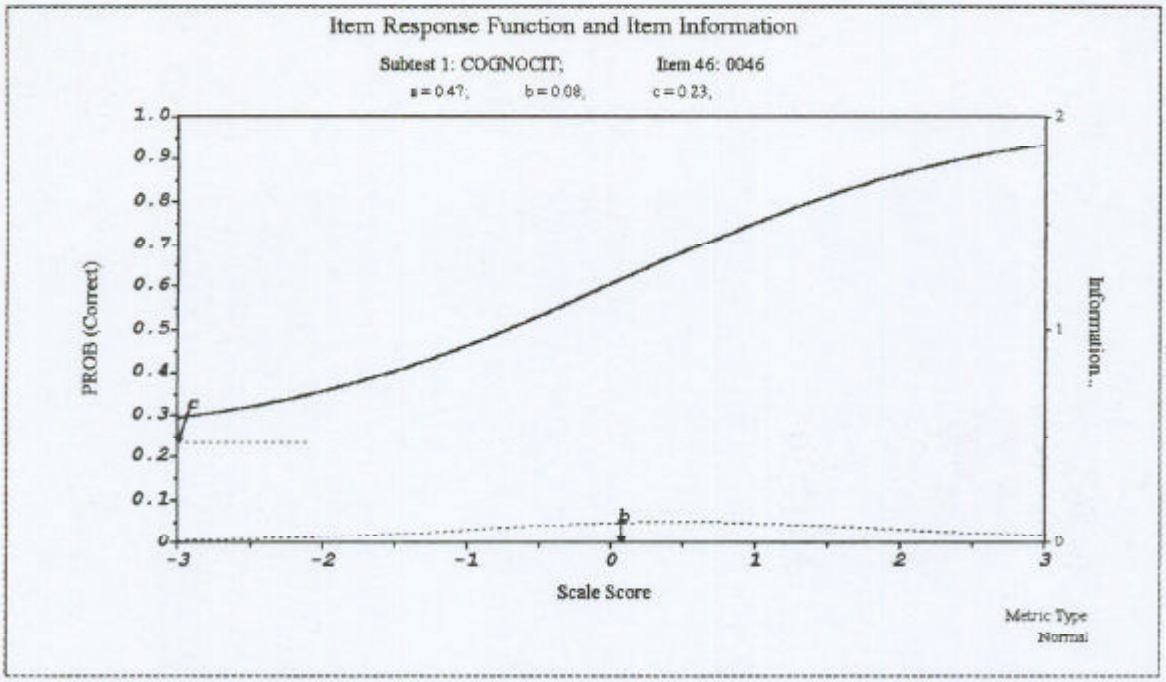

Item Response Function and Item Information
Subtest 1: COGNOCTI;
Item 40: 0040
a = 0.29;
b = -0.15;
c = 0.24;
PROB (Correct)
Information
Scale Score
Metric Type
Normal

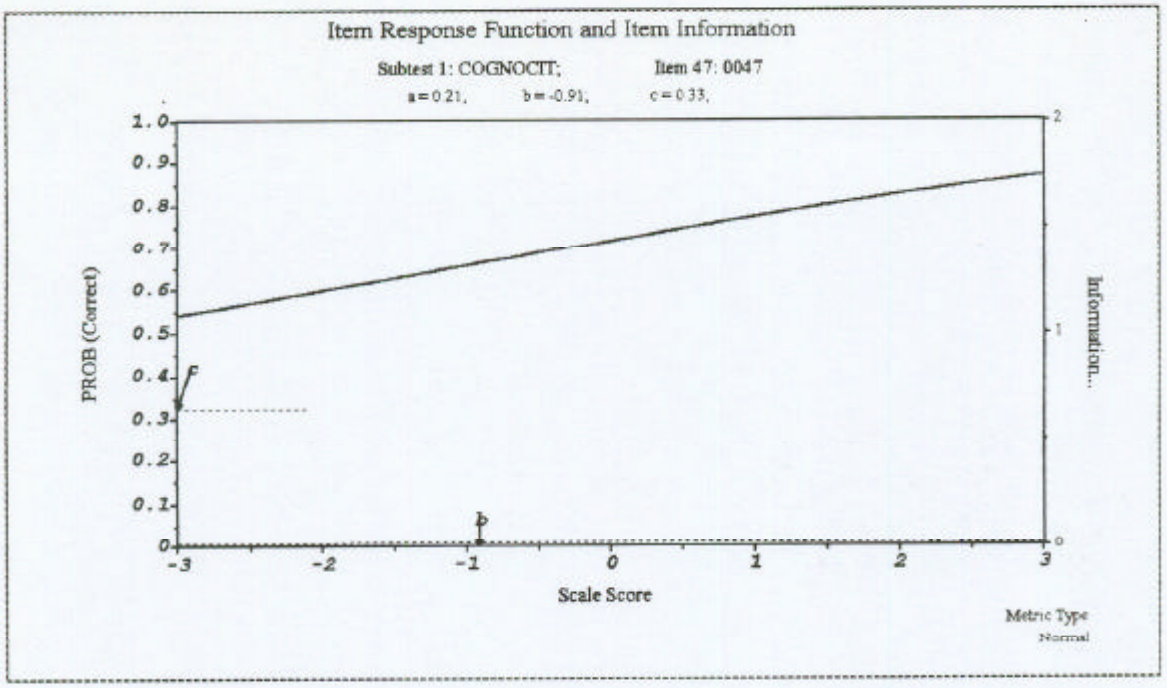

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 42: 0042
a = 0.23;
b = 0.50;
c = 0.27;
PROB (Correct)
Information
Scale Score
Metric Type
Normal

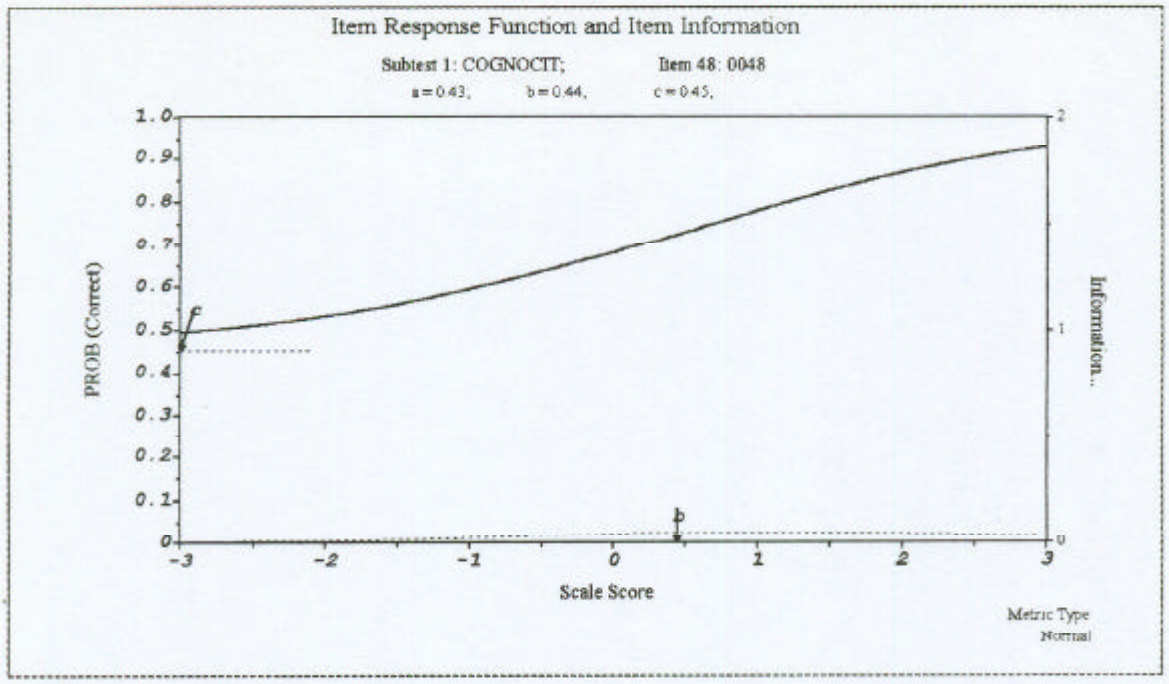

Item Response Function and Item Information
Subtest 1: COGNOCTT;
Item 43: 0043
a = 0.36;
b = 1.58;
c = 0.40;
PROB (Correct)
Information
Scale Score
Metric Type
Normal

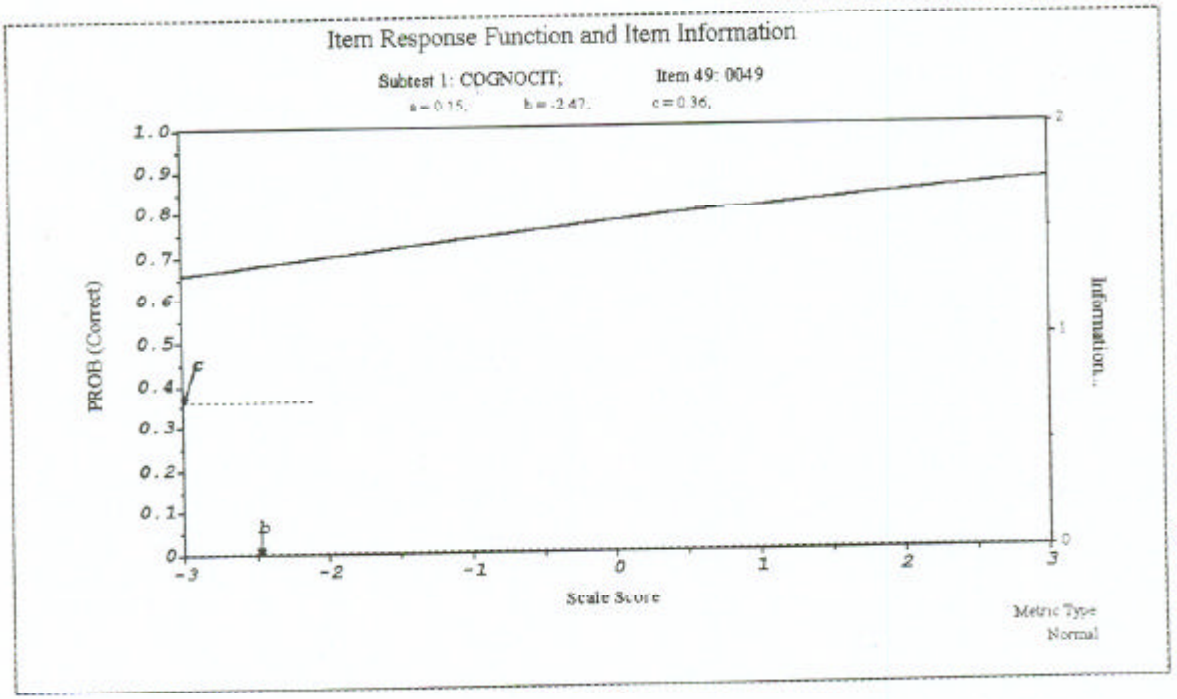

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 44: 0044
a = 0.49;
b = 1.53;
c = 0.19;
PROB (Correct)
Information...
Scale Score
Metric Type
Normal

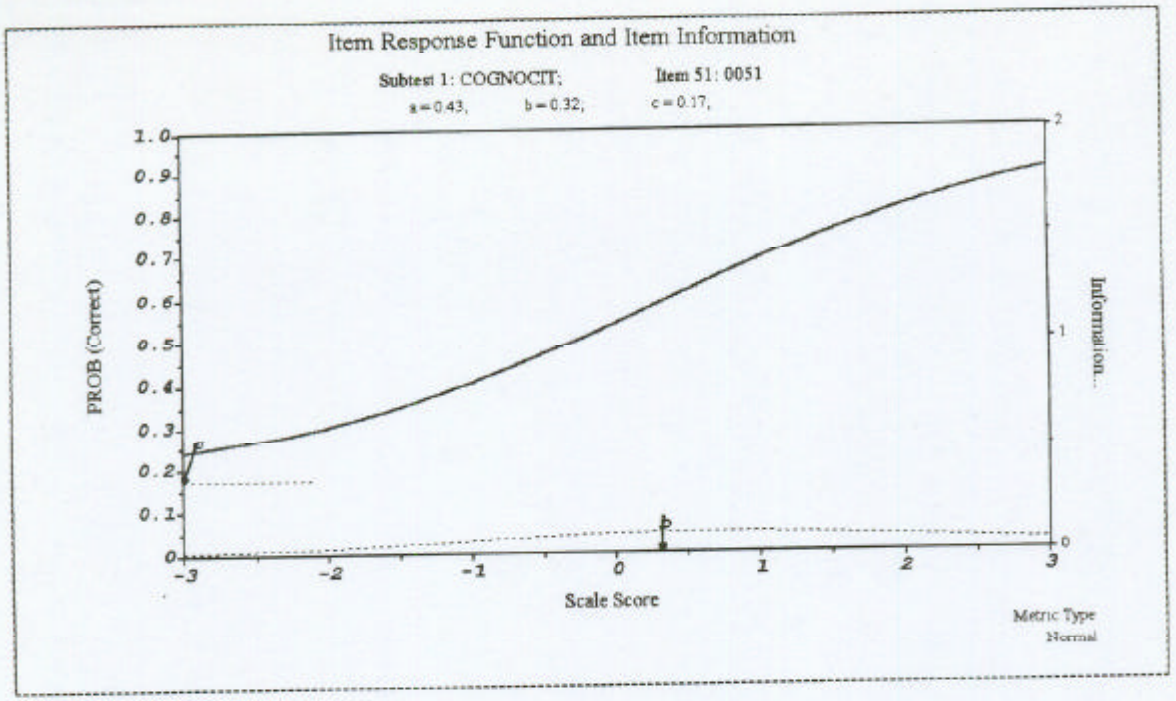

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 45: 0045
a=0.52,
b=2.30,
c=0.42,
PROB(Correct)
Information
Scale Score
Metric Type
Normal

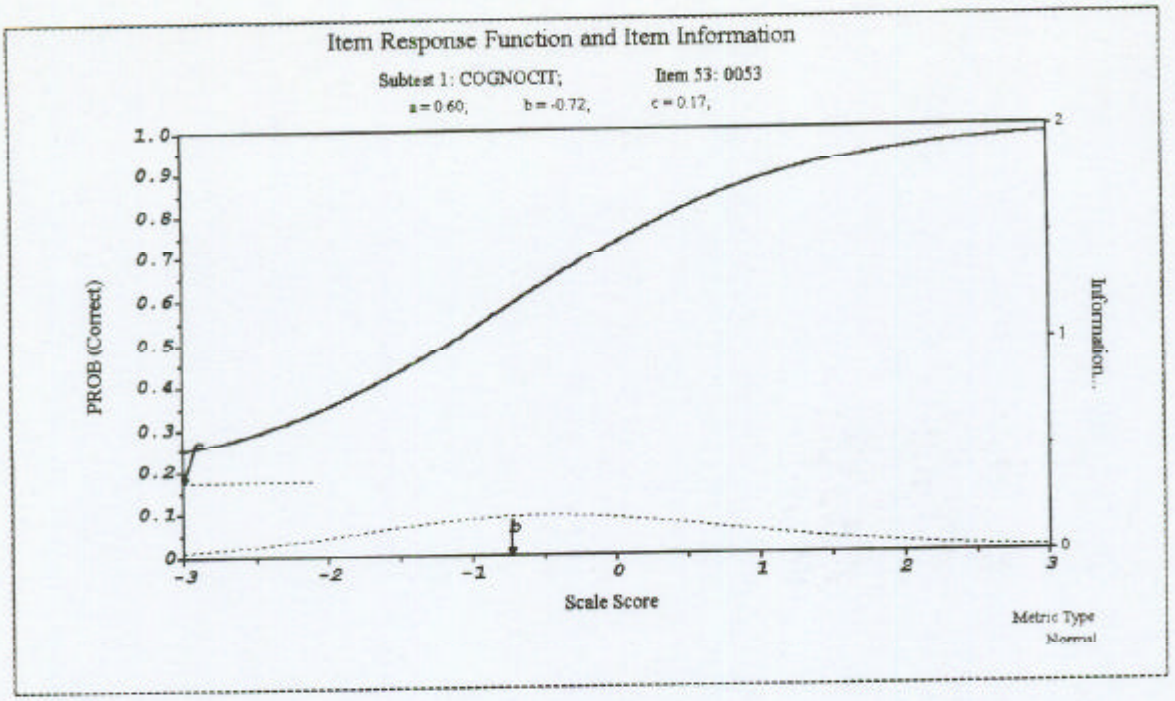

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 46: 0046
a=0.47;
b=0.08;
c=0.23;
PROB (Correct)
Information
Scale Score
Metric Type
Normal

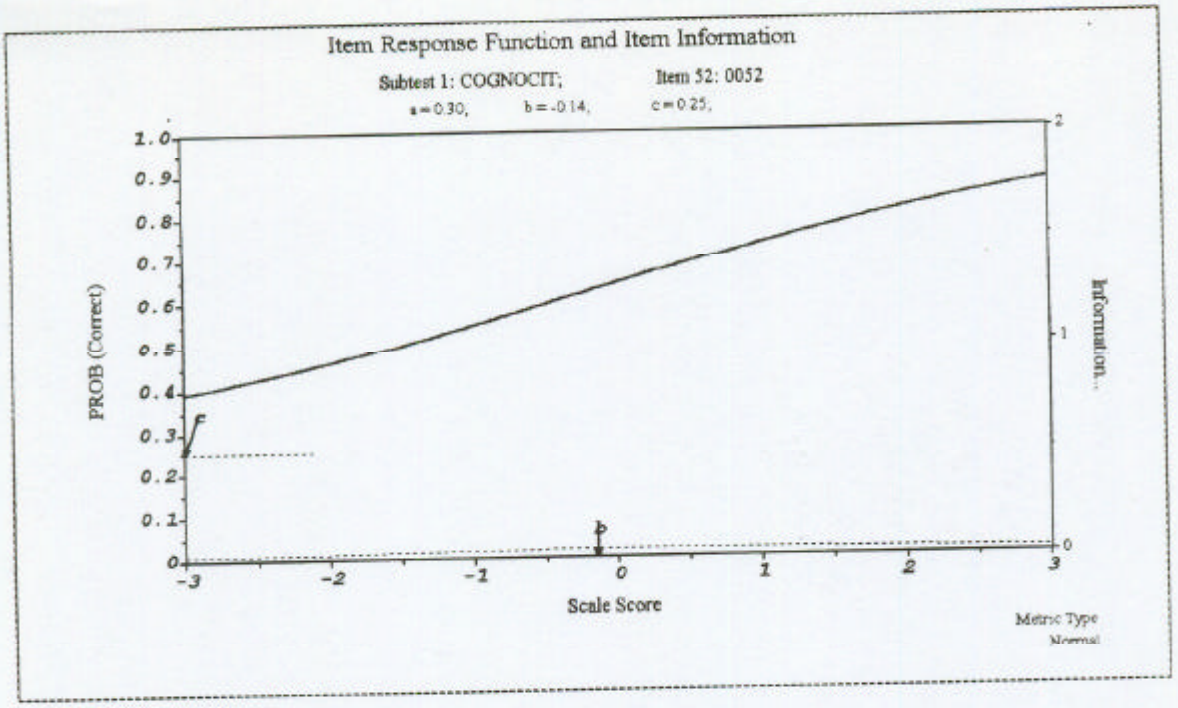

Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 47: 0047
a = 0.21,
b = -0.91,
c = 0.33,
PROB (Correct)
Information
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2
1
0
c
b
-3
-2
-1
0
1
2
3
Scale Score
Metric Type
Normal

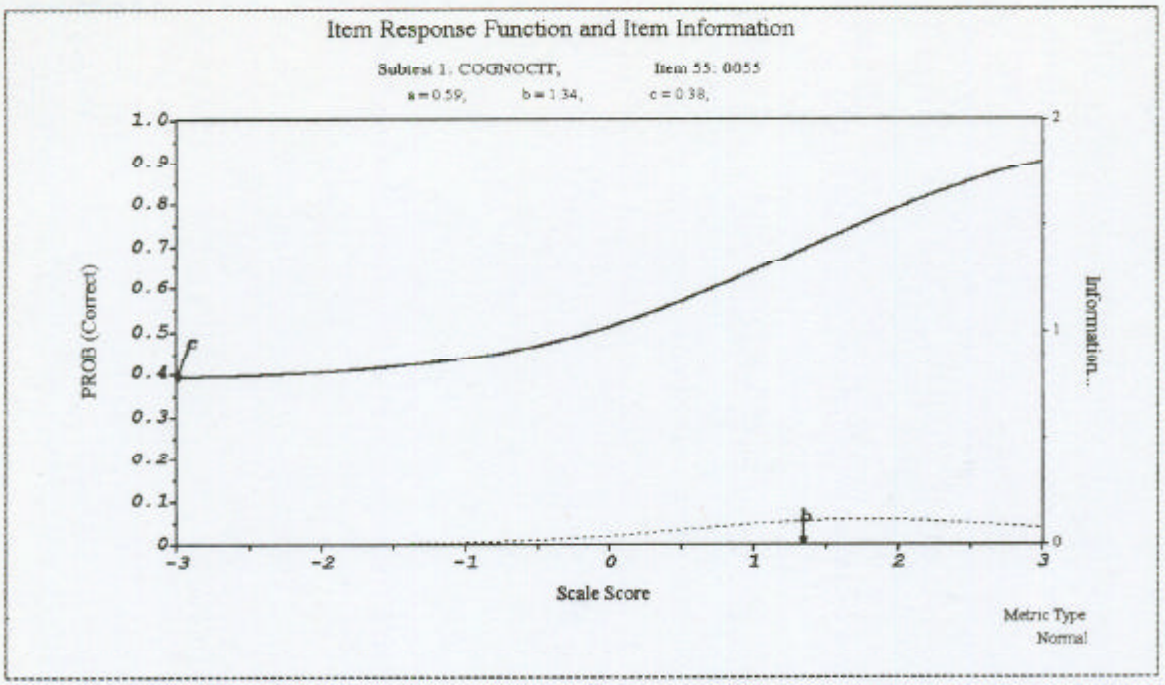

Item Response Function and Item Information
Subtest 1: COGNOCTT;
Item 48: 0048
a = 0.43,
b = 0.44,
c = 0.45,
PROB (Correct)
Information
Scale Score
Metric Type
Normal


Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 49: 0049
a = 0.15,
b = -2.47,
c = 0.36,
PROB (Correct)
Information
Scale Score
Metric Type
Normal


Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 51: 0051
a = 0.43;
b = 0.32;
c = 0.17;
PROB (Correct)
Information
Scale Score
Metric Type
Normal


Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 53: 0053
a = 0.60;
b = -0.72;
c = 0.17;
PROB (Correct)
Information
Scale Score
Metric Type
Normal


Item Response Function and Item Information
Subtest 1: COGNOCIT;
Item 52: 0052
a = 0.30,
b = -0.14,
c = 0.25,
PROB (Correct)
Information
Scale Score
Metric Type
Normal


Item Response Function and Item Information
Subtest 1: COGNOCTT,
Item 55: 0055
a = 0.59,
b = 1.34,
c = 0.38,
PROB (Correct)
Information
Scale Score
Metric Type
Normal