

識別性検査 A-1001の「関係判断力・応用力」領域 および「記憶」領域のIRT尺度化

野 口 裕 之

1. はじめに

識別性検査 A-1001は、鉄道事業の運転関係従事員を対象とする一種の知的能力検査で、

- ① 検査項目に鉄道に関係の深い素材を用いる
- ② 検査を実施するのにパーソナルコンピュータを用いる
- ③ 検査を支えるテスト理論に項目反応理論を用いる
- ④ 検査の実施は適応型テスト方式を用いる

という条件の下で、新たに開発するものである。そして、測定領域および問題は、

- I 「知覚の速さ・正確さ」領域
 - ① 「図形・記号の探索」問題
 - ② 「図形・記号の異同弁別」問題
- II 「関係判断力・応用力」領域
 - ① 「空間図形の推理」問題
 - ② 「図形・記号系列の作成」問題
- III 「記憶」領域
 - ① 「図形・記号系列の順唱・逆唱」問題

で構成されている。

これまでに、項目プールの作成、予備テスト、項目分析が実施され、既に野口(1993a, 1993b, 1994a, 1994b)で報告されているが、さらに、「記憶」領域についてはIRT尺度化が試みられ(野口, 1995a), 「知覚の速さ・正確さ」領域に属する「図形・記号の探索」問題および「図形・記号の異同弁別」問題についても項目反応時間情報を組み込んでIRT尺度化が試みられた(野口, 1995b)。

本研究では、「関係判断力・応用力」領域の「空間図形の推理」問題および「図形・記号系列の作成」問題、そして「記憶」領域の「図形・記号系列の順唱・逆唱」問題についてIRT尺度を構成する。「記憶」領域につ

いては既に野口(1995a)の報告があるが、今回は新たな受験者データを加えた上で再びIRT尺度化を実施する。

2. 項目反応モデル

項目反応モデルと総称されるモデル群のなかで、野口(1995b)では「知覚の速さ・正確さ」領域に属する「図形・記号の探索」問題および「図形・記号の異同弁別」問題をIRT尺度化するのに、これらの問題が速度検査になっているため、受験者の各項目に対する解答時間と解答の正誤とを組み合わせで順序カテゴリー化した上で『段階反応モデル』を用いてIRT尺度を構成した。これに対して、本研究で対象とする「関係判断力・応用力」領域の「空間図形の推理」問題および「図形・記号系列の作成」問題、そして「記憶」領域の「図形・記号系列の順唱・逆唱」問題は、基本的には力量検査であるため解答速度を特に問題にする必要はなく、『2パラメタ・ロジスティック・モデル』を用いてIRT尺度化する。

すなわち、項目特性曲線にロジスティック曲線と呼ばれる曲線形を仮定し、項目jの特性関数は、潜在特性尺度値を θ 、正答確率を $P_j(\theta)$ とすると、

$$P_j(\theta) = \{1 + \exp[-1.7a_j(\theta - b_j)]\}^{-1} \quad (1)$$

で与えられる。ここで、 a_j および b_j は曲線の形状を決定するパラメタで、 a_j は曲線の立ち上がりの強さを、 b_j は横軸上における曲線の位置を表わし、これらのパラメタで項目の特性は全て記述される。前者は『識別力パラメタ』、後者は『困難度パラメタ』と呼ばれる。

テストに含まれる各項目のパラメタ値は受験者集団の項目反応データをもとに推定し、さらに、パラメタ推定値を利用してテスト情報量を計算することによって各テストの測定精度が評価できる。

詳しくは、例えば Lord, F. M. & Novick, M. R. (1968), Hambleton, R. K. & Swaminathan, H. (1985), 芝 (1991), 渡辺・野口(印刷中)などを参照

されたい。

3. 空間図形の推理問題

3.1. 問題の概要

本問題では、鉄道線路に用いられる枕木に似た形状の積み木を並べたり積み上げたものを呈示し、受験者はその中にひとつだけ存在する緑色の積み木に接している積み木の数解答する。本検査では“鉄道に関係の深い素材”を用いて問題項目を作成するため、通常の心理テストによく用いられる直方体の積み木ではなくてPC枕木のイメージを持つ積み木を用いている。

各項目毎に、積み木の総数および接する積み木の数変化する。このことによって、項目困難度を変化させて

いる。本問題は“反応の速さ”ではなく、受験者の“正誤反応”を測定することを目的としている。全部で30項目から構成されるが、各項目に含まれる積み木の総数および接する積み木数は表1に示したとおりである。

各受験者毎に、解答・正誤および反応時間(0.01秒単位)が記録される。

詳しくは、野口(1993b)を参照されたい。

3.2. 項目分析の結果

本研究に用いる受験者は、1993年6月から1996年6月に本検査を実施した名古屋市内および岡山市内の大学生290名と一般社会人88名の合わせて378名である。本問題の各項目は、野口(1993b)で既に項目分析が実施され

表1 「空間図形の推理」問題各項目の積み木の総数・接する積み木の数、通過率・点双列相関係数、反応時間(秒)の平均・標準偏差・総反応時間との相関係数および項目パラメタ推定値

項目番号	積み木数	接する数	通過率	点双列相関係数	反応時間平均	反応時間標準偏差	総反応時間との相関係数	識別力	困難度
1	6	3	.677	.339	6.931	3.755	.408	.463	-1.080
2	6	4	.638	.293	6.132	4.245	.334	.375	-.983
3	9	6	.347	.443	9.525	5.970	.437	.680	.674
4	9	5	.865	.214	10.637	5.997	.479	.399	-2.996
5	9	6	.931	.371	9.440	5.986	.530	.883	-2.334
6	9	4	.690	.553	9.429	6.326	.563	.870	-.756
7	10	6	.743	.266	15.273	9.620	.583	.352	-1.930
8	10	5	.553	.552	15.481	10.080	.572	.833	-.208
9	8	4	.585	.383	15.697	10.413	.688	.487	-.487
10	8	2	.447	.086	9.592	7.143	.527	.180	.685
11	10	6	.865	.409	10.957	6.456	.501	.729	-1.907
12	10	5	.619	.564	14.637	9.430	.503	.898	-.450
13	10	5	.780	.515	10.966	6.963	.504	.849	-1.204
14	18	8	.622	.587	17.756	9.709	.604	1.085	-.415
15	18	7	.640	.625	15.431	8.199	.594	1.247	-.452
16	18	11	.690	.630	21.156	11.174	.602	1.237	-.634
17	20	10	.643	.567	18.563	11.309	.635	.941	-.530
18	20	11	.553	.517	19.634	11.042	.638	.721	-.229
19	20	6	.516	.578	21.457	13.542	.690	.947	-.057
20	20	7	.762	.628	14.605	9.506	.600	1.450	-.866
21	27	10	.402	.406	23.329	13.093	.682	.584	.465
22	27	11	.561	.419	20.431	13.008	.615	.543	-.324
23	27	9	.497	.593	21.329	11.524	.679	1.070	.013
24	27	7	.378	.466	20.702	13.649	.552	.738	.502
25	27	6	.624	.500	15.864	8.171	.676	.701	-.550
26	27	6	.574	.358	24.792	15.571	.502	.424	-.477
27	27	6	.272	.445	19.332	14.092	.577	.766	.964
28	27	8	.622	.597	17.441	8.685	.657	1.026	-.427
29	27	7	.675	.357	14.879	8.735	.540	.467	-1.056
30	27	6	.632	.601	16.962	8.799	.576	.965	-.482

ているが、その後182名の受験者のデータが得られたため、今回改めて項目分析を実施する。

①項目通過率

各項目の通過率は表1に示した通りである。項目27の.272が最も小さく、項目3の.347が次に小さい。また、項目5の.931が最も大きく、項目4および11の.865が次いで大きい。易しい項目から難しい項目まで広い範囲の困難度の項目が含まれている。

②点双列相関係数

各項目の点双列相関係数は表1に示した通りである。項目16の.630が最も大きく、項目20の.628が次いで大きい。また、項目10の.086が最も小さく、項目4の.214が次に小さい。項目10の値については、前回の分析(野口, 1993b)でも.139と低い値を示していたが、今回ほどには極端に低くはなかった。項目の内容など詳細に検討する必要がある。その他の項目については特に問題となるものはなかった。

③項目間相関行列の因子分析

本問題中の項目の等質性を確認するために、項目間相関行列に対して因子分析を実施した。主因子解の計算が収束した時点で得られた固有値を大きいものから順に5つ示すと、9.97, 1.85, 1.31, 1.21, 1.09となり、第I固有値と第II固有値との比が5.4と大きく、しかも、第II固有値以下は値そのものが小さく高い1因子性を示している。したがって本問題に含まれる項目はかなり高い等質性を持つと言える。

④正答数得点の分布

正答数得点の分布は表2に示した通りである。最高が全項目に正答した30点が1名、次いで29点が2名、最低は1点が2名、次いで2点が1名と広い範囲に分布している。正答数得点の平均は18.4, 標準偏差は6.4, 歪度は-0.53, 尖度は-0.56であった。全体としてわずかに左に裾を引く分布を示している。

⑤総反応時間の分布

各受験者毎に全ての項目に要した反応時間の合計した総反応時間の分布について、平均, 標準偏差, 歪度, 尖度を計算すると、順に、469.4秒, 165.1秒, 0.99, 1.01である。やや右に裾を引く形状を示している。総反応時間の短い受験者及び長い受験者各5名分の実測値は、前者が最短の方から、146.1, 173.3, 181.1, 202.6, 209.3秒であり、後者が最長の方から、1111.6, 1043.4, 1034.5, 1009.2, 999.2秒であった。3分程度で本問題を終了した受験者がいる一方で、20分近くかけて解答した受験者もあり、総反応時間には大きな個人差が見られた。

⑥項目反応時間

各項目毎の受験者の反応時間の平均・標準偏差及び総反応時間との相関係数は表1に示した通りである。

反応時間の平均は、項目2で6.13秒と最も短く、項目1の6.93秒が次いで短く、項目26の24.79秒が最も長く、次いで項目21の23.33秒が長い。反応時間の標準偏差は、項目1で3.76秒と最も小さく、項目2の4.25秒が次いで小さく、項目26の15.57秒が最も大きく、項目27の14.09秒が次いで大きくなっている。また、

表2 「空間図形の推理」問題の正答数得点の分布

正答数得点	人数	百分率	正答数得点	人数	百分率
0	0	0.0	16	14	3.7
1	2	0.5	17	16	4.2
2	1	0.3	18	18	4.8
3	3	0.8	19	26	6.9
4	2	0.5	20	20	5.3
5	6	1.6	21	20	5.3
6	3	0.8	22	20	5.3
7	7	1.9	23	25	6.6
8	10	2.6	24	22	5.8
9	12	3.2	25	31	8.2
10	8	2.1	26	20	5.3
11	11	2.9	27	16	4.2
12	13	3.4	28	8	2.1
13	14	3.7	29	2	0.5
14	9	2.4	30	1	0.3
15	18	4.8			

総反応時間との相関係数は、項目19で0.69と最も大きく、項目9の0.69が次いで大きく、項目2の0.33が最も小さく、項目1の0.41が次いで小さい値を示した。

さらに、全項目間の反応時間間相関行列に対して因子分析を実施した。主因子解の計算が収束した時点で得られた固有値を大きいものから順に5つ示すと、9.41, 1.26, 0.72, 0.65, 0.53となり、第I固有値と第II固有値の比が7.5と大きく、さらに、第II固有値以降でも値に段差が見られず、高い1因子性を示した。

3.3. 項目反応モデルによる分析

3.3.1. 方法

本研究では、受験者の各項目に対する反応（正答または誤答）を基に、2パラメタ・ロジスティック・モデルを適用して各項目の識別力および困難度パラメタの推定値を計算する。計算プログラムにはパーソナル・コンピュータ用のBILOG3を利用する。尺度は、受験者母集団の特性尺度値の分布の平均が0.0、標準偏差が1.0となるように原点と単位とを設定する。

さらに、得られた項目パラメタ推定値にもとづいて、テスト情報量を計算する。

3.3.2. 結果および評価

項目パラメタ推定値は表1に示したとおりである。また、これらの30項目から得られるテスト情報量は表3および図1に示したとおりである。

表3 「空間図形の推理」問題のテスト情報量

尺度値	-3.0	-2.5	-2.0	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5	2.0	2.5	3.0
情報量	1.45	1.80	2.23	2.78	3.33	3.55	3.30	2.80	2.26	1.79	1.40	1.09	0.84
標準誤差	0.69	0.56	0.45	0.36	0.30	0.28	0.30	0.36	0.44	0.56	0.71	0.92	1.19

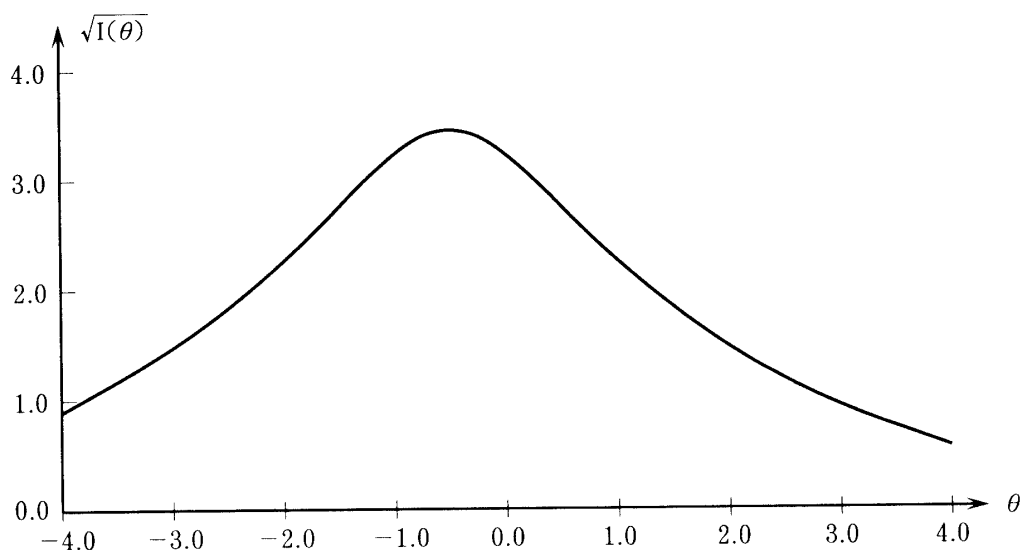


図1 「空間図形の推理問題」のテスト情報量

識別力パラメタの推定値は、最小値が項目10の0.18、最大値が項目20の1.45、平均が0.764、標準偏差が0.300であり、30項目中6項目が1.00を超えた。0.50未満の項目が8項目あるが、項目10を除くと極端に低い識別力を示す項目はなく、全体として十分な識別力を示している。項目10については点双列相関係数の値も極端に低く、既に述べたように改めて検討する必要がある。

困難度パラメタの推定値は、最小値が項目4の-2.996、最大値が項目27の0.964、平均が-0.584、標準偏差が0.881であり、30項目中24項目が負の値を示した。今回分析に用いた受験者集団にとっては比較的解答が容易な項目が多かったと言える。

テスト情報量については、尺度値が-0.5で最大の情報量3.55（標準誤差0.28）を示し、-1.3から0.3の範囲で少なくとも情報量で3.00（標準誤差0.33）を示し、-1.7から0.7の範囲で少なくとも情報量で2.50（標準誤差0.40）を示している。標準誤差が0.33および0.40というのは、信頼性係数に直すとそれぞれおよそ0.90および0.86になり、測定精度として十分に満足できる値である。

4. 図形・記号系列の作成問題

4.1. 問題の概要

本問題は、ある規則に従って並んだ図形・記号の系列が受験者に呈示され、受験者はその系列の並び方の規則を発見して、呈示された系列の次ぎ及びその次ぎに入る

べき図形・記号を解答する。本検査では、“鉄道に関係の深い素材”を用いて項目を作成するため、時刻表記号を系列の要素に用いている。系列の要素となる時刻表記号の種類が2, 3そして4種類のもの各10項目、全部で30項目から構成される。本問題も“反応の速さ”ではなく、受験者の“正誤反応”を測定することを目的としている。

各受験者毎に、解答の正誤および反応時間（0.01秒単位）が記録される。

詳しくは、野口（1993b）を参照されたい。

4.2. 項目分析の結果

本研究に用いる受験者は、1993年11月から1996年6月

に本検査を実施した名古屋市内および岡山市内の大学生192名である。本問題の各項目は、野口（1993b）で既に項目分析が実施されているが、その後項目の一部を改訂したため、今回改めて項目分析を実施した。

なお、前回の項目分析の結果から通過率が全体として高く（30項目中20項目が0.9を超える）、テストがやさしすぎる傾向が見られたので、今回からは各項目毎に反応時間が20秒を超えるケースについては受験者の解答が正答である場合でも誤答として取り扱うことにした。

①項目通過率

各項目の通過率は表4に示した通りである。項目29の.036が最も小さく、項目20の.073が次に小さい。また、項目1および12の.990が最も大きく、項目3および

表4 「図形・記号系列の作成」問題各項目の通過率・点双列相関係数、反応時間（秒）の平均・標準偏差・総反応時間との相関係数および項目パラメタ推定値

項目番号	通過率	点双列相関係数	反応時間平均	反応時間標準偏差	総反応時間との相関係数	識別力	困難度
1	.990	.176	7.847	2.560	.364	1.237	-2.795
2	.979	.091	7.224	6.844	.236	.694	-3.630
3	.984	.138	7.071	11.298	.179	.935	-3.095
4	.917	.301	11.858	5.710	.347	.844	-2.040
5	.984	.034	9.130	2.825	.324	.652	-4.091
6	.807	.200	12.491	4.954	.344	.396	-2.261
7	.792	.220	12.566	7.594	.415	.407	-2.062
8	.536	.410	22.017	15.156	.446	.657	-.155
9	.510	.511	26.020	25.898	.471	.838	-.040
10	.385	.541	31.399	25.276	.499	.996	.380
11	.891	.260	13.700	5.191	.452	.660	-2.147
12	.990	.128	8.306	2.681	.321	.951	-3.320
13	.964	.161	9.761	3.783	.450	.734	-2.995
14	.969	.113	10.717	3.685	.320	.658	-3.412
15	.865	.226	12.748	6.243	.339	.498	-2.389
16	.948	-.063	10.540	3.498	.349	.349	-5.066
17	.370	.367	32.746	22.092	.327	.545	.657
18	.563	.491	21.216	12.671	.430	.751	-.245
19	.464	.263	24.663	15.541	.411	.373	.251
20	.073	.251	51.103	38.123	.491	.665	2.595
21	.875	.170	13.428	5.362	.327	.446	-2.751
22	.458	.352	25.616	15.594	.339	.468	.235
23	.630	.431	19.640	9.454	.349	.670	-.559
24	.120	.192	50.698	32.038	.487	.438	2.889
25	.969	.169	10.666	5.173	.308	.820	-2.880
26	.255	.494	31.345	20.626	.464	.951	.892
27	.531	.540	21.851	12.156	.526	.918	-.111
28	.432	.356	25.456	16.020	.410	.519	.351
29	.036	.191	78.396	62.524	.525	.751	3.001
30	.250	.354	34.234	24.545	.401	.579	1.282

表5 「図形・記号系列作成」問題の正答数得点の分布

正答数得点	人数	百分率	正答数得点	人数	百分率
0	0	0.0	16	17	8.9
1	0	0.0	17	20	10.4
2	0	0.0	18	20	10.4
3	0	0.0	19	18	9.4
4	0	0.0	20	26	13.5
5	0	0.0	21	15	7.8
6	0	0.0	22	20	10.4
7	0	0.0	23	14	7.3
8	0	0.0	24	11	5.7
9	0	0.0	25	7	3.6
10	0	0.0	26	3	1.6
11	0	0.0	27	1	0.5
12	1	0.5	28	0	0.0
13	4	2.1	29	1	0.5
14	7	3.6	30	0	0.0
15	7	3.6			

び5の.984が次いで大きい。易しい項目から難しい項目まで広い範囲の困難度の項目が含まれているが、通過率が0.9を超える項目が10項目、0.8以上0.9未満が4項目と合わせてほぼ半数にのぼり、全体として易しい項目から構成された問題になっている。

②点双列相関係数

各項目の点双列相関係数は表4に示した通りである。項目10の.541が最も大きく、項目27の.540が次いで大きい。また、項目16の-.063が最も小さく、項目5の.034が次に小さい。全体として点双列相関係数の値が小さいが、これは通過率が全体的に高い値を示すことの影響であると考えられる。

実際に、項目5および16の通過率はそれぞれ.984、.948と高く、項目10および27の通過率はそれぞれ.385、.531と中程度の値を示している。

③項目間相関行列の因子分析

本問題中の項目の等質性を確認するために、項目間相関行列に対して因子分析を実施した。主因子解の計算が収束した時点で得られた固有値を大きいものから順に5つ示すと、2.23, 1.66, 0.70, 0.65, 0.56となり、第I固有値の優越性は見られるが、第II固有値も相対的に大きく、1因子性というよりも2因子性を示したと言える。ただし、今回は項目間相関行列に ϕ 係数を用いて因子分析を行なったため、通過率の高いことが項目間相関係数の値を見かけ上下げてしまい、その影響で1因子性が明らかにならなかった可能性がある。項目間相関係数にこの問題の生じないテトラコリッ

表6 「図形・記号系列の作成」問題における時刻表記号の種類数別の平均通過率および平均反応時間

種類数	平均通過率	平均反応時間(秒)
2	.648	14.761
3	.669	19.548
4	.524	24.162

ク相関係数を用いて再分析する必要がある。本研究では便宜上1因子性を仮定してIRT分析を進める。

④正答数得点の分布

正答数得点の分布は表5に示した通りである。最高が29点で1名、次いで27点が1名、26点が3名と続き、最低は12点で1名、次いで13点が4名、14点が7名と広い範囲に分布している。正答数得点の平均は19.5、標準偏差は3.2、歪度は0.09、尖度は-0.43であった。全体としておおよそ左右対象形の分布をしている。

⑤総反応時間の分布

各受験者毎に全ての項目に要した反応時間の合計した総反応時間の分布について、平均、標準偏差、歪度、尖度を計算すると、順に、654.4秒、183.9秒、0.57、0.62である。やや右に裾を引く形状を示している。総反応時間の短い受験者及び長い受験者各5名分の実測値は、前者が最短の方から、275.7、284.7、299.3、313.4、361.4秒であり、後者が最長の方から、1374.5、1118.4、1117.4、1097.5、1063.0秒であった。5分弱で本問題を終了した受験者がいる一方で、23分近くかけて解答した受験者もあり、「空間図形の推理問題」

と同様に、総反応時間には大きな個人差が見られた。

⑥項目反応時間

各項目毎の受験者の反応時間の平均・標準偏差及び総反応時間との相関係数は表4に示した通りである。

反応時間の平均は、項目3で7.07秒と最も短く、項目2の7.22秒が次いで短く、項目29の78.40秒が最も長く、次いで項目20の51.10秒が長い。反応時間の標準偏差は、項目1で2.56秒と最も小さく、項目12の2.68秒が次いで小さく、項目29の62.52秒が最も大きく、項目20の38.12秒が次いで大きくなっている。また、総反応時間との相関係数は、項目27および29で0.53と最も大きく、項目10の0.50が次いで大きく、項目3の0.18が最も小さく、項目2の0.24が次いで小さい値を示した。

さらに、全項目間の反応時間間相関行列に対して因子分析を実施した。主因子解の計算が収束した時点で得られた固有値を大きいものから順に5つ示すと、5.70, 2.34, 1.77, 1.59, 1.37となり、第I固有値と第II固有値の比が2.4である。項目間相関行列と同様に明確な1因子性は確認できないが、第II以下の固有値が大きな段差なく並んでいることから、項目反応時間に関しては項目得点とは異なり、1因子性が十分に満たされていると見做せる。

⑦時刻表記号の種類数別の分析

時刻表記号の種類数別に平均通過率および平均反応時間を求めた結果は表6に示したとおりである。

平均通過率を種類数別に見ると、種類数2, 3, 4の順に.648, .669, .524, となり、種類数2よりも3の方がわずかに高い値を示しているが、種類数4では明らかに下がっている。平均反応時間を種類数別に見ると、種類数2, 3, 4の順に14.761秒, 19.548秒, 24.162秒, と順に値が大きくなっている。これらを合わせると、種類数の増加が項目の課題をより複雑にしていることが明らかである。

4.3. 項目反応モデルによる分析

4.3.1. 方法

本問題でも「空間図形の推理問題」と同様に、受験者の解答に基づいて2パラメタ・ロジスティック・モデルの項目パラメタ値を推定するが、本問題の場合、各項目の解答時間が20秒を超えた場合には受験者の解答が正しい場合にも誤答として取り扱う点が異なっている。

そして、「空間図形の推理問題」と同様に受験者の項目反応データからBLI OG 3を用いて各項目の識別力および困難度パラメタの推定値を計算する。尺度は、受験者母集団の特性尺度値の分布の平均が0.0, 標準偏差が1.0となるように原点と単位とを設定する。

さらに、得られた項目パラメタ推定値にもとづいて、

表7 「図形・記号系列の作成」問題のテスト情報量

尺度値	-3.0	-2.5	-2.0	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5	2.0	2.5	3.0
情報量	2.36	2.33	2.17	2.05	2.04	2.12	2.18	2.14	1.98	1.75	1.54	1.37	1.23
標準誤差	0.42	0.43	0.46	0.49	0.49	0.47	0.46	0.47	0.51	0.57	0.65	0.73	0.81

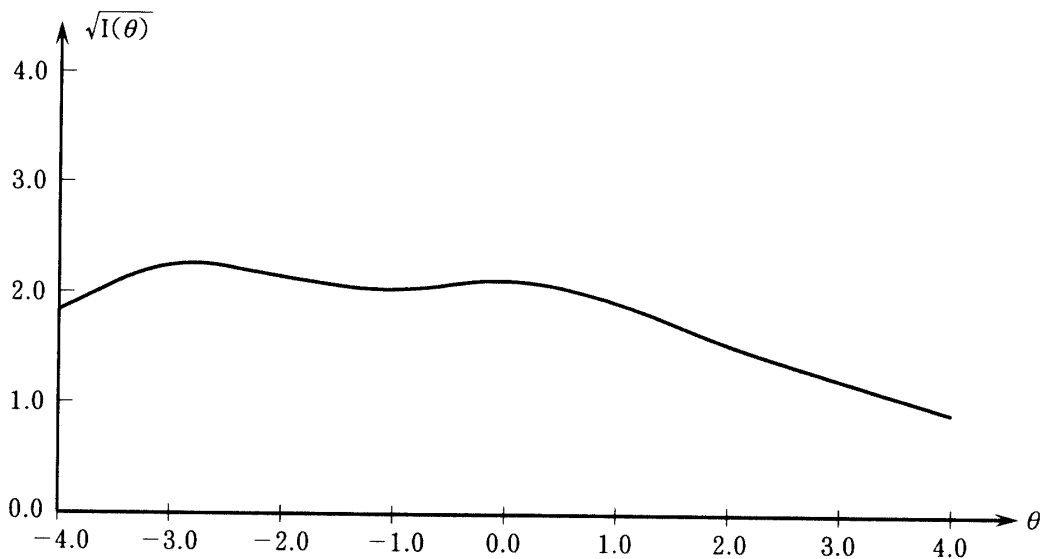


図2 「図形・記号等の系列作成問題」のテスト情報量

テスト情報量を計算する。

4.3.2. 結果および評価

項目パラメタ推定値は表4に示したとおりである。また、これらの30項目から得られるテスト情報量は表7および図2に示したとおりである。

識別力パラメタの推定値は、最小値が項目16の0.35、最大値が項目1の1.24、平均が0.680、標準偏差が0.217である。全体として識別力パラメタの値が均質になっており、十分な識別力を示している。

困難度パラメタの推定値は、最小値が項目16の-5.066、最大値が項目29の3.001、平均が-1.117、標準偏差が2.155であり、30項目中20項目が負の値を示した。今回分析に用いた受験者集団にとっては比較的解答が容易な項目が多かったが、パラメタ値が2.0を超える項目も3項目あり、困難度に広がりを持つ問題になっている。

テスト情報量については、尺度値が0.1で最大の情報量2.19（標準誤差0.46）を示し、-3.8から0.9の範囲で少なくとも情報量で2.00（標準誤差0.50）を示している。標準誤差が0.50というのは、信頼性係数に直すと0.80になり、測定精度としては必ずしも十分に満足できる値となっていないが、尺度値の広い範囲でこのような情報量を示すことが特徴的である。

5. 図形・記号等系列の記憶問題

5.1. 問題の概要

本問題は、順唱課題と逆唱課題の二つの部分から構成される。受験者は、画面中に順次呈示された複数の図形・記号等（いずれも呈示時間は1.0秒）について、その系列を呈示順に記憶し、順唱課題では図形・記号等の系列をそのままの順序で再生し、逆唱課題では図形・記号等の系列を逆順で再生して解答する。従来の識別性検査には含まれていなかった課題であり、パーソナル・コンピュータをテスト過程の制御に用いることによって実施が可能になったものである。

図形・記号等には具体的には、数字、色、時刻表記号、駅名、列車種別、の5種類を用いる。呈示する図形・記号等の数を、順唱課題では4、5、6、7と、逆唱課題では3、4、5、6と変えることによって項目の困難度を予め変化させてある。項目数は、順唱課題が20、逆唱課題が20の、全部で40項目から構成される。

各受験者毎に、解答・正誤および反応時間（0.01秒単位）が記録される。

詳しくは、野口（1993b）を参照されたい。

5.2. 項目分析の結果

本研究に用いる受験者は、1993年11月から1996年6月に本検査を実施した名古屋市内および岡山市内の大学生192名である。本問題の各項目は、野口（1993b）で既に項目分析が実施されているが、その後項目を構成する各図形・記号等の呈示時間を1.5秒から1.0秒に変更したため、今回改めて項目分析を実施した。

①項目通過率

各項目の通過率は表8に示した通りである。項目8の.099が最も小さく、項目12の.115が次に小さい。また、項目21の.984が最も大きく、項目1および17の.943が次いで大きい。そして、0.0以上で0.2未満が3項目、0.2以上で0.4未満が4項目、0.4以上で0.6未満が10項目、0.6以上で0.8未満が10項目、0.8以上で1.0までが13項目あり、易しい項目から難しい項目まで広い範囲の困難度の項目が含まれている。

②点双列相関係数

各項目の点双列相関係数は表8に示した通りである。項目26の.509が最も大きく、項目31の.502が次いで大きい。また、項目1の.107が最も小さく、項目25の.124が次に小さい。全体として点双列相関係数が比較的高い値を示している。

③項目間相関行列の因子分析

本問題中の項目の等質性を確認するために、項目間相関行列に対して因子分析を実施した。主因子解の計算が収束した時点で得られた固有値を大きいものから順に5つ示すと、4.30、1.63、1.15、1.04、1.00となり、第1固有値の優越性は見られるが、第2固有値も相対的にやや大きい値を示している。しかしながら第1固有値と第2固有値との比は2.64であり、第1固有値の優越性は明らかで、1因子性が示されたと言える。

④正答数得点の分布

正答数得点の分布は表9に示した通りである。最高が38点で1名、次いで37点が2名、36点が3名と続き、最低は9点で1名、次いで10点が1名、12点が1名と広い範囲に分布している。正答数得点の平均は25.8、標準偏差は5.5、歪度は-0.36、尖度は0.09であった。全体としてわずかに左に裾を引く分布を示している。

⑤総反応時間の分布

各受験者毎に全ての項目に要した反応時間の合計した総反応時間の分布について、平均、標準偏差、歪度、尖度を計算すると、順に、418.1秒、84.2秒、0.38、-0.14である。やや右に裾を引く形状を示している。総反応時間の短い受験者及び長い受験者各5名分の実測値は、前者が最短の方から、224.8、229.0、252.4、262.4、272.5秒であり、後者が最長の方から、631.7、

表 8 「図形記号等系列の記憶」問題各項目の図形・記号等の種類・数、通過率・点双列相関係数、反応時間(秒)の平均・標準偏差・総反応時間との相関係数および項目パラメタ推定値

項目番号	種類	接する数	通過率	点双列 相関係数	反応時間(秒) の平均	標準偏差	総反応時間と の相関係数	識別力	困難度
1	数字	4	943	107	11.239	5.105	343	.517	-3.471
2	数字	5	927	397	10.109	4.851	400	1.265	-1.722
3	数字	6	760	340	12.374	6.291	383	.591	-1.336
4	数字	7	755	392	13.766	7.443	301	.724	-1.125
5	色	4	714	288	6.858	2.629	316	.476	-1.263
6	色	5	599	256	9.285	4.549	337	.426	-.625
7	色	6	339	355	11.671	6.608	383	.594	.737
8	色	7	099	369	13.974	7.810	430	1.075	1.612
9	時刻表	4	536	383	8.479	4.723	372	.599	-.191
10	時刻表	5	490	318	9.196	4.564	450	.490	.030
11	時刻表	6	417	257	10.364	5.748	365	.428	.477
12	時刻表	7	115	314	13.469	6.767	347	.854	1.725
13	駅名	4	932	151	7.151	2.412	301	.551	-3.092
14	駅名	5	734	297	8.808	5.613	410	.480	-1.392
15	駅名	6	385	309	14.379	7.867	356	.513	.572
16	駅名	7	198	340	16.727	10.446	366	.692	1.381
17	列車種	4	943	144	5.887	2.927	353	.585	-3.133
18	列車種	5	719	476	8.373	5.175	372	.954	-.794
19	列車種	6	635	410	9.605	5.062	328	.666	-.606
20	列車種	7	458	316	11.592	6.073	435	.492	.196
21	数字	3	984	195	6.811	2.813	421	1.040	-2.939
22	数字	4	938	275	9.442	3.972	346	.872	-2.258
23	数字	5	917	146	11.145	7.085	397	.539	-2.895
24	数字	6	745	464	14.769	6.359	373	.928	-.914
25	色	3	885	124	6.234	2.915	309	.437	-2.966
26	色	4	656	509	10.095	5.887	277	.991	-.548
27	色	5	552	484	12.658	6.818	438	.809	-.218
28	色	6	479	451	15.684	10.711	417	.733	.055
29	時刻表	3	708	355	5.975	3.215	348	.567	-1.068
30	時刻表	4	651	371	8.356	4.208	417	.640	-.696
31	時刻表	5	469	502	12.195	8.572	446	.970	.075
32	時刻表	6	391	416	13.671	8.853	443	.680	.432
33	駅名	3	901	200	6.982	4.269	336	.552	-2.622
34	駅名	4	844	325	9.340	3.694	391	.686	-1.717
35	駅名	5	464	418	13.196	6.293	337	.627	.134
36	駅名	6	276	466	16.753	9.114	415	.981	.752
37	列車種	3	922	395	5.665	2.298	322	1.084	-1.818
38	列車種	4	927	182	5.203	2.607	346	.577	-2.891
39	列車種	5	891	163	7.693	2.960	318	.496	-2.729
40	列車種	6	547	416	12.942	7.133	383	.662	-.224

表9 正答数得点の分布

正答数得点	人数	百分率	正答数得点	人数	百分率
9	1	0.5	25	12	6.3
10	1	0.5	26	13	6.8
11	0	0.0	27	13	6.8
12	1	0.5	28	13	6.8
13	1	0.5	29	14	7.3
14	3	1.6	30	11	5.7
15	0	0.0	31	7	3.6
16	3	1.6	32	11	5.7
17	4	2.1	33	9	4.7
18	6	3.1	34	3	1.6
19	3	1.6	35	3	1.6
20	8	4.2	36	3	1.6
21	6	3.1	37	2	1.0
22	8	4.2	38	1	0.5
23	17	8.9	39	0	0.0
24	15	7.8	40	0	0.0

628.9, 617.2, 616.9, 615.3秒であった。4分弱で本問題を終了した受験者がいる一方で、10分以上かけて解答した受験者もあり、「空間図形の推理問題」と同様に、総反応時間には大きな個人差が見られた。

⑥項目反応時間

各項目毎の受験者の反応時間の平均・標準偏差及び総反応時間との相関係数は表8に示した通りである。

反応時間の平均は、項目38で5.20秒と最も短く、項目37の5.67秒が次いで短く、項目36の16.75秒が最も長く、次いで項目16の16.73秒が長い。反応時間の標準偏差は、項目37で2.30秒と最も小さく、項目13の2.41秒が次いで小さく、項目28の10.71秒が最も大きく、項目16の10.45秒が次いで大きくなっている。また、総反応時間との相関係数は、項目10および31で0.45と最も大きく、項目32の0.44が次いで大きく、項目26の0.28が最も小さく、項目4および13の0.30が次いで小さい値を示した。

さらに、全項目間の反応時間間相関行列に対して因子分析を実施した。主因子解の計算が収束した時点で得られた固有値を大きいものから順に5つ示すと、5.23, 2.30, 1.72, 1.20, 1.11となり、第I固有値と第II固有値の比がおよそ2.3である。第II固有値がそれ以下に比べてやや大きい値を示しているが、第I固有値との段差が大きいことから1因子性が満たされていると見做せる。

⑦図形・記号等および順唱・逆唱別の分析

図形・記号等および順唱・逆唱別に平均通過率および平均反応時間を求めた結果は表10に示したとおりで

表10 図形・記号等および順唱・逆唱別の平均通過率および平均反応時間

数 字	順唱課題	逆唱課題
数 字	0.846	0.896
	11.872	10.542
色	0.438	0.643
	10.447	11.168
時刻表記号	0.390	0.555
	10.377	10.049
駅 名	0.562	0.621
	11.766	11.568
列車種別	0.689	0.822
	8.864	7.876

上段：平均通過率 下段：平均反応時間（秒）

ある。

平均通過率を図形・記号等別に見ると、順唱課題では数字から順に0.846, 0.438, 0.390, 0.562, 0.689, 逆唱課題では同じく順に0.896, 0.643, 0.555, 0.621, 0.822であり何れの場合にも数字・列車種別で高く、色・時刻表記号および駅名で低くなっている。「数字」が両方で特に高い値を示すのは、記憶する際にコード化する必要がなく、記憶が容易であるためと思われる。

平均反応時間を図形・記号等別に見ると、順唱課題では数字から順に11.872, 10.447, 10.377, 11.766, 8.864, 逆唱課題では同じく順に10.542, 11.168, 10.049, 11.568, 7.876であり何れの場合にも列車種別やや短い傾向が見られる以外は大きな差は見られず

表11 「図形・記号等系列の記憶」問題のテスト情報量

尺度値	-3.0	-2.5	-2.0	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5	2.0	2.5	3.0
情報量	2.25	2.55	2.83	2.98	3.02	2.99	2.91	2.71	2.50	2.25	1.93	1.57	1.22
標準誤差	0.44	0.39	0.35	0.34	0.33	0.33	0.34	0.37	0.40	0.44	0.52	0.64	0.82

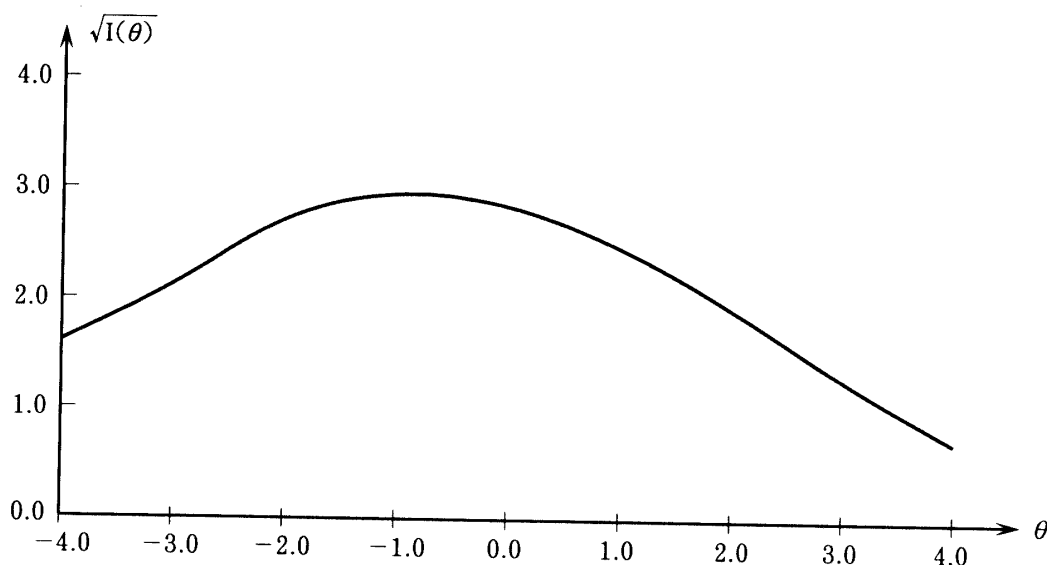


図3 「図形・記号等の系列記憶問題」のテスト情報量

図形・記号等の違いが反応時間に実質的な影響を与えているとは思われない。

順唱・逆唱の間では、同一の図形・記号等の中で比較すると、通過率については、ほぼ等しいかまたは逆唱課題の方が高い値を示している。これは呈示する図形・記号等の数が、順唱課題で4, 5, 6, 7, 逆唱課題で3, 4, 5, 6, と逆唱課題の方が1つつ少なくなっていることの影響である。

5.3. 項目反応モデルによる分析

5.3.1. 方法

本問題でも「空間図形の推理問題」と同様に、2パラメタ・ロジスティック・モデルを適用して、BILOG 3を用いて計算する。尺度は、受験者母集団の特性尺度値の分布の平均が0.0, 標準偏差が1.0となるように原点と単位とを設定する。

さらに、得られた項目パラメタ推定値にもとづいて、テスト情報量を計算する。

5.3.2. 結果および評価

項目パラメタ推定値は表8に示したとおりである。また、これらの40項目から得られるテスト情報量は表11および図3に示したとおりである。

識別力パラメタの推定値は、最小値が項目6の0.43, 最大値が項目2の1.27, 平均が0.696, 標準偏差が0.216

である。全体として識別力パラメタの値が均質になっており、十分な識別力を示している。

困難度パラメタの推定値は、最小値が項目1の-3.47, 最大値が項目12の1.73, 平均が-0.927, 標準偏差が1.43であり、40項目中27項目が負の値を示した。今回分析に用いた受験者集団にとっては比較的解答が容易な項目が多かったと言える。

テスト情報量については、尺度値が-1.0で最大の情報量3.02(標準誤差0.33)を示し、-1.4から-0.6の範囲で少なくとも情報量で3.00(標準誤差0.33)を示し、-2.5から1.0の範囲で少なくとも情報量で2.50(標準誤差0.40)を示している。標準誤差が0.33および0.40というのは、信頼性係数に直すとそれぞれおよそ0.90および0.86になり、測定精度として十分に満足できる値である。

6. まとめ

本研究では、識別性検査A-1001の「空間図形の推理問題」, 「図形・記号等の系列作成問題」および「図形・記号等系列の記憶問題」に対して、2パラメタ・ロジスティック・モデルを適用してIRT尺度を構成することを試みた。その結果、何れの項目についてもパラメタ推定値が特に極端な値を示すこともなく、テスト情報量の観点からも実用化に向けて今後の作業を進める方向性が示された。

しかしながら、今後引き続き検討しなければならない問題も多数残されている。

第一に、項目パラメタ値の推定に用いた受験者数である。本研究では、「空間図形の推理問題」から順に378名、192名、192名であった。Hulin, Lissak & Drasgow (1982) では2パラメタ・ロジスティック・モデルの場合に項目数が30で受験者数は500程度は必要であるとしている。もちろん、受験者集団の性質やパラメタ推定法によりパラメタ値推定に必要な受験者数は変動するであろうが、「図形・記号等の系列作成問題」および「図形・記号等系列の記憶問題」の192名は、さらに受験者数を増やして再尺度化する必要がある。

次に、「空間図形の推理問題」と「図形・記号等の系列作成問題」はいずれも「関係判断力・応用力領域」に属する問題であるが、今回は両者を独立にIRT尺度を構成した。しかしながら、実際に本検査を適用する場合には各問題毎に尺度値を表示するのではなく、「知覚の速さ・正確さ領域」、「関係判断力・応用力領域」、「記憶領域」の3領域のそれぞれ毎に受験者の測定結果を表示することが望まれる。したがって、今後例えば、両者に共通する尺度を構成するか、あるいは独立に構成された尺度の測定結果から合成得点を求めるのか等、「空間図形の推理問題」と「図形・記号等の系列作成問題」の結果を総合する方法について検討する必要がある。

さらに、本検査の測定結果を表示する尺度が確定したとして、尺度値を解釈する規準を設定する作業が必要である。これには、本研究でご協力を得た一般社会人や大学生を受験者とするのではなく、実際に本検査が適用される集団、すなわち鉄道事業の運転関係従事員およびその候補者に対して本検査の受験をお願いする必要がある。

また、従来から用いられてきた識別性検査、例えば「鉄道総研式識別性検査J-1001」(小笠原, 1991)などとの併存的妥当性についても検討しておく必要がある。本研究はIRT尺度の構成を中心にして進めており、いわば測定の信頼性の側面が主として強調されたが、測定の妥当性の観点からも十分に検討を進める必要がある。

そして、識別性検査A-1001を開発するにあたっての条件のひとつである、適応型テスト方式による検査実施法について、例えば、項目固定型多段階テスト方式を用いるのか項目可変型多段階テスト方式を用いるのか、検査の切りは実施項目数で行なうのか推定尺度値の精度で行なうのか等、具体的な仕様を決定し実際のデータに基づいてその効果について検討を進めなければならない。

文 献

- Lord, F. M. & Novick, M. R. 1968 Statistical theories of mental test scores., Addison-Wesley, Reading, MA.
- Hambleton, R. K. & Swaminathan, H. 1985 Item response theory - Principles and applications -., Kluwer-Nijhoff, Boston, MA.
- Hulin, C. L., Lissak, R. I. & Drasgow, F. 1982 Recovery of two- and three-parameter logistic item characteristic curves: A Monte Carlo study., Applied Psychological Measurement, 6, 249-260.
- 野口 裕之 1993a コンピュータを用いた適応型知能検査の開発(1), 東日本旅客鉄道安全研究所委託研究報告書.
- 野口 裕之 1993b パーソナル・コンピュータをベースとする識別性検査A-1001用項目プールの作成., 名古屋大学教育学部紀要——教育心理学科——, 40, 141-158.
- 野口 裕之 1994a コンピュータを用いた適応型知能検査の開発(2), 東日本旅客鉄道安全研究所委託研究報告書.
- 野口 裕之 1994b パーソナル・コンピュータをベースにした識別性検査に対する受験者の反応., 名古屋大学教育学部紀要——教育心理学科——, 41, 25-38.
- 野口 裕之 1995a 識別性検査A-1001における記憶問題のIRT尺度化., 日本教育心理学会第37回総会発表論文集.
- 野口 裕之 1995b 識別性検査A-1001の「知覚の速さ・正確さ」領域のIRT尺度化., 名古屋大学教育学部紀要——教育心理学科——, 42, 59-71.
- 小笠原 春彦 1991 鉄道総研式識別性検査J-1001の開発., 鉄道総研報告第5巻 3号, 19-26.
- 芝 祐順 編著 1991 項目反応理論——基礎と応用——., 東京大学出版会.
- 渡辺 直登・野口 裕之 編著 印刷中 項目反応理論——組織心理測定へのニューアプローチ——
(1996年□月□日 受稿)

ABSTRACT

IRT scaling for Spatial reasoning problems, Figure series
completion problems and Memory problems in adult
intelligence test : A-1001

NOGUCHI Hiroyuki

In this paper, IRT scaling was attempted for subtests in adult intelligence test : A-1001, that were intended to measure the reasoning and spatial cognitive abilities and the ability of memory.

2-parameter logistic model was adopted in IRT scaling and computer program BILOG was used for estimating item discrimination and item difficulty parameters. Estimated item parameters and test information curve in each subtest showed satisfactory results for actual testing situations. This results indicate the applicability of these items in each subtest to the item pools for adaptive testing version of adult intelligence test : A-1001.

Key words : adult intelligence test, spatial reasoning, memory, time response theory, 2-parameter logistic model