

識別性検査 A-1001の「知覚の速さ・正確さ」 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、新たに開発されたものである。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
 - ③ 「空間図形の推理」問題
 - ④ 「図形・記号系列の作成」問題
- III 「記憶」領域
 - ⑤ 「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尺度化の試みが実施されている(野口,1995)。

ところで、「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」
問題は、受験者の誰もが時間をかけさえすれば正答を得
られるような困難度の低い項目から構成されており、速
く正確に解答することが要求される、いわゆる速度検査
になっている。速度検査課題について通常は一定の制限
時間の下での作業量が問題にされる。例えば、「鉄道総
研式識別性検査 J-1001」(小笠原,1991)では、「記号
と数字の置き換え」と「図形の探索」の2問題が含まれ

ているが、採点にあたっては作業量が問題にされてい
る。

これに対して、「識別性検査 A-1001」は最終的には、
項目反応理論に基づいて尺度を構成し、適応型テスト方
式で検査を実施するため、受験者毎に解答する項目およ
び検査所要時間が異なり、従って作業量をもとに得点化
することができない。

一方、「識別性検査 A-1001」はパーソナル・コンピュ
ータを用いて検査を実施するため、各項目毎に受験者の反
応時間を記録することが可能であり、これを直接利用し
て得点化するという考え方もある。しかしながら、測定
された反応時間は、

- ①反応時間には、解答の際の動作時間(例えば、マウ
スの操作に要する時間)も含まれていて、測定目的
からすると誤差になる
- ②時間という比尺度が、そのまま直ちに測定しようと
する能力の比尺度に対応する保証がない
- ③正答に対する反応時間と誤答に対する反応時間の取
り扱いを分ける必要がある

などの問題がある。従って、反応時間を情報として組み
込んだ能力尺度構成モデルが必要になる。

項目反応モデルに解答時間を組み込む試みは、これま
でも提案されていたが、実用場面で利用された例は見
られない。これは、

- ①受験者の各項目に対する解答時間がデータとして得
にくかった
- ②モデルを実際に適用するのに、例えば、モデル中に
含まれるパラメタ数が多いために精度の良い推定値
を得るのに膨大な受験者数のデータを集めなければ
ならない、パラメタ推定用のコンピュータ・プログ
ラムが入手しにくい、など種々の困難があった

などの理由による。

そこで、本研究では広く普及している項目反応モデル及びそのパラメタ推定プログラムを利用し、その適用の仕方に工夫をこらすことによって、各項目に対する解答時間情報も組み込んだIRT尺度を構成することを試みる。

具体的には、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ることでカテゴリを設定し、既存の多値型項目反応モデル用の分析プログラム(MULTILOG)を利用し、「段階反応モデル(Graded response model)」に基づいて、「識別性検査A-1001」の「図形・記号の探索」問題および「図形・記号の異同弁別」問題に対して、時間情報を組み込んだIRT尺度を構成することを試みる。

2. 段階反応モデル

テストの分析に用いられる項目反応モデルは、観測変量が2値型であるか多値型であるかによって大きく2つに分類される。観測変量が2値型のものには、「1パラメタ・ロジスティック・モデル」、「2パラメタ・ロジスティック・モデル」、「3パラメタ・ロジスティック・モデル」などがあり、従来はこのタイプのモデルがよく用いられてきた。これに対して、観測変量が多値型のものには、「名義反応モデル」、「段階反応モデル」等がある。前者は、多値型の反応が名義的なカテゴリで与えられている場合に、後者は、多値型の反応が段階(順序)づけられたカテゴリで与えられる場合に用いられる(例えば、Baker:1992 参照)。ここで、本研究に用いる「段階反応モデル」について簡単にまとめておく。

いま、受験者の反応が{1, 2, 3, 4}の4段階で得られ、特性(能力)尺度値が大きい受験者ほどカテゴリ4に、小さい受験者ほどカテゴリ1に反応する確率が高くなるものとする。この時、2パラメタ・ロジスティック

・モデルにおける項目特性曲線と同様に、各項目のカテゴリ毎に特性(能力)尺度値と当該カテゴリに反応する確率との関係を表わす項目反応カテゴリ特性曲線(Item Response Category Characteristic Curve, IRCCCと略す)を設定する。

しかしながら、全てのIRCCCを統一的に表現でき、しかも数学的に取り扱い易い関数が存在しない為、境界特性曲線(Boundary Characteristic Curve, BCCと略す)を設定し、それにロジスティック曲線を用いる。BCCは、特性(能力)尺度値と当該カテゴリ以上の値をもつカテゴリに反応する確率との関係を表わす。すなわち、特性(能力)尺度値が θ で、カテゴリ{1, 2, 3, 4}に反応する確率を $P_0^*(\theta)$ 、カテゴリ{2, 3, 4}に反応する確率を $P_1^*(\theta)$ 、カテゴリ{3, 4}に反応する確率を $P_2^*(\theta)$ 、カテゴリ{4}に反応する確率を $P_3^*(\theta)$ 、カテゴリ{・}に反応する確率を $P_4^*(\theta)$ として、 $P_1^*(\theta)$ 、 $P_2^*(\theta)$ 、 $P_3^*(\theta)$ にロジスティック曲線を仮定する。すなわち、

$$P_k^*(\theta) = [1 + \exp \{-Da(\theta - b_k)\}]^{-1} \quad (1)$$

$$k = 1, 2, 3$$

で表わされる。パラメタaは同一項目のBCCでは同じ値をとることがモデル上仮定される。なお、ここでは式が繁雑になるのを避ける為、項目を表わす添字は省略した。 $P_0^*(\theta)$ は常に1.0、 $P_4^*(\theta)$ は常に0.0であり、これらはモデル表現上の必要性から設定される(図1)。 $P_1^*(\theta)$ はカテゴリ{2, 3, 4}を正答、 $P_2^*(\theta)$ はカテゴリ{3, 4}を正答、 $P_3^*(\theta)$ はカテゴリ{4}を正答として採点した時に得られる項目特性曲線と考えれば、段階反応モデルが2パラメタ・ロジスティック・モデルの自然な拡張モデルであることが明らかになる。

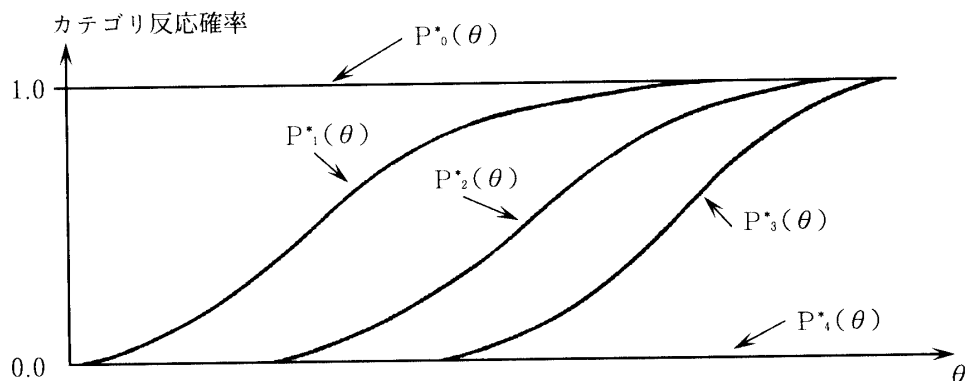


図1 境界特性曲線(BCC)

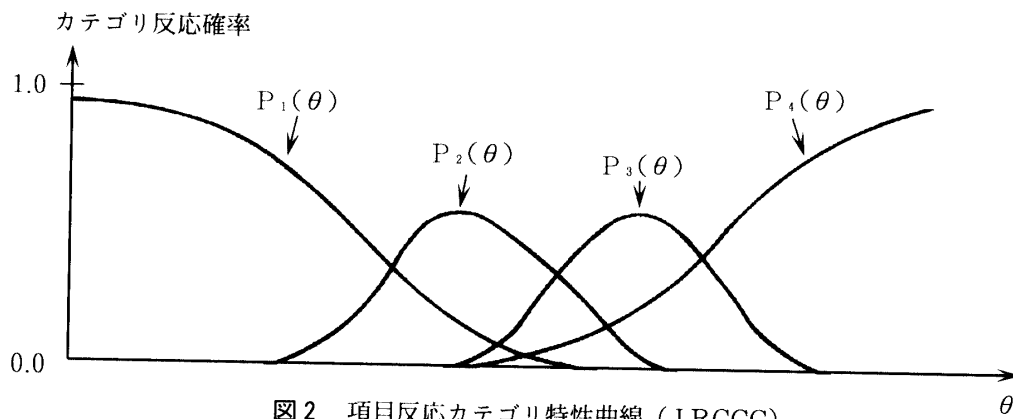


図2 項目反応カテゴリ特性曲線 (IRCCC)

このようにBCCが定義されると、各カテゴリのIRCCCはBCCの差で表わされる。例えば、特性尺度値 θ でカテゴリ2に反応する確率 $P_2(\theta)$ は、カテゴリ{2, 3, 4}に反応する確率 $P_1^*(\theta)$ とカテゴリ{3, 4}に反応する確率 $P_3^*(\theta)$ との差で与えられる。すなわち、

カテゴリ {1} に反応する確率は、

$$P_1(\theta) = P_0^*(\theta) - P_1^*(\theta), \quad (2)$$

カテゴリ {2} に反応する確率は、

$$P_2(\theta) = P_1^*(\theta) - P_2^*(\theta), \quad (3)$$

カテゴリ {3} に反応する確率は、

$$P_3(\theta) = P_2^*(\theta) - P_3^*(\theta), \quad (4)$$

カテゴリ {4} に反応する確率は、

$$P_4(\theta) = P_3^*(\theta) - P_4^*(\theta), \quad (5)$$

で与えられる。

従って、IRCCCの具体的な曲線型は、カテゴリ1が右下がりの単調減少型、カテゴリ2及び3が左右対称な単峰型、カテゴリ4が右上がりの単調増加型になる(図2)。各カテゴリの位置を表すのに、カテゴリ1及び4についてはIRCCCの確率0.5に対応する横軸 θ 上の値を、カテゴリ2及び3についてはIRCCCのピークに対応する横軸 θ 上の値を用いる。カテゴリ1から4の位置を b_1' , b_2' , b_3' , b_4' とすると、これらはBCCのパラメタを用いて、

$$\begin{aligned} b_1' &= b_1 \\ b_2' &= (b_1 + b_2) / 2 \\ b_3' &= (b_2 + b_3) / 2 \\ b_4' &= b_3 \end{aligned}$$

で得られる (Baker, 1992)。

結局、段階反応モデルは、各項目毎に設定されるパラ

メタ a 及び b_1, b_2, \dots, b_m (m は当該項目の段階数)を全項目にわたって推定する必要がある。パラメタ推定プログラムには例えば、MULTILOG (Thissen, 1991)が利用できる。

また、受験者個人の実験特性(能力)尺度値は、当該受験者の項目カテゴリ反応パターンと各反応カテゴリのIRCCCを用いて最尤推定する。この場合もMULTILOGが利用できる。

3. 図形・記号の探索問題

3.1. 問題の概要

本問題では、指定された図形・記号等と同一のものを、選択枝中から見出して解答する。各項目は全て課題としては非常に易しいものであり、受験者の“正誤反応”よりもむしろ“反応の速さ”を測定することを主たる目的としている。図形・記号等には具体的には、機関車の前頭部に装着されるトレインマーク(「トレインマーク」と略す)、時刻表で用いられる記号(「時刻表記号」と略す)、JRの文字をデザイン化し色を変化させたもの(「JRマーク」と略す)、鉄道車両の正面、の4種類を用いる。各種類毎に20項目、全部で80項目から構成される。そして、各受験者毎に、解答選択枝・正誤及び反応時間(0.01秒単位)が記録される。なお、本問題は、野口(1993a)以降に項目の呈示時間を若干変更し、改訂している。

3.2. 項目分析の結果

本研究に用いる受験者は、改訂後の1993年10月から1995年3月に集められた名古屋市内および岡山市内の大学生240名である。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1に示した通りである。項目79の.944が最も小さく、項目75の.948が次に小さい。項目4, 5, 10, 13, 14, 15, 17, 18, 21, 22, 25,

29, 31, 33, 35, 36, 38, 39, 40, 41, 44, 46, 48, 53, 54, 58, 66, 68, 70, 73, 74, 77, 78の33項目が1.00であり、全体として極めて高い値を示した。

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。ただし、通過率の値が1.00である上記の33項目については点双列相関係数を計算する事ができない。その他の47項目については、項目3, 7, 34, 57, 76の.018が最も小さく、中には項目52の.527, 項目55の.450, 項目51の.439などの様に大きい値をとるものもあるが、全体として点双列相関係数の値は小さい。これは項目通過率の値が全体として極めて高い値を示している事が影響しているものである。

③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。74点が3名, 75点が1名, 76点が3名, 77点が11名, 78点が23名, 79点が56名, 80点が153名と、約3/5の受験者が全項目に正答しており、全体として、正答数得点は極めて高い。なお、正答数得点の分布の平均は79.3, 標準偏差は1.1, 歪度は-2.3, 尖度は6.3で、左に裾を引く尖った分布を示している。

④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を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、97.1秒, 16.6秒, 1.1, 2.9, 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、63.5, 66.6, 68.6, 68.8, 69.09秒であり、後者が最長の方から、179.8, 166.1, 146.5, 140.8, 140.4秒であった。

⑤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22で0.81秒と最も短く、項目11で1.79秒と最も長く、反応時間の標準偏差は、項目70で0.19秒と最も小さく、項目37で0.75秒と最も大きくなっている。項目に用いられている図形・記号等の種類毎(各20項目)に平均項目反応時間の平均値を計算すると、トレインマークで1.339秒, 時刻表記号で1.171秒, JRマークで1.126秒, そして鉄道車両の正面で1.221秒になる。より複雑な刺激で大きな値となる傾向が見られるが絶対的な差はそれ程大きなものではない。また、総反応時間との相関係数は、項目8で0.70と最も大きく、項目69で0.11と最も小さくなっている。

さらに、全項目間の反応時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、22.0, 3.5, 2.8, 2.3, 2.2となり、第I固有値と第II固有値の比が6.3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。

表1 「図形・記号の探索」問題各項目の通過率・点双列相関係数及び反応時間の平均・標準偏差・総反応時間との相関係数

| 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 |
|------|------|---------|--------|----------|-----------|
| 1 | 1000 | — | 1.555 | .493 | 499 |
| 2 | 988 | 130 | 1.691 | .578 | 571 |
| 3 | 996 | 018 | 1.213 | .387 | 590 |
| 4 | 1000 | — | 1.108 | .412 | 457 |
| 5 | 1000 | — | 1.145 | .425 | 538 |
| 6 | 996 | 075 | 1.633 | .545 | 660 |
| 7 | 996 | 018 | 1.661 | .491 | 594 |
| 8 | 992 | 146 | 1.602 | .643 | 700 |
| 9 | 992 | 227 | 1.433 | .623 | 545 |
| 10 | 1000 | — | 1.116 | .363 | 685 |
| 11 | 984 | 151 | 1.786 | .602 | 548 |
| 12 | 988 | 130 | 1.211 | .372 | 638 |
| 13 | 1000 | — | 1.167 | .383 | 593 |
| 14 | 1000 | — | 1.346 | .495 | 636 |
| 15 | 1000 | — | 1.023 | .297 | 335 |
| 16 | 992 | 066 | 1.298 | .380 | 534 |
| 17 | 1000 | — | .834 | .259 | 341 |
| 18 | 1000 | — | 1.332 | .336 | 594 |
| 19 | 988 | 196 | 1.331 | .490 | 403 |
| 20 | 988 | 130 | 1.295 | .710 | 234 |
| 21 | 1000 | — | 1.177 | .445 | 451 |

原 著

| | | | | | |
|----|------|-----|-------|---------|-------|
| 22 | 1000 | — | .808 | .237 | 337 |
| 23 | 992 | 066 | 1.292 | .395 | 512 |
| 24 | 996 | 075 | 1.394 | .392 | 533 |
| 25 | 1000 | — | 1.621 | .455 | 633 |
| 26 | 992 | 106 | 1.549 | .472 | 508 |
| 27 | 992 | 026 | 1.603 | .595 | 526 |
| 28 | 988 | 130 | 1.031 | .298 | 493 |
| 29 | 1000 | — | 1.242 | .376 | 467 |
| 30 | 988 | 163 | 1.072 | .270 | 609 |
| 31 | 1000 | — | 1.095 | .299 | 560 |
| 32 | 996 | 075 | 1.135 | .281 | 621 |
| 33 | 1000 | — | 1.232 | .321 | 510 |
| 34 | 996 | 018 | 1.007 | .258 | 366 |
| 35 | 1000 | — | .902 | .252 | 577 |
| 36 | 1000 | — | 1.252 | .309 | 647 |
| 37 | 996 | 302 | 1.210 | .753 | 248 |
| 38 | 1000 | — | .896 | .243 | 252 |
| 39 | 1000 | — | .851 | .211 | 584 |
| 40 | 1000 | — | 1.051 | .348 | 373 |
| 41 | 1000 | — | 1.183 | .263 | 511 |
| 42 | 988 | 327 | 1.189 | .376 | 591 |
| 43 | 984 | 436 | 1.398 | .489 | 501 |
| 44 | 1000 | — | 1.037 | .352 | 301 |
| 45 | 976 | 419 | 1.368 | .565 | 429 |
| 46 | 1000 | — | 1.028 | .252 | 491 |
| 47 | 992 | 066 | 1.000 | .248 | 651 |
| 48 | 1000 | — | .898 | .270 | 319 |
| 49 | 996 | 132 | 1.181 | .308 | 501 |
| 50 | 992 | 066 | 1.078 | .317 | 576 |
| 51 | 972 | 439 | .841 | .350 | 181 |
| 52 | 980 | 527 | 1.195 | .644 | 357 |
| 53 | 1000 | — | .998 | .290 | 526 |
| 54 | 1000 | — | 1.000 | .231 | 530 |
| 55 | 980 | 450 | 1.127 | .273 | 672 |
| 56 | 956 | 428 | 1.379 | .667 | 428 |
| 57 | 996 | 018 | .953 | .304 | 509 |
| 58 | 1000 | — | .908 | .273 | 473 |
| 59 | 992 | 227 | 1.506 | (1.030) | (651) |
| 60 | 984 | 179 | 1.258 | .713 | 239 |
| 61 | 988 | 130 | 1.642 | .484 | 467 |
| 62 | 992 | 267 | 1.041 | .301 | 552 |
| 63 | 964 | 286 | 1.614 | .467 | 430 |
| 64 | 956 | 149 | 1.349 | .443 | 215 |
| 65 | 992 | 066 | 1.232 | .325 | 606 |
| 66 | 1000 | — | 1.036 | .309 | 171 |
| 67 | 988 | 130 | 1.147 | .361 | 523 |
| 68 | 1000 | — | .964 | .280 | 625 |
| 69 | 984 | 122 | 1.174 | .497 | 114 |
| 70 | 1000 | — | .822 | .190 | 271 |
| 71 | 996 | 131 | 1.268 | .341 | 457 |
| 72 | 968 | 194 | 1.176 | .419 | 233 |
| 73 | 1000 | — | 1.211 | .352 | 636 |
| 74 | 1000 | — | 1.265 | .314 | 645 |
| 75 | 948 | 277 | 1.632 | .508 | 469 |
| 76 | 996 | 018 | .996 | .247 | 538 |
| 77 | 1000 | — | .773 | .263 | 218 |
| 78 | 1000 | — | 1.105 | .234 | 674 |
| 79 | 944 | 428 | 1.680 | (.495) | (644) |
| 80 | 980 | 245 | 1.288 | (.771) | (647) |

表2 「図形・記号の探索」問題の正答数得点の分布

| 正答数得点 | 人数 | 百分率 |
|-------|-----|------|
| 74 | 3 | 1.2 |
| 75 | 1 | 0.4 |
| 76 | 3 | 1.2 |
| 77 | 11 | 4.4 |
| 78 | 23 | 9.2 |
| 79 | 56 | 22.4 |
| 80 | 153 | 61.2 |

3. 3 段階反応モデルによる分析

3. 3. 1. 方法

本研究では、受験者の解答と反応時間とを合わせて4つのカテゴリに割り当てる。すなわち、正答で反応時間の短いものから順にカテゴリ1, 2, 3, 4とし、さらに誤答の場合はカテゴリ4を割り当てる。カテゴリの境界には、全ての受験者の全正答項目の反応時間を大小順に並べ、第1四分位数、中央値、第3四分位数にあたる反応時間を求め、それらの値を丸めてきりのよい数値としたものを用いる。実際には、順に0.90秒、1.10秒、1.40秒をカテゴリ境界値とした。

そして、①各受験者毎に各項目の解答および反応時間から、各項目毎にカテゴリ値{1, 2, 3, 4}の何れかを割り当てる。②これらのデータに対して、多値型IRTモデルによる分析用プログラムのMULTILOG(PC版)を利用して、項目カテゴリ・パラメタの推定値を計算する。③得られた項目カテゴリ・パラメタ推定値にもとづいて、テスト情報量を計算する、という手順で分析を進める。

ただし、今回はパーソナル・コンピュータ用のMULTILOGを利用するために、分析対象とする項目の数が50項目以下に制限されてしまう。そこで、全80項目の中から、項目の材料のバランス等に配慮して、48項目を選択して分析を行なった。

なお、本研究で当初は、受験者の解答と反応時間とを合わせてカテゴリ化するのに、正答の場合に反応時間に応じて4段階のカテゴリ、誤答の場合は反応時間に関わらずひとつのカテゴリとして、全部で5カテゴリを設定したが、項目カテゴリパラメタの推定精度が十分ではなかったため、正答で反応時間の最も長いカテゴリと誤答とを合わせてひとつのカテゴリとした。

3. 3. 2. 結果

項目カテゴリ・パラメタ推定値は表3に示したとおりである。ただし、項目37, 43, 47の3項目は十分な推定精度が得られなかったため、削除して改めて45項目で推定値を求めた。なお、計算にあたってプログラム上の制限に合わせるためにカテゴリ値を逆順にした。すなわち、誤答または正答で反応時間が最も長いカテゴリを“1”，正答で反応時間が最も短いカテゴリを“4”，それらの中間のカテゴリも元の“3”を“2”に、元の“2”を“3”に割り当て直した。表3も割り当て直したカテゴリ値で示した。

また、これらの45項目から得られるテスト情報量は表4に示したとおりである。

3. 3. 3. 評価

パラメタaの推定値は、最小値が項目41の0.28、最大値が項目15の2.29で、45項目中32項目が1.00代であった。0.50未満の項目も3項目あるが、全体として十分な識別力を示している。

パラメタbの推定値は、項目カテゴリが全体として正の値を示す項目(項目1, 4, 5など)、全体として負の値を示す項目(項目11, 24, 35など)、カテゴリ間の差の大きい項目(項目12, 38, 41など)、カテゴリ間の差の小さい項目(項目7, 32, 40など)が適宜存在しており、全体として偏りが無い精度をもつ項目配置となっている。

なお、パラメタbの推定値が絶対値で4.00を超える値を持つ項目は、項目5, 10, 11, 37, 38, 39, 41, 45の8項目あるが、これらについては概してパラメタaの推定値が低く、項目5の1.11を除いては全て1.00未満、特に項目38, 39, 41は0.50未満である。0.50未満の値を示す項目については、項目の内容等についてさらに詳しく検討する必要がある。

テスト情報量については、尺度値が0.0で最大の情報量26.2(標準誤差0.20)を示し、-1.5から2.0の範囲で少なくとも情報量で18.6(標準誤差0.23)を示している。標準誤差が0.23というのは、信頼性係数に直すとおよそ0.95弱の値になり、測定精度として十分に満足できる値である。この値は項目数を実際よりも減らした上で計算されたものであるから、実際の「図形・記号の探索問題」ではより高い精度の測定が実現されているものと思われる。

表3 「図形・記号の探索」問題各項目の段階反応モデル・パラメタ推定値

| 項目番号 | 項目カテゴリパラメタ | | | |
|------|------------|-------|-------|-------|
| | A | B(1) | B(2) | B(3) |
| 1 | 0.92 | 0.30 | 1.93 | 3.67 |
| 3 | 1.10 | -1.17 | 0.17 | 1.74 |
| 4 | 0.81 | -2.07 | -0.75 | 0.73 |
| 6 | 1.66 | 0.43 | 1.96 | 3.17 |
| 7 | 1.11 | 0.85 | 2.85 | 5.39 |
| 9 | 2.03 | -0.27 | 0.78 | 2.13 |
| 10 | 2.07 | -1.08 | -0.19 | 0.54 |
| 13 | 1.56 | -1.47 | -0.12 | 1.18 |
| 14 | 1.80 | -0.37 | 0.59 | 1.56 |
| 16 | 1.69 | -0.79 | 0.63 | 1.89 |
| 17 | 0.61 | -5.64 | -3.52 | -1.67 |
| 19 | 0.60 | -1.61 | 1.56 | 4.56 |
| 21 | 1.30 | -1.59 | -0.15 | 1.32 |
| 23 | 1.64 | -0.69 | 0.67 | 1.90 |
| 24 | 2.29 | -0.31 | 0.86 | 2.24 |
| 26 | 1.42 | 0.32 | 1.66 | 2.82 |
| 27 | 1.44 | 0.39 | 1.96 | 3.43 |
| 29 | 1.23 | -1.28 | 0.29 | 1.90 |
| 30 | 1.20 | -2.11 | -0.37 | 1.02 |
| 33 | 1.32 | -1.36 | 0.52 | 2.11 |
| 34 | 1.00 | -3.03 | -1.10 | 0.54 |
| 36 | 1.90 | -0.81 | 0.40 | 1.66 |
| 39 | 1.91 | -1.24 | 0.10 | 1.12 |
| 41 | 1.59 | -3.14 | -1.85 | -0.69 |
| 44 | 1.06 | -1.84 | 0.27 | 2.32 |
| 46 | 1.15 | -0.54 | 0.92 | 2.73 |
| 49 | 0.83 | -2.99 | -1.16 | 0.84 |
| 50 | 1.01 | -2.70 | -0.84 | 0.90 |
| 53 | 1.86 | -1.98 | -0.70 | 0.34 |
| 54 | 1.22 | -1.39 | 0.07 | 1.73 |
| 56 | 1.66 | -1.73 | -0.33 | 0.71 |
| 57 | 1.62 | -1.95 | -0.99 | 0.24 |
| 59 | 1.26 | -2.70 | -1.01 | 0.34 |
| 61 | 1.70 | -0.52 | 0.51 | 2.07 |
| 63 | 1.38 | -2.13 | -1.19 | -0.16 |
| 64 | 1.65 | -1.17 | -0.05 | 1.20 |
| 66 | 0.94 | 0.82 | 2.76 | 5.05 |
| 67 | 0.29 | -2.13 | 2.73 | 7.95 |
| 69 | 0.33 | -6.22 | -2.78 | 1.54 |
| 70 | 1.80 | -1.17 | -0.10 | 0.81 |
| 73 | 0.28 | -4.35 | -0.85 | 3.04 |
| 74 | 1.78 | -1.10 | 0.18 | 1.55 |
| 76 | 1.96 | -0.88 | 0.57 | 1.87 |
| 77 | 1.24 | -2.53 | -1.04 | 0.44 |
| 79 | 0.86 | 0.74 | 2.94 | 5.44 |

4. 図形・記号の異同弁別問題

4. 1. 問題の概要

本問題は、画面中の左右2つの矩形領域内の各々に呈示された複数の図形・記号について、2つの領域の間でその内容が同じであるか、異なるかを判断して解答する。図形・記号の種類・個数のみならず配列順序も考慮に入れて判断することが求められる。各項目は全て課題としては易しいものであるが、「図形・記号の探索問題」と同様、受験者の“正誤反応”よりもむしろ“反応の速さ”を測定することを主たる目的としている。図形・記号等には具体的には、トレインマーク、時刻表記号、JRで用いられている警戒警報発令時に掲示される標識(以下、警戒標識図と略す)、鉄道車両の正面、の4種類を用いる。各項目を構成するこれらの図形・記号の中で、トレインマーク、時刻表記号、鉄道車両の正面、の3種類は「図形・記号の探索問題」で用いたものと同じである。なお、各領域内に含まれる図形・記号の数を3, 4, 5, 6, 8と変えることによって項目の困難度を予め変化させた。実際の各項目に含まれる、図形・記号の種類及び数は表5に示した通りである。各受験者毎に、正誤及び反応時間(0.01秒単位)が記録される。

4. 2. 項目分析の結果

本問題の各項目は、野口(1993a)で既に予備テストを実施しているが、受験者数が188名と少なかったことに加えて、さらに、1993年10月から1995年3月にかけて、「図形・記号の探索問題」と同じく240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した。

①項目通過率

各項目の通過率は表6に示した通りである。項目6の.70が最も小さく、項目36の.72、項目52の.76、項目35の.80、項目20の.81、項目44の.89と続く、通過率が1.00の項目は42のみであり、前回の結果では6項目あったのが減少している。その他の項目は0.90代の値を示している。全体として高い値を示すが、「図形・記号の探索」問題に比べると、相対的に低い値を示し、しかも項目間でのバラツキが大きくなっている。

表4 「図形・記号の探索問題」(短縮版)のテスト情報量

| 尺度値 | -2.0 | -1.5 | -1.0 | -0.5 | 0.0 | 0.5 | 1.0 | 1.5 | 2.0 |
|------|------|------|------|------|------|------|------|------|------|
| 情報量 | 14.5 | 18.9 | 22.8 | 25.3 | 26.2 | 26.0 | 24.3 | 21.6 | 18.6 |
| 標準誤差 | 0.26 | 0.23 | 0.21 | 0.20 | 0.20 | 0.20 | 0.20 | 0.22 | 0.23 |

表 5 各項目に含まれる図形・記号の種類及び個数

| | 図 形 ・ 記 号 の 種 類 | | | | | | | | | | | | | | | |
|---|-----------------|----|-------|----|----|----|-------|----|----|----|---------|----|----|----|----|----|
| | トレインマーク | | 時刻表記号 | | | | 警戒標識図 | | | | 鉄道車両の正面 | | | | | |
| 個 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| 3 | | | | | | | | | | | | | | | | |
| 4 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
| 5 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 |
| 6 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 |
| 8 | 57 | 58 | | | 59 | 60 | | | 61 | 62 | | | 63 | 64 | | |

②点双列相関係数

各項目の点双列相関係数は表6に示した通りである。ただし、通過率の値が1.00である項目42については点双列相関係数を計算する事ができない。その他の63項目については、項目46の-.04が最も小さく、項目29の-.01、項目41及び59の-.00が次に小さい値を示している。また、項目6の.49が最も大きい値を示し、項目36の.49、項目35の.48が次に大きい値を示している。全体として点双列相関係数の値は小さく、「図形・記号の探索」問題に比べて、全体的に小さくなっている。

③正答数得点の分布

正答数得点の分布は表7に示した通りである。41点、46点、49点、51点、52点が各1名、54点が6名、55点が7名、56点が11名、57点が22名、58点が27名、59点が39名、60点が44名、61点が66名、62点が79名、63点が72名、64点が58名となっている。全体として、正答数得点は比較的高い。正答数得点の分布の平均は60.8、標準偏差は2.8、歪度は-1.8、尖度は7.3で、左に裾を引く尖った分布を示している。ただ、その程度は「図形・記号の探索」問題に比べて緩やかである。

④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を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、190.4秒、28.0秒、0.404、0.378、である。わずかに右に裾を引くが、正規分布に比較的近い形状を示している。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、118.3、127.1、131.6、132.7、133.1、秒であり、後者が最長の方から、312.3、268.3、264.5、264.4、258.7秒であった。

⑤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6に示した通りである。

反応時間の平均は、項目32で1.89秒と最も短く、項目61で4.88秒と最も長く、反応時間の標準偏差は、項

目4で0.42秒と最も小さく、項目63で2.30秒と最も大きくなっている。標準偏差が、項目28、51、63で1.82、2.14、2.30と他の項目に比べて特に大きな値をとっているが、これは受験者の中に特に大きな反応時間（順に、34.6秒、40.4秒、43.1秒）を示した者がいた為である。総反応時間との相関係数は、項目9と21で0.63と最も大きく、項目52で0.17と最も小さくなっている。

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、14.0、2.6、1.9、1.8、1.7となり、第I固有値と第II固有値の比が5.4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。

⑥図形・記号の種類及び個数別の分析

図形・記号の種類及び個数別に平均通過率及び平均反応時間を求めた結果は表8に示した通りである。

平均通過率を図形・記号の種類毎に見ると、鉄道車両の正面では、個数3から順に.976、.975、.969、.967、.924、と個数が増えるに従って、通過率が減少しているが、トレインマークでは個数3から順に.975、.959、.973、.954、.955、時刻表記号では同じく、.970、.972、.985、.963、.973、警戒標識図では同じく、.816、.933、.866、.912、.943というように刺激個数との間に明確な関係は見られない。個数毎に図形・記号の種類間で比較すると、個数8で鉄道車両の正面が最も小さい値を示すのを除いて、その他の個数では全て警戒標識図が最も小さい値を示している。特に、個数3及び5でこの傾向が顕著である。ただ、警戒標識図の中で刺激個数との間に明確な関係が見られない為、この結果から、警戒標識図が特に他と比べて受験者にとって困難であると言いつけることはできない。

平均反応時間を図形・記号の種類毎に見ると、トレインマークでは個数3から順に2.208、2.438、2.835、3.289、3.666、時刻表記号では同じく順に、2.020、2.328、2.555、2.905、3.542、鉄道車両の正面では、

同じく順に2.264, 2.654, 3.071, 3.060, 3.798, と個数が増えるに従って、反応時間が増加しているが、警戒標識図では個数3から順に、2.800, 2.644, 3.765, 3.729, 3.624, というように刺激個数4と5の間に段差が見られる。この場合でも、全体としては個数の増加が反応時間の増加をもたらすと言える。個数毎に図形・記号の種類間で比較すると、個数3, 5, 6で警戒標識図が最も大きい値を示すのに対して、個

数4と8では鉄道車両の正面が最も大きい値を示している。すなわち、本問題に用いている図形・記号の中で最も単純なもので、かえって反応時間が長くなっている。この事は、トレインマーク、時刻表記号、鉄道車両の正面については、各刺激が漢字の様にひとまとまりの意味を持ち、コード化が容易であるのに対して、警戒標識図では各刺激に特別な意味が無くコード化が必ずしも容易ではなかった事によるものと思われる。

表6 「図形・記号の異同弁別」問題各項目の通過率・点双列相関係数及び反応時間の平均・標準偏差・総反応時間との相関係数

| 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 |
|------|------|---------|--------|----------|-----------|
| 1 | 998 | 099 | 2.258 | .566 | 465 |
| 2 | 952 | 192 | 2.157 | .540 | 423 |
| 3 | 995 | 043 | 2.097 | .639 | 367 |
| 4 | 945 | 149 | 1.942 | .423 | 430 |
| 5 | 929 | 232 | 2.737 | .858 | 488 |
| 6 | 702 | 494 | 2.862 | .965 | 412 |
| 7 | 993 | 023 | 2.433 | .629 | 534 |
| 8 | 959 | 246 | 2.094 | .585 | 395 |
| 9 | 995 | 224 | 2.892 | .709 | 630 |
| 10 | 989 | 122 | 2.672 | .692 | 554 |
| 11 | 920 | 349 | 2.102 | .771 | 320 |
| 12 | 933 | 234 | 2.084 | .711 | 365 |
| 13 | 993 | 151 | 2.565 | .708 | 519 |
| 14 | 993 | 062 | 2.479 | .787 | 548 |
| 15 | 954 | 267 | 1.965 | .639 | 297 |
| 16 | 947 | 262 | 2.303 | .616 | 443 |
| 17 | 956 | 083 | 3.461 | 1.186 | 504 |
| 18 | 975 | 195 | 3.048 | 1.188 | 519 |
| 19 | 989 | 206 | 2.441 | .940 | 415 |
| 20 | 812 | 467 | 3.141 | 1.178 | 453 |
| 21 | 989 | 122 | 3.081 | .839 | 631 |
| 22 | 986 | 026 | 2.570 | .767 | 574 |
| 23 | 954 | 240 | 2.470 | .728 | 422 |
| 24 | 970 | 293 | 2.494 | .734 | 381 |
| 25 | 995 | 200 | 3.581 | .886 | 583 |
| 26 | 989 | 060 | 3.173 | .908 | 466 |
| 27 | 940 | 283 | 2.128 | .595 | 388 |
| 28 | 966 | 227 | 2.456 | 1.823 | 231 |
| 29 | 984 | -010 | 2.883 | .721 | 586 |
| 30 | 989 | 153 | 3.017 | .887 | 478 |
| 31 | 984 | 276 | 2.422 | .709 | 409 |
| 32 | 984 | 217 | 1.898 | .891 | 225 |
| 33 | 961 | 102 | 4.308 | 1.448 | 595 |
| 34 | 979 | 109 | 3.906 | 1.281 | 600 |
| 35 | 807 | 481 | 3.079 | 1.198 | 392 |
| 36 | 718 | 487 | 3.767 | 1.328 | 432 |
| 37 | 984 | 081 | 3.652 | 1.171 | 555 |
| 38 | 991 | 112 | 3.486 | 1.067 | 521 |
| 39 | 945 | 185 | 2.273 | .798 | 299 |
| 40 | 954 | 295 | 2.872 | .958 | 379 |
| 41 | 998 | -004 | 4.287 | 1.061 | 623 |
| 42 | 1000 | — | 3.453 | .953 | 558 |
| 43 | 929 | 226 | 2.410 | .987 | 251 |
| 44 | 888 | 163 | 3.006 | 1.014 | 360 |

識別性検査 A-1001の「知覚の速さ・正確さ」領域の I R T 尺度化

| | | | | | |
|----|-----|------|-------|-------|-----|
| 45 | 979 | 086 | 3.675 | 1.041 | 580 |
| 46 | 998 | -038 | 3.036 | .889 | 512 |
| 47 | 956 | 087 | 1.964 | .711 | 256 |
| 48 | 920 | 325 | 2.944 | .932 | 312 |
| 49 | 972 | 111 | 4.365 | 1.443 | 535 |
| 50 | 970 | 144 | 3.873 | 1.318 | 592 |
| 51 | 947 | 226 | 3.092 | 2.143 | 273 |
| 52 | 759 | 461 | 3.586 | 1.622 | 173 |
| 53 | 986 | 061 | 3.726 | .994 | 543 |
| 54 | 984 | 055 | 3.665 | 1.295 | 474 |
| 55 | 943 | 261 | 2.369 | .759 | 363 |
| 56 | 956 | 159 | 2.480 | 1.136 | 285 |
| 57 | 993 | 003 | 3.820 | .842 | 607 |
| 58 | 917 | 288 | 3.512 | 1.278 | 489 |
| 59 | 998 | -004 | 4.146 | 1.004 | 559 |
| 60 | 947 | 262 | 2.938 | .833 | 423 |
| 61 | 961 | 161 | 4.882 | 1.549 | 595 |
| 62 | 924 | 227 | 2.366 | 1.416 | 280 |
| 63 | 979 | 069 | 4.809 | 2.301 | 371 |
| 64 | 869 | 204 | 2.786 | .946 | 325 |

表7 「図形・記号の異同弁別」
問題の正答数得点の分布

| 正答数得点 | 人数 | 百分率 |
|-------|----|------|
| 41 | 1 | 0.2 |
| 42 | 0 | 0 |
| 43 | 0 | 0 |
| 44 | 0 | 0 |
| 45 | 0 | 0 |
| 46 | 1 | 0.2 |
| 47 | 0 | 0 |
| 48 | 0 | 0 |
| 49 | 1 | 0.2 |
| 50 | 0 | 0 |
| 51 | 1 | 0.2 |
| 52 | 1 | 0.2 |
| 53 | 0 | 0 |
| 54 | 6 | 1.4 |
| 55 | 7 | 1.6 |
| 56 | 11 | 2.5 |
| 57 | 22 | 5.0 |
| 58 | 27 | 6.2 |
| 59 | 39 | 8.9 |
| 60 | 44 | 10.1 |
| 61 | 66 | 15.1 |
| 62 | 79 | 18.1 |
| 63 | 72 | 16.5 |
| 64 | 58 | 13.3 |

表8 図形・記号の種類及び個数別の平均通過率及び平均反応時間
上段：平均通過率 下段：平均反応時間(秒)

| | | 図形・記号の種類 | | | |
|---|---|----------|-------|-------|---------|
| | | トレインマーク | 時刻表記号 | 警戒標識図 | 鉄道車両の正面 |
| 3 | 個 | .775 | .970 | .816 | .976 |
| | | 2.208 | 2.020 | 2.800 | 2.264 |
| 4 | 個 | .959 | .972 | .933 | .975 |
| | | 2.438 | 2.328 | 2.644 | 2.654 |
| 5 | 個 | .973 | .985 | .866 | .969 |
| | | 2.835 | 2.555 | 3.765 | 3.071 |
| 6 | 数 | .954 | .963 | .912 | .967 |
| | | 3.289 | 2.905 | 3.729 | 3.060 |
| 8 | 数 | .955 | .973 | .943 | .924 |
| | | 3.666 | 3.542 | 3.624 | 3.798 |

4. 3. 段階反応モデルによる分析

4. 3. 1. 方法

本問題でも「図形・記号の探索問題」と同じ手続きにより、受験者の解答と反応時間とを合わせて4つのカテゴリに割り当てる。実際には、順に2.10秒、2.80秒、3.60秒をカテゴリ境界値とした。

そして、①各受験者毎に各項目の解答および反応時間から、各項目毎にカテゴリ値 {1, 2, 3, 4} の何れかを割り当てる。②これらのデータに対して、多値型 I R T モデルによる分析用プログラムの MULT ILOG (PC版) を利用して、項目カテゴリ・パラメタの推定値を計算する。③得られた項目カテゴリ・パラメタ推定値にもとづいて、テスト情報量を計算する、という手順

で分析を進める。

ただし、「図形・記号の探索問題」と同様に分析プログラムの制限から、全64項目の中から、項目の材料および刺激数のバランス等に配慮して、50項目を選択して分析を行なった。

なお、本問題についても当初は、全部で5カテゴリを設定したが、項目カテゴリパラメタの推定精度が十分ではなかったため、正答で反応時間の最も長いカテゴリと誤答とを合わせてひとつのカテゴリ、全部で4カテゴリとした。

4.3.2. 結果

項目カテゴリ・パラメタ推定値は表9に示したとおりである。ただし、項目16, 28, 29, 33, 41, 49の6項目は十分な推定精度が得られなかったため、削除して改めて44項目で推定値を求めた。なお、「図形・記号の探索問題」と同様に、計算にあたってプログラムの制限に合わせるためにカテゴリ値を逆順にした。

また、これらの44項目から得られるテスト情報量は表10に示したとおりである。

4.3.3. 評価

パラメタaの推定値は、最小値が項目21の0.31, 最大値が項目25の2.13で、44項目中19項目が1.00代であった。0.50未満の項目も13項目あり、「図形・記号の探索問題」と比較して全体としてやや低い識別力を示している。

パラメタbの推定値は、項目カテゴリが全体として正の値を示す項目(項目25, 26, 34など), 全体として負の値を示す項目(項目3, 4, 7など), カテゴリ間の差の大きい項目(項目17, 31, 44など), カテゴリ間の差の小さい項目(項目25, 26, 35など)が適宜存在しており、全体として偏りが無い精度をもつ項目配置となっている。

なお、パラメタbの推定値が絶対値で4.00を超える値を持つ項目は、項目1, 4, 7, 9, 10など18項目あるが、これらについては項目1, 32, 40, 41の4項目を除いてパラメタaの推定値が低く、全て1.00未満である。0.50未満の値を示す9項目(項目9, 10, 12, 21, 28, 30, 31, 38, 44)については、項目の内容等についてさらに詳しく検討する必要がある。

表9 「図形・記号の異同弁別」問題各項目の段階反応モデル・パラメタ推定値

| 項目番号 | 項目カテゴリパラメタ | | | |
|------|------------|-------|-------|-------|
| | A | B(1) | B(2) | B(3) |
| 1 | 1.00 | -4.16 | -1.93 | 0.19 |
| 2 | 0.77 | -3.81 | -2.63 | 0.17 |
| 3 | 1.05 | -3.89 | -2.64 | -0.31 |
| 4 | 0.52 | -5.42 | -4.55 | -1.57 |
| 5 | 1.08 | -1.37 | -0.09 | 1.53 |
| 7 | 1.28 | -2.62 | -1.18 | 0.71 |
| 8 | 0.60 | -4.89 | -3.13 | -0.13 |
| 9 | 1.40 | -1.72 | -0.21 | 2.32 |
| 11 | 0.37 | -5.55 | -4.07 | -1.05 |
| 12 | 0.41 | -5.47 | -3.62 | -0.26 |
| 13 | 1.34 | -2.23 | -0.80 | 1.06 |
| 15 | 0.43 | -6.60 | -4.17 | -1.41 |
| 17 | 0.79 | -3.42 | -2.04 | 0.78 |
| 19 | 1.32 | -0.24 | 1.01 | 2.60 |
| 20 | 0.81 | -2.87 | -1.37 | 0.54 |
| 21 | 1.71 | -1.10 | 0.26 | 2.08 |
| 23 | 0.49 | -4.38 | -1.93 | 1.83 |
| 24 | 0.56 | -4.30 | -1.60 | 1.62 |
| 25 | 1.22 | -0.47 | 1.79 | 4.95 |
| 27 | 0.62 | -4.02 | -3.09 | -0.10 |
| 31 | 0.31 | -6.75 | -3.34 | 0.96 |
| 32 | 1.48 | -1.70 | -0.08 | 2.21 |
| 34 | 0.74 | -3.53 | -1.70 | 1.18 |
| 35 | 0.42 | -6.57 | -4.24 | -1.58 |
| 36 | 2.13 | 0.63 | 1.47 | 2.29 |
| 37 | 1.81 | 0.27 | 1.33 | 1.89 |
| 39 | 1.42 | -0.23 | 1.34 | 3.41 |
| 40 | 0.38 | -5.38 | -3.31 | 0.47 |
| 43 | 0.53 | -2.28 | 0.33 | 2.58 |
| 44 | 0.29 | -5.25 | -3.17 | 1.05 |
| 45 | 0.35 | -2.90 | 1.18 | 5.50 |
| 47 | 1.21 | -0.20 | 1.63 | 4.06 |
| 50 | 0.39 | -6.56 | -4.22 | -1.76 |
| 51 | 1.65 | 0.90 | 1.68 | 2.50 |
| 52 | 1.58 | 0.40 | 1.41 | 1.99 |
| 53 | 0.75 | -1.07 | 0.46 | 2.03 |
| 55 | 1.26 | 0.02 | 1.80 | 3.72 |
| 56 | 0.39 | -5.39 | -2.72 | 1.25 |
| 57 | 0.43 | -3.58 | -2.15 | 0.31 |
| 59 | 1.15 | 0.27 | 2.57 | 5.16 |
| 60 | 1.42 | 0.80 | 2.50 | 4.14 |
| 61 | 0.68 | -2.01 | 0.66 | 3.61 |
| 63 | 1.99 | 1.28 | 1.75 | 2.29 |
| 64 | 0.34 | -3.13 | -0.45 | 5.76 |

表10 「図形・記号の異同弁別問題」(短縮版)のテスト情報量

| 尺度値 | -2.0 | -1.5 | -1.0 | -0.5 | 0.0 | 0.5 | 1.0 | 1.5 | 2.0 |
|------|------|------|------|------|------|------|------|------|------|
| 情報量 | 7.3 | 8.5 | 9.7 | 11.1 | 12.6 | 13.7 | 14.1 | 14.1 | 13.5 |
| 標準誤差 | 0.37 | 0.34 | 0.32 | 0.30 | 0.28 | 0.27 | 0.27 | 0.27 | 0.27 |

テスト情報量については、尺度値が1.0および1.5で最大の情報量14.1（標準誤差0.27）を示し、-1.5から2.0の範囲で少なくとも情報量で8.5（標準誤差0.34）を示している。全体として「図形・記号の探索問題」に比べてテスト情報量が低くなっている。これはパラメタaの推定値が全体として低い値を示すことと関係している。それでも標準誤差が0.34というのは、信頼性係数に直すとおよそ0.90弱の値になり、測定精度として満足できる値である。この値は項目数を実際よりも減らした上で計算されたものであるから、実際の「図形・記号の異同弁別問題」ではより高い精度の測定が実現されているものと思われる。

5. まとめ

本研究では、識別性検査A-1001の「図形・記号の探索問題」と「図形・記号の異同弁別問題」に対して、受験者の項目反応時間と解答の正誤に基づいてカテゴリ化した項目反応データに対して、多値型IRTモデルのひとつである「段階反応モデル」を適用してIRT尺度を構成することを試みた。その結果、ごく一部の項目を除いて極端な項目カテゴリ・パラメタ推定値が得られることもなく、テスト情報量についても実用上満足な精度を持つ尺度が構成されていることが示された。

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、本研究では利用するコンピュータ資源の関係から、実際の検査よりも項目数を減じて分析しているが、さらに全ての項目を含んだ分析を実施する必要がある。

次に、項目反応時間と項目反応に基づいてカテゴリ化する際に、本研究では全ての受験者の正答項目の反応時間の第一四分位数、中央値、第三四分位数を分割点としたが、これが最適なカテゴリ化であるかさらに検討しなければならない。

また、誤答の場合の取り扱いについても、例えば独立したカテゴリを設定する可能性についても検討する必要がある。ただし、この場合に検査の特徴から誤答が生起することが極めて少ないため、かなり多数の受験者からデータを得なければならず、現実的には困難が予想される。

さらに、多値型IRTの他のモデル、具体的には「名義反応モデル」を適用した場合とどちらがより有効であるか比較する必要もある。この場合も「名義反応モデル」におけるパラメタ数から考えて、かなり多数の受験者からデータを得なければならず、現実的な困難が予想される。

最後に、しかしながら最も重要な検討課題として、従来の作業量に基づく得点化法による結果と比較する必要がある。これによって、本研究で試みた多値型IRTモデルに基づく尺度の妥当性を検証することができるからである。

文 献

- Baker, F. B. 1992 Item Response Theory - Parameter estimation techniques -, Marcel Dekker Inc., New York, N.Y., U.S.A.
- 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 1995 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, 239.
- 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.
- Thissen, D. 1991 MULTILOG Version 6 User's Guide., Scientific Software Inc., Mooresville, IN., U.S.A.

(1995年9月13日 受稿)

ABSTRACT

IRT scaling for subtests intended to measuring the speed
and accuracy of perception 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 speed and accuracy of perception of subject. One of four categories was assigned to the answer based on subject's item response (correct or incorrect) and item response time in each item. Graded response model was adopted in IRT scaling and computer program MULILOG was used for item category parameter estimation. Estimated item category parameter values and test information curves showed satisfactory results for actual testing situations. This results indicate the possibility of using IRT model for scaling a speed test.

Key words : adult intelligence test, speed test, item response time, item response theory, graded response model