

# Partial Credit モデルにおける項目個別弁別力異同と 閾値母数推定値増大順序性違反についての探査一考察

## Item Discrimination and Disordering of the Threshold Estimates on the Partial Credit Model

井澤 廣行<sup>\*</sup>  
Hiroyuki Izawa

The purpose is to understand the following relationships specified by Andrich, regarding the threshold estimates on the Partial Credit model: 1) the farther apart the threshold estimates in their natural order, the weaker the item discrimination; 2) the closer the threshold estimates, the sharper the item discrimination; and 3) when they show a disorder, the item discriminates more sharply. It has been observed that Andrich's term of 'discrimination' has no direct relationship with point-biserial correlations on the raw data. Furthermore, it can be seen that the Partial Credit model is not in line with Rasch's theory.

**Key words:** Partial Credit model, threshold estimates, discrimination.

### 1. はじめに

井澤<sup>1)</sup>(2007)において、Andrich (1978<sup>2)</sup>, 2004<sup>3)</sup> and 2005<sup>4)</sup>)により洞察された Rasch 評定尺度モデル規定としての潜在閾値順序性がモデル測定出力上での閾値母数推定値増大順序性をカテゴリー順序尺度化妥当性の上で必然化しているとそのモデル解題で以って理解された。その一つの例示として、右脳と左脳の機能区別化に基づく筆者作成による性格二極分別性を一元的潜在特性とする 15 項目から成る四段階五範疇順序尺度質問紙調査データの Facets (Linacre<sup>5)</sup>, 1989-2001, Winsteps)上での閾値母数推定値出力が参照されて、 $\tau_1 = -0.68$ 、 $\tau_2 = -0.48$ 、 $\tau_3 = 0.37$ 、 $\tau_4 = 0.79$  が各数値順次増大性で以って出力されている。そのデータは、2004 年度と 2005 年度での流通科学大学外国語センターによる特別講義「言葉と文化」上での筆者担当授業受講者(調査研究承諾者)合計 709 名への質問紙調査結果であり、欠損値個数を 0 とする完全有効被験者数 692 名から成っている。提示された表 1(井澤<sup>1)</sup>, 2007, p. 16)により、Rating Scale モデル測定上では 15 項目全体としての四段階順序尺度化妥当性は否定されなかったけれども、Partial Credit モデル(Masters<sup>6)</sup>, 1982)測定上で全体 15 項目中の 9 項目が四段階順序尺度化妥当性に欠けるとの示唆が与えられた。表 1

---

<sup>\*</sup>流通科学大学サービス産業学部、〒651-2188 神戸市西区学園西町 3-1

に付された素点上での各項目点双列相関係数大小が以下に再度参照する Andrich<sup>7)</sup> (1988)と Linacre<sup>8)</sup> (1999)によって指摘された Partial Credit モデル測定上での閾値母数推定値順序性違反と項目弁別力との関係の上で整合性に欠けると意識された(井澤<sup>1)</sup>, 2007, p. 21)。

- 1) 『閾値母数推定値の数値増大性が保持されていることを以ってして、その値の差が大きくなればなる程、その項目弁別力は順じて小さくなる』 (Andrich<sup>7)</sup>, 1988, p. 370)。
- 2) 『閾値母数推定値間の差が小さくなればなる程、その項目弁別力は順じて大きくなる』 (Andrich<sup>7)</sup>, 1988, p. 370)。
- 3) 『閾値母数推定値順序性に違反があれば、その項目弁別力は更に大きくなる』 (Andrich<sup>7)</sup>, 1988, p. 370)。
- 4) 『閾値順序性への違反は、その項目が潜在特性に関する限られた範囲の上で過度に弁別的であることの示唆である』 (Linacre<sup>8)</sup>, 1999, p. 675)。

上記の各項目点双列相関係数大小の上での項目素点弁別力指標値と Partial Credit モデル上での閾値母数推定値順序性違反との不整合性への認知で以って、その不整合性を次節で明示して、第三節で、閾値母数推定値順序性違反と Rasch 測定出力上での項目弁別力との関係を査察する。第四節において、Andrich<sup>7)</sup> (1988)により供与された例示(pp. 370–371)の理解の上で、当該データ Partial Credit モデル測定出力上での閾値母数推定値分布と項目個別弁別力異同との関係を探查考察する。

## 2. 素点上での項目弁別力と Partial Credit モデル測定上での閾値母数推定値増大順序性違反との不整合性

表 1 当該データ Partial Credit モデル測定 Facets 出力上での閾値母数推定値とその増大順序性

	項目番号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	平均値
	Rating 位置母数 推定値	0.37	-0.13	0.19	0.53	-0.12	-0.05	0.58	-0.87	0.06	0.26	0.67	-0.64	-0.27	0.04	-0.61	0.00
	Partial 位置母数 推定値	0.41	-0.26	0.14	0.48	-0.20	-0.05	0.51	-0.67	0.08	0.21	0.62	-0.53	-0.24	0.00	-0.50	0.00
		関 値 母 数 推 定 値															
	閾値 順序性	○	○	○	X	○	○	○	X	X	X	X	X	X	X	X	*
項目 カテゴリ 段階	1	-1.05	-1.13	-0.93	-0.62	-0.86	-0.76	-0.48	-0.29	-0.41	-0.94	-0.74	-0.19	-0.31	-0.44	-0.11	-0.62
	2	-0.57	-0.28	-0.69	0.03	-0.49	-0.42	-0.21	0.08	-0.71	-0.88	0.02	-0.51	-0.59	-1.09	-0.67	-0.47
	3	0.30	0.65	0.66	-0.22	0.62	0.08	0.08	-0.23	0.31	0.95	-0.19	0.23	0.35	0.86	-0.12	0.29
	4	1.32	0.76	0.96	0.80	0.73	1.10	0.61	0.43	0.82	0.87	0.91	0.47	0.54	0.67	0.90	0.79
	点双列 相関係数	0.26	0.28	0.21	0.15	0.46	0.19	0.23	0.25	0.33	0.15	0.18	0.26	0.46	0.18	0.28	0.26

本節での考察内容理解促進のために、井澤<sup>1)</sup> (2007)提示表 1(p. 16)を上にも再掲する。簡易な項目弁別力指標値としての 15 項目に関する素点上での SPSS (Version15.0, SPSS Inc., 2006)出力に基づく点双列相関係数は被験者全員による各項目への素点応答列と各項目自体のその素点応答列を除く項目列合計全被験者得点列とのピアソン相関係数であり、SPSS 出力小数点四桁表示の値がその提示の上で小数点二桁表示に四捨五入されている。点双列相関係数は全項目に渡って全般的に低

いことが観察されて、15 項目に関するその平均値は 0.26 である。Partial Credit モデル規定としての閾値順次増大性に適合している○の記号が付された 6 項目、並びに、閾値母数推定値が少なくとも一つのカテゴリー順序段階においてモデル規定としての閾値順次増大性に反している×の記号が付された 9 項目それぞれの一群についての点双列相関係数平均値は 0.27 及び 0.25 と算出される。従って、素点に基づく点双列相関係数を指標値とする項目弁別力の上では上記二つの項目群の間に傾向的な差異は見出されない。

表 2 は、素点上での得点等間隔五段階被験者応答強度群別平均値に基づく 15 項目間での弁別力異同を提示するものである。各項目は五つのカテゴリー別とされており、左脳機能識別標示「強」から右脳機能識別標示「強」に渡る五範疇「強」・「弱」・「中間」・「弱」・「強」のそれぞれに 0・1・2・3・4 の配点とされている(井澤<sup>1)</sup>、2007, p. 25、参照)。各項目の五段階被験者応答強度群別平均値は、692 名素点得点昇順並びの上で、11 点を各段階での得点範囲として各段階に含まれる被験者群の各項目上での総点を各段階構成人数で除された平均値である。又、表 3 は、表 2 に与えられた各項目の五段階被験者応答強度群別平均値における最大値と最小値の差としての範囲を示す値を昇順で項目番号と併せて横に並べられたものである。

表 2 15 項目素点得点上での五段階被験者応答強度群別による平均値

被験者 群別	得点 範囲	構成 人数	ア ン ケ ー ト 項 目															平均値
			○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	
最低位	1-12	4	0.75	1.25	0.75	0.25	0.00	1.00	0.25	0.75	0.75	0.50	0.50	0.25	0.75	0.25	0.75	0.58
低位	13-24	54	1.00	1.67	1.28	1.06	1.20	1.44	0.87	2.07	1.19	1.50	0.78	2.02	1.09	1.48	2.07	1.38
中位	25-36	379	1.66	2.22	1.92	1.48	2.11	2.23	1.33	3.09	1.92	1.88	1.36	2.85	2.37	2.17	2.84	2.10
高位	37-48	244	2.25	2.91	2.45	2.04	3.15	2.73	2.10	3.50	2.84	2.21	1.81	3.35	3.25	2.50	3.28	2.69
最高位	49-60	11	3.18	3.73	2.91	2.45	3.82	3.09	3.45	3.55	3.91	3.27	2.82	3.91	3.91	3.45	3.91	3.42
最 小 値			0.75	1.25	0.75	0.25	0.00	1.00	0.25	0.75	0.75	0.50	0.50	0.25	0.75	0.25	0.75	0.58
平 均 値			1.77	2.36	1.86	1.46	2.06	2.10	1.60	2.59	2.12	1.87	1.45	2.48	2.27	1.97	2.57	2.04
最 大 値			3.18	3.73	2.91	2.45	3.82	3.09	3.45	3.55	3.91	3.27	2.82	3.91	3.91	3.45	3.91	3.42
範 囲			2.43	2.48	2.16	2.20	3.82	2.09	3.20	2.80	3.16	2.77	2.32	3.66	3.16	3.20	3.16	2.84

表 3 15 項目素点得点上での五段階被験者応答強度群別平均値における  
最大値と最小値の差としての範囲を示す値の昇順横並び

項目番号	6	3	4	11	1	2	10	8	9	13	15	7	14	12	5
範 囲	2.09	2.16	2.20	2.32	2.43	2.48	2.77	2.80	3.16	3.16	3.16	3.20	3.20	3.66	3.82

前掲表 1 に観察された 15 項目の閾値母数推定値増大順序性において、少なくとも一つのカテゴリー順序段階の上で Rasch 評定尺度モデル規定としての閾値順次増大性に反している 9 項目は項目番号 4、8、9、10、11、12、13、14、15 である。然しながら、表 2 において各項目への記号○の付加により明示されている様に、素点得点上での全項目の五段階被験者応答強度群別平均値は完全な昇順並びとなっており、素点上では弁別力増大順序性に違反している項目は存在しない。更に、表 3 における各項目の素点五段階被験者応答強度群別平均値における最大値と最小値の差としての範囲値の大小順位上で、前掲表 1 での Partial Credit モデル測定上での閾値母数推定値増大順序性への準拠と違反に基づく項目群折半との関連性は明確には観察されない。ちなみに、図

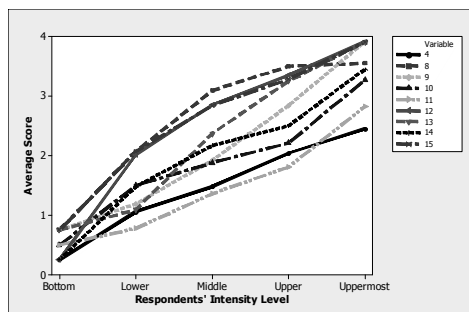


図1 モデル規定としての閾値順次増大性に反する  
9項目の素点上での被験者群応答強度弁別性

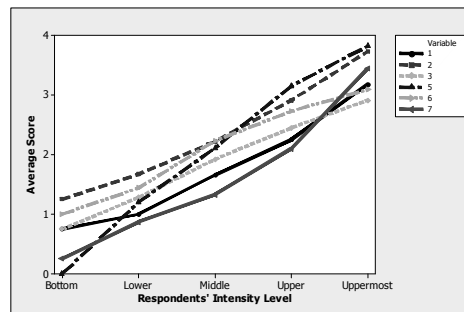


図2 モデル規定としての閾値順次増大性に準ずる  
6項目の素点上での被験者群応答強度弁別性

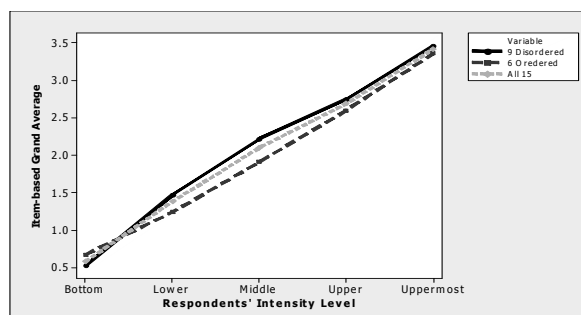


図3 モデル規定閾値母数推定値順序性違反9項目、準拠6項目、及び、全体15項目の  
素点上でのそれぞれ全体平均一項目あたりでの被験者群応答強度弁別性傾向線

1に与えたモデル規定としての閾値順次増大性に反している9項目の素点上での弁別力が図2でのモデル規定に準ずる6項目の弁別力に全体的に大きく優位であるとも目視され難い。

前掲表2での最右列に提示された平均値は各項目上での各段階被験者応答強度群別平均値の総点を項目数15で除された一項目あたりでの算出値であり、五段階を通して15項目全体平均としての素点上での被験者群応答強度弁別性を示すものである。その図示が図3での三本の傾向線において中位に位置しているものであり、その上位傾向線と下位傾向線はそれぞれ前掲表1でのPartial Creditモデル規定閾値母数推定値順序性への違反9項目群と準拠6項目群に関して同様に算出された一項目あたり平均での被験者群応答強度弁別性を示している。三線はほぼ平行であり、いずれも傾斜角45度に近いほぼ直線的な傾向線として描かれており、素点上での五段階被験者応答強度弁別性は、モデル規定閾値母数推定値順序性違反9項目群、準拠6項目群、及び、全体15項目それぞれの上で同様に弱くはないとの示唆である。

念を期することを目的として、前掲表2と3、並びに、図3のそれぞれに対応する別の新たな試行結果を示すものが表4と5、並びに、図4である。その提示は、692名素点得点昇順並びの上で各段階での異なる得点範囲内にほぼ均等な人数が含まれる五段階被験者応答強度群別に基いている。

表 4 別試行上での素点得点 15 項目五段階被験者応答強度群別平均値

被験者 群別	得点 範囲	構成 人数	アンケート項目															平均値
			○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	
最低位	7-28	136	1.16	1.74	1.44	1.15	1.40	1.68	0.95	2.39	1.43	1.57	0.90	2.39	1.45	1.79	2.29	1.58
低位	29-32	138	1.67	2.23	1.91	1.30	2.07	2.21	1.23	3.20	1.86	1.91	1.20	2.78	2.33	2.04	2.67	2.04
中位	33-35	123	1.80	2.34	2.06	1.72	2.44	2.37	1.57	3.15	2.08	1.92	1.59	2.86	2.63	2.33	3.07	2.26
高位	36-39	155	2.08	2.74	2.25	1.73	2.86	2.68	1.65	3.43	2.59	1.99	1.63	3.30	3.10	2.25	3.15	2.50
最高位	40-56	140	2.44	3.09	2.61	2.34	3.26	2.75	2.54	3.54	3.01	2.48	2.11	3.43	3.38	2.81	3.47	2.88
最小値			1.16	1.74	1.44	1.15	1.40	1.68	0.95	2.39	1.43	1.57	0.90	2.39	1.45	1.79	2.29	1.58
平均値			1.83	2.43	2.05	1.65	2.41	2.34	1.59	3.14	2.19	1.97	1.49	2.95	2.58	2.24	2.93	2.25
最大値			2.44	3.09	2.61	2.34	3.26	2.75	2.54	3.54	3.01	2.48	2.11	3.43	3.38	2.81	3.47	2.88
範囲			1.28	1.35	1.17	1.19	1.86	1.07	1.59	1.15	1.58	0.91	1.21	1.04	1.93	1.02	1.18	1.30

表 5 表 4 での素点得点 15 項目五段階被験者応答強度群別平均値における最大値と最小値の差の昇順横並び

項目番号	10	14	12	6	8	3	15	4	11	1	2	9	7	5	13
範囲	0.91	1.02	1.04	1.07	1.15	1.17	1.18	1.19	1.21	1.28	1.35	1.58	1.59	1.86	1.93

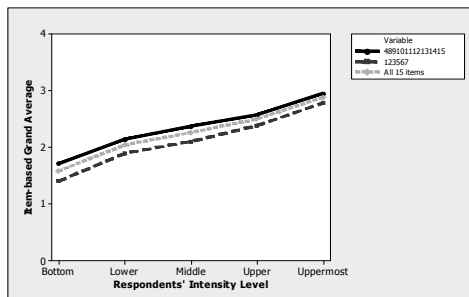


図 4 モデル規定閾値母数推定値順序性違反 9 項目、準拠 6 項目、及び、全体 15 項目の素点上での表 4 に関するそれぞれ全体平均一項目あたりでの被験者群応答強度弁別性傾向線

表 4 と 5 が前掲表 2 と 3 のそれぞれに大きく異なる処は、各項目の素点上での弁別力が格段に小さくなっていることのみであり、併せて、表 4 には項目 8 と 14 にそれぞれ一つの順位段階で僅かばかりの値減少が観察される。然しながら、素点上での項目群全体に渡って前掲図 3 との比較の上でのその弁別力減少を示す三本の傾斜角が緩やかな被験者群応答強度弁別性傾向線は僅少な差を以てしての上下平行移動として図 4 に示されている。従って、表 4 と 5、並びに、図 4 においても、前掲表 2 と 3、並びに、図 3 においてと同じく、Partial

Credit モデル規定閾値母数推定値順序性違反 9 項目群と準拠 6 項目群との間にそれぞれ項目群一体としての素点上での被験者群応答強度弁別性に大きな差異は観察されない。これは、前掲表 1 に与えた点双列相関係数による項目素点弁別力指標値上で観察された結果と同様である。又、表 1 での 15 項目点双列相関係数の平均値が 0.26 となっており、その低い値が図 3 よりもむしろ図 4 での全体 15 項目についての全体的に傾斜角の緩やかな直線的傾向線で示された一項目平均としての低い被験者群応答強度弁別性程度に相応していることが目視される。いずれにしても、当該データに関しては、Partial Credit モデル測定上での閾値母数推定値増大順序性への違反 9 項目群と準拠 6 項目群の間には素点の上では全体的な項目弁別力についての識別可能な大きい傾向的差異は存在しない。従って、この一事例の上では、Andrich<sup>7)</sup> (1988, p. 370) と Linacre<sup>8)</sup> (1999, p. 675) により指摘された Partial Credit モデル規定閾値母数推定値増大順序性への準拠と違反に基づく項目弁別力異同と素点で以て定義される項目弁別力異同の間に両者の弁別力概念として通ずる一義的な整合性が認められない。

### 3. 閾値母数推定値順序性違反と Rasch 測定出力上での項目弁別力との関係査察

前節で筆者に意識されたことは、当該データ Partial Credit モデル測定出力上での閾値母数推定値増大順序性への違反か準拠に基づくデータ折半項目群の間に素点に依拠する項目弁別力指標値において異同の程度が明確に観察されないということである。従って、Partial Credit モデル閾値母数推定値順序性違反と Rasch 測定出力上での項目弁別力指標との関係を観ることになるが、筆者に唯一つ既知の方法による結果を以下に示す。それは、当該データ Partial Credit モデル測定出力上での閾値母数推定値順序性違反9項目群とその準拠6項目群への折半に基づく二分各下位データの Rating Scale モデル測定出力上での同一被験者群応答強度位置母数推定値の分布幅異同程度を観察するものである。表6がその出力結果についての基本統計量であり、図5がその分布比較箱ひげ図である。なお、6項目群データには0点と満点の被験者がそれぞれ一名ずつ存在することからその有効被験者数は690名となっており、又、全体15項目データによる出力結果も併せて参照までに図表内の提示に含めている。

表6 閾値母数推定値順序性違反9項目群データ、閾値母数推定値順序性準拠6項目群データ、並びに、全体15項目データの Rating Scale モデル測定出力上での被験者群位置母数推定値分布基本統計量

	閾値母数推定値順序性違反 9項目群データによる 被験者群位置母数推定値	閾値母数推定値順序性準拠 6項目群データによる 被験者群位置母数推定値	全体15項目データによる 被験者群位置母数推定値
有効被験者数	692	690	692
最小値	-2.02	-2.87	-1.70
平均値	0.31	0.11	0.22
最大値	2.46	2.91	2.26
最大値と最小値の差	4.48	5.78	3.96
標準偏差	0.50	0.62	0.44

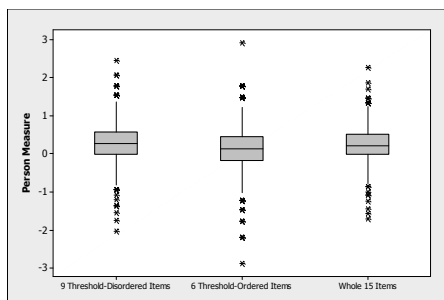


図5 表6での各データ Rating Scale モデル測定出力上での被験者群位置母数推定値分布箱ひげ図

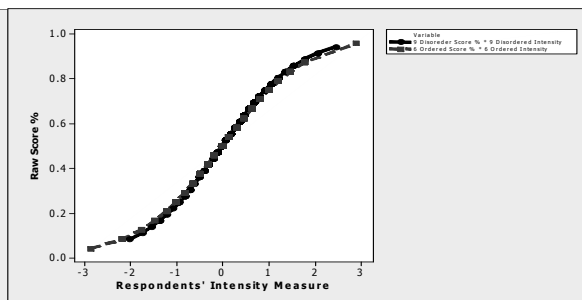


図6 違反9項目群データと準拠6項目群データの Rating Scale 測定出力上でのデータ特性曲線

表6と図5から、当該データ Partial Credit モデル測定出力上での閾値母数推定値順序性違反9項目群データがその準拠6項目群データに被験者群位置母数推定値分布のばらつきにおいて下回っていることが観察される。両データに関する分散同等性検定を参照すれば、標本がかなり大きいこともあり、F-test と Levene's Test のいずれによっても有意である( $p < 0.001$ )。これは、各折半デー

タ Rating Scale モデル測定上での被験者位置母数推定値のその同一被験者各自素点得点率に対する各折半データ特性曲線を想定すれば、閾値母数推定値順序性違反 9 項目群データがその準拠 6 項目群データよりも被験者群位置母数推定値分布範囲において僅かにしても狭小であるとの顕示である。それは、図 6 に示されている様に、各データ特性曲線交点での違反 9 項目群データ特性曲線傾斜角度が準拠 6 項目群データのそれ以上に僅かにしても鋭いということであり、被験者群応答強度弁別性において違反 9 項目群データが準拠 6 項目群データに僅かにしても優ることを意味している。従って、その一項目あたりでの平均弁別力を想定すれば、『閾値母数推定値順序性に違反があれば、その項目弁別力は更に大きくなる』(Andrich<sup>7)</sup>, 1988, p. 370)ということが当該データ Rasch モデル測定上で示唆されている。

表 7 閾値母数推定値順序性違反 9 項目群データ、閾値母数推定値順序性準拠 6 項目群データ、並びに、全体 15 項目データそれぞれの Rating Scale モデル測定上での閾値母数推定値分布

閾 値	閾値母数推定値順序性違反 9項目群データRating Scale 測定上での閾値母数推定値	閾値母数推定値順序性準拠 6項目群データRating Scale 測定上での閾値母数推定値	全体15項目データ Rating Scale測定上での 閾値母数推定値
タウ 1	-0.63	-1.00	-0.68
タウ 2	-0.50	-0.52	-0.48
タウ 3	0.34	0.46	0.37
タウ 4	0.78	1.06	0.79
タウ4とタウ1の差	1.41	2.06	1.47
標準偏差	0.68	0.93	0.70

表 7 は、上記各折半データと全体 15 項目データの Rating Scale モデル測定上でのカテゴリー順序四段階閾値母数推定値分布を比較するものである。当該全体 15 項目データの Partial Credit モデル測定出力上での閾値母数推定値順序性に違反している 9 項目群データの Rating Scale モデル測定による閾値母数推定値群が比較されている二つのデータの対応閾値母数推定値群よりもその分布幅において僅かにしても小さいと示されている。前掲表 6 からは、違反 9 項目群データの準拠 6 項目群データに僅かにしても上回る被験者群応答強度位置母数推定値弁別性が示唆されている。要するに、Rating Scale モデル測定上での閾値母数推定値の分布幅が小さくなるに従って、項目群一体としての弁別力総体は大きくなる傾向一片が当該データについて観察される。これは、項目群一体としての弁別力総体という観点の上で、第一節で参照された『閾値母数推定値間の差が小さくなればなる程、その項目弁別力は順じて大きくなる』との Andrich<sup>7)</sup> (1988, p. 370)による指摘に合致するものである。従って、前節での素点上での項目弁別力と Partial Credit モデル測定上での閾値母数推定値増大順序性違反との当該データにおける不整合性が本節での Rating Scale モデル測定上での被験者群位置母数推定値分布と閾値母数推定値分布への関連考察により理解一端の上で解消される。付記として、素点上での項目弁別力指標である点双列相関係数は少なくとも Rasch 測定との関連性の上では安易に参照されるべきではないと意識される。

なお、本節での考察課題から離れてはいるが、興味ある事項を以下に観察しておく。潜在閾値増大順序性を規定する Rasch 評定尺度モデルの観点からは、当該データ 15 項目の中で Partial Credit

モデル測定出力上での閾値母数推定値順序性に準ずる 6 項目がその規定違反の 9 項目以上に項目群総体として Rasch モデルへの適合性が高いと推量される。各項目が五範疇から構成される順序尺度データの Rating Scale モデル測定上での閾値母数推定値に関して、隣接値との差が少なくとも 1.00 であることがモデル理論として望まれると Linacre<sup>9)</sup> (2004, p. 274)により指摘されている。従って、表 7 における閾値母数推定値順序性準拠 6 項目群データがその違反 9 項目群データと全体 15 項目データに優って閾値母数推定値分布幅が大きくなっていることは道理に適う。ちなみに、表 8 と 9 がその閾値母数推定値順序性の観点上で、その準拠 6 項目群データがその違反 9 項目群データに Rasch モデル適合性において優る証左となる。各データの Partial Credit モデル測定出力上での各項目に付された閾値母数推定値とその順序性確認を与えており、前節冒頭掲示表 1 に対応する提示要領である。

表 8 準拠 6 項目群データ Partial Credit 測定上での閾値母数推定値とその増大順序性、並びに、点双列相関係数

	項 目 番 号	1	2	3	5	6	7	平均 値
	Rating 位置母数推定値	0.28	-0.32	0.06	-0.30	-0.23	0.52	0.00
	Partial 位置母数推定値	0.35	-0.40	0.06	-0.33	-0.17	0.48	0.00
	閾 値 母 数 推 定 値							
	閾 値 順 序 性	○	○	○	○	○	○	*
項目 カテゴリー 段階	1	-1.25	-1.36	-1.13	-1.08	-0.96	-0.64	-1.07
	2	-0.62	-0.31	-0.73	-0.52	-0.46	-0.25	-0.48
	3	0.36	0.72	0.72	0.69	0.14	0.12	0.46
	4	1.51	0.94	1.14	0.91	1.27	0.77	1.09
	点双列相関係数	0.22	0.21	0.18	0.35	0.14	0.20	0.22

表 9 違反 9 項目群データ Partial Credit 測定上での閾値母数推定値とその増大順序性、並びに、点双列相関係数

	項 目 番 号	4	8	9	10	11	12	13	14	15	平均 値
	Rating 位置母数推定値	0.63	-0.78	0.16	0.35	0.77	-0.55	-0.18	0.13	-0.52	0.00
	Partial 位置母数推定値	0.57	-0.63	0.15	0.28	0.71	-0.49	-0.19	0.07	-0.46	0.00
	閾 値 母 数 推 定 値										
	閾 値 順 序 性	X	X	X	X	X	X	X	X	X	*
項目 カテゴリー 段階	1	-0.67	-0.35	-0.46	-1.00	-0.79	-0.24	-0.36	-0.50	-0.17	-0.50
	2	0.01	0.07	-0.73	-0.90	-0.01	-0.52	-0.60	-1.10	-0.68	-0.50
	3	-0.21	-0.21	0.32	0.97	-0.18	0.25	0.37	0.88	-0.11	0.23
	4	0.86	0.49	0.87	0.93	0.98	0.52	0.59	0.72	0.96	0.77
	点双列相関係数	0.12	0.22	0.21	0.11	0.13	0.26	0.38	0.22	0.29	0.22

表 8 と 9 における Rating Scale と Partial Credit のモデル測定出力間での項目群位置母数推定値の相似性は井澤<sup>1)</sup>(2007, p. 17)における観察と同様の体である。Partial Credit モデル測定上での閾値母数推定値順序性への準拠と違反が全体 15 項目データに関する前節掲示表 1 での峻別と完全に同



一であることに加えて、表 1 との比較の上で各項目位置母数推定値のみならず各項目閾値母数推定値の近似性が明確な傾向として示されている。これは、Henning<sup>10)</sup> (1988)によって検証され、井澤(2003<sup>11)</sup>, pp. 16–17; 2005<sup>12)</sup>, p. 8)によっても例証確認された 0・1 データ Rasch 項目分析モデル測定上での部分項目群データと全体項目群データ間におけるその同一部分項目群の困難度推定値同一性に準ずるものと理解される。但し、Rasch 項目分析モデル測定に較べて Rasch 評定尺度モデル測定の上では変数増大による起因結果として、その同一性についての誤差変動が存在すると推察される。なお、前節での分析結果に準じて、表 8 と 9 における一項目あたりの点双列相関係数平均値が同値 0.22 であることに注目されるが、本節で明示された様に 01 データ上での項目弁別力と Rasch 測定上での項目弁別力との間に存在するその実体内容相違への留意が必要とされる。

Rasch 評定尺度モデル規定としての潜在閾値増大順序性(Andrich, 1978<sup>2)</sup>, 2004<sup>3)</sup> and 2005<sup>4)</sup>)に基づく観点の上では、表 8 での閾値母数推定値順序性準拠 6 項目群データが表 9 でのその違反 9 項目群データに Rasch モデル適合性において明確に優っている。それは、Rasch 評定尺度モデルカテゴリー一生起確率構造のガットマン完全順序尺度への依拠が閾値母数推定値増大順序性を必然としている(Andrich<sup>4)</sup>, 2005, p. 49)ことによる。然しながら、順序尺度データ潜在特性一次元性は閾値母数推定値増大順序性のみによっては説明され得ない(Linacre<sup>13)</sup>, 2001, p. 794, 参照)。井澤<sup>1)</sup>(2007)提示表 4(p. 19)での当該 15 項目全体データに関する被験者群 690 名位置母数推定値不変性成立程度により明示された項目群内容一次元性充足度の低さにその示唆もある。又、上掲表 8 での閾値母数推定値順序性準拠 6 項目群データの素点に基づく簡易な項目群内容一次元性指標値としての SPSS 出力による Cronback の内的一貫性係数 0.4392 という低い値の上でも、その 6 項目群データ潜在特性に関する一次元性充足度の低さが窺われる。以上が本節での付加情報である。

#### 4. Andrich (1988)例示の査察と Partial Credit モデルにおける項目個別弁別力異同

前節において当該データ Partial Credit モデル測定出力上での閾値母数推定値順序性違反 9 項目群とその準拠 6 項目群への折半に基づく二分各下位データ Rating Scale モデル測定出力上での被験者群応答強度位置母数推定値分布と閾値母数推定値分布を比較参照した。それを以って、本稿考察課題としての Andrich<sup>7)</sup> (1988, p. 370)により指摘された Partial Credit モデル測定出力上での閾値母数推定値分布と項目弁別力の相対関係が一例上で大略理解される。然しながら、それは、上記二分各データ Rating Scale モデル測定出力上での項目群一体としての弁別力総体という観点に基づく理解一片であり、Partial Credit モデル測定出力上での閾値母数推定値分布と項目個別弁別力異同との関係への査察に基づく納得性の高い理解ではない。その理解向上に具する説明が Andrich<sup>7)</sup> (1988)により供与されており(pp. 370–371)、本節でそれを例示詳察した上で、当該データ Partial Credit モデル測定出力上での閾値母数推定値分布と項目個別弁別力異同との関係を査察する。

『Partial Credit モデルにおける項目間弁別力異同は各項目位置母数  $\delta_i$  を限定する各閾値母数  $\tau_{ij}$  の項目間での相違から生ずる』(Andrich<sup>7)</sup>, 1988, p. 371)と明記され、その項目間弁別力異同との解釈類似性が一つの 0・1 データにおける二つの下位テストの Rasch 項目分析モデル測定上で説明されている。0・1 データに適用される Rasch 項目分析モデルは各項目弁別力の等位性を規定している。Rasch 項目分析モデル測定上での二つの下位テストそれぞれについての項目群全体としての弁別力総体における異同が Partial Credit モデル測定上での項目間弁別力異同に相当類似するとされる。更に、各下位テスト項目群における Rasch 項目分析モデル測定上での各項目困難度の差異が Partial Credit モデル測定上での各閾値母数の差異に比較該当するとされる。Andrich<sup>7)</sup> (1988)によるその要点は、大きい分布幅を持つ項目困難度推定値群が所与とされた全体 0・1 データの Rasch 項目分析モデル測定上での隣接困難度推定値間の差が小さい推定値中間範囲集中項目群のみから成る部分テストは隣接困難度推定値間の差が大きい推定値全体範囲網羅項目群のみから成る部分テストに較べて受験者群能力弁別性の程度が大きい(pp. 370–371)ということである。従って、Partial

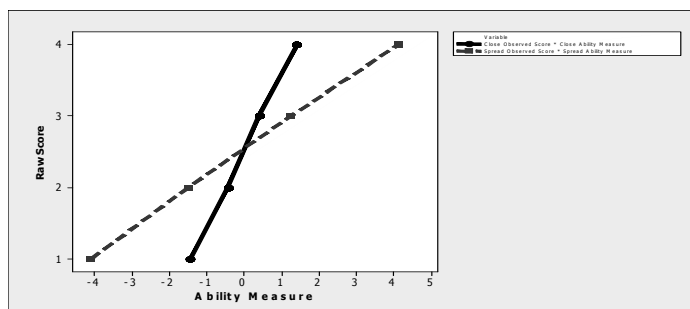


図7 二つの下位テスト間でのテスト特性曲線比較

Credit モデル測定との比較類似性で以って、『閾値母数推定値間の差が小さくなればなる程、その項目弁別力は順じて大きくなる』(Andrich<sup>7)</sup>, 1988, p. 370)と図示を含めて帰結されている。

Rasch 項目分析モデル測定上での全体 0・1 データにおけ

る隣接困難度推定値間の差が小さい推定値中間範囲に集中する 5 項目のみから成る部分テストと隣接困難度推定値間の差が大きい推定値全体範囲に渡る 5 項目のみから成る部分テストについての Andrich<sup>7)</sup> (1988)により与えられたテスト特性曲線比較図(p. 371)に対応するものが上図7である。前者困難度推定値中間範囲集中 5 項目群テストの後者困難度推定値全体範囲網羅 5 項目群テストにかなり上回る鋭い特性曲線傾斜角が示されており、Andrich<sup>7)</sup> (1988)の指摘通りである。以下に、図7作成法に併せて、Rasch 項目分析モデル測定出力上での二つの下位テストに付された数量指標を参照比較する。

利用 0・1 データは井澤<sup>14)</sup>(2006)においてカイ二乗検定による母数推定値同源性検証で以って Rasch モデル適合度が高くはないと判定された英語構文知識に関する 100 項目への流通科学大学生 305 名によるテスト応答である。その Rasch 項目分析モデル測定上での項目困難度推定値として -0.40、-0.20、0.01、0.17、0.40 が付された 5 項目と -3.60、-1.79、0.01、1.68、3.36 が付された前者項目群と重複しない 5 項目の項目群二組を二つの下位テストとしている。なお、満点得点者

と 0 点得点者を除く有効受験者数は前後者各部分テストにおいて 211 名と 288 名である。Andrich<sup>7)</sup> (1988)により図示されている(p. 371)様に、各部分テストにおける受験者群素点得点と各 0・1 データ Rasch 項目分析モデル測定上での受験者群能力母数推定値を縦横軸にそれぞれ配置したものが上図 7 である。

表 10 困難度推定値中間範囲集中 5 項のみから成る部分テストの Rasch 測定指標値と点双列相関係数

項目番号	元データにおける 困難度推定値	近接5項目データ 困難度推定値	近接5項目データ 点双列相関係数	近接5項目データ 受験者群 点双列相関係数 平均値	近接5項目データ 5項目による 受験者群分離 信頼性係数	近接5項目データ 211名による 項目群分離 信頼性係数
50	-0.40	-0.51	-0.02	0.14	0.00	0.82
14	-0.20	-0.26	-0.14			
73	0.01	0.01	-0.17			
70	0.17	0.22	0.03			
76	0.40	0.54	-0.11			
平均値	0.00	0.00	-0.08			

表 11 困難度推定値全体範囲網羅 5 項のみから成る部分テストの Rasch 測定指標値と点双列相関係数

項目番号	元データにおける 困難度推定値	広範5項目データ 困難度推定値	広範5項目データ 点双列相関係数	広範5項目データ 受験者群 点双列相関係数 平均値	広範5項目データ 5項目による 受験者群分離 信頼性係数	広範5項目データ 288名による 項目群分離 信頼性係数
78	-3.60	-5.19	0.11	0.76	0.44	0.99
36	-1.79	-2.90	0.05			
16	0.01	-0.15	0.20			
91	1.68	2.71	0.06			
38	3.36	5.53	-0.03			
平均値	-0.07	0.00	0.08			

表 10 と 11 において観察される様に、元データから選出された困難度推定値中間範囲集中 5 項目群と困難度推定値全体範囲網羅 5 項目群のそれぞれから成る各部分テスト 0・1 データの Rasch 測定上で付された 5 項目群困難度推定値分布範囲もその選出基準に呼応してそれぞれ中間範囲集中と全体範囲網羅となっている。従って、前後者各部分テスト 5 項目群による受験者群能力分離程度の違いが受験者群分離信頼性係数として付されたそれぞれ 0.00 と 0.44 の値に示されている。それは、前後者各部分テスト 5 項目群によって Rasch 項目分析モデル測定の上で推定された受験者群能力分布が前者においては狭い範囲内にある-1.42、-0.42、0.42、1.42、そして、後者においては広い範囲内にある-4.08、-1.48、1.27、4.14 から成るものに準じている。前後者各部分テスト 5 項目群への応答上におけるその受験者能力推定値分布それぞれが、受験者素点得点上での 1、2、3、4 に対応している故に、上図 7 において前者困難度推定値中間範囲集中 5 項目群テストの後者困難度推定値全体範囲網羅 5 項目群テストにかなり上回る鋭い特性曲線傾斜角が示されている訳である。

前後者各部分テストの素点上での受験者群点双列相関係数分布範囲はそれぞれ-0.86～0.84、-0.14～0.93 と与えられて、表 10 と 11 に示されている様にそれぞれにおける平均値は 0.14、0.76 となる。その受験者群点双列相関係数平均値それぞれと較べて、項目群分離信頼性係数それぞれの値が 0.82、0.99 と出力されており、受験者数が多いこともあり、後者の大変に高い値について

は容易に納得される。前者部分テストに関する受験者群点双列相関係数平均値 0.14 と項目群分離信頼性係数 0.82 との大きな数値乖離は、その受験者群点双列相関係数分布範囲が $-0.86$  から  $0.84$  に渡っており、負の相関係数が高い受験者群がかなり含まれているからである。なお、上記の受験者群と項目群に関する分離信頼性係数は以下の様に算出されており、その算出への基となる概念は、古典的テスト理論においてと同様に、[信頼性係数 = 真の得点分散 / 観測得点分散]との定義に基づく(Linacre<sup>5)</sup>, 1989-2001, p. 77)ものである。又、受験者群能力と項目群困難度のいずれについてもそれぞれ項目数と受験者数の大きさに依存して推定値標準誤差が押し並べて小さな値になるにつれて、その分離信頼性係数は 1 に近づくことがその数理上の特徴である。

$$\begin{aligned} \text{受験者群分離信頼性係数} &= \text{誤差調整済み能力推定値分散} / \text{能力推定値分散} \\ &= (\text{能力推定値分散} - \text{受験者平均能力推定値誤差分散}) / \text{能力推定値分散} \\ &= (\text{能力推定値の分散} - \text{受験者群モデル標準誤差二乗和の平均値}) / \text{能力推定値の分散} \\ \text{項目群分離信頼性係数} &= \text{誤差調整済み困難度推定値分散} / \text{困難度推定値分散} \\ &= (\text{困難度推定値分散} - \text{項目平均困難度推定値誤差分散}) / \text{困難度推定値分散} \\ &= (\text{困難度推定値の分散} - \text{項目群モデル標準誤差二乗和の平均値}) / \text{困難度推定値の分散} \end{aligned}$$

表 10 と 11 に観察される最も興味ある事項は、困難度推定値中間範囲集中 5 項目群テストと困難度推定値全体範囲網羅 5 項目群テストそれぞれに関する 5 項目総体としての図 7 に示された受験者群能力弁別性強弱異同とは対称的にそれぞれの 5 項目点双列相関係数平均値絶対値が同一極小値として示されていることである。しかも、前者困難度推定値中間範囲集中 5 項目群総体としての強い弁別力が素点上での項目群弁別力平均として極小的な負の値で以って逆転表示されている。従って、Andrich<sup>7)</sup> (1988, p. 371)による図 7 作成法をも考慮すれば、Andrich<sup>7)</sup> (1988)における discrimination という言葉はデータ素点上での項目弁別力とは概念として異なり、Rasch 測定上での項目群により顕示される受験者群能力弁別性であると理解される。Andrich<sup>7)</sup> (1988)によるその意味概念の上では、Rasch 測定上での項目群弁別力総体大小は表 10 と 11 に示された受験者群分離信頼性係数小大の反転理解に通ずる。すなわち、表 10 における受験者群分離信頼性係数が 0 であり、それは困難度推定値中間範囲集中 5 項目群テストにより 211 名受験者能力がほとんど分離されていないということであり、その 211 名受験者群能力推定値分布幅が狭小である故に、5 項目群総体としての受験者群能力弁別性が大変に強いということになる。以上で、Andrich<sup>7)</sup> (1988)におけるテスト特性曲線比較図(p. 371)に対応する図 7 での困難度推定値中間範囲集中 5 項目群テストの困難度推定値全体範囲網羅 5 項目群テストに上回る特性曲線傾斜角、及び、前後者部分テスト 0・1 データの Rasch 項目分析モデル測定出力上での関連指標値の考察を終える。

次に、本稿分析当該データの Partial Credit モデル測定出力上での閾値母数推定値分布と個別項目弁別力異同との関係を査察する。Partial Credit モデル上での項目弁別力についての有益な参照事項として、Andrich<sup>7)</sup> (1988)により、『被験者位置母数推定値  $\beta_n$  に対応する項目弁別力、それは

つまり  $\beta_n$  に関する  $E[x_{ni}]$  の変化率であり、 $\partial E[x_{ni}] / \partial \beta_n = V[x_{ni}] = \sum_{x=0}^m x^2 \pi_{nix} - (\sum_{x=0}^m x \pi_{nix})^2$  で以って算出される』(pp. 371–372)と指摘されている。項目弁別力は、『単一の数値ではなく、 $\theta_n[\beta_n]$  の関数である』(静<sup>15)</sup>, 2007, p. 280)ことに留意される。 $E[x_{ni}]$  は、被験者  $n$  が項目  $i$  に対して順序カテゴリ  $x$  を選択する Partial Credit モデル上での期待値であり、 $E[x_{ni}] = \sum_{x=0}^m x \pi_{nix}$  として算出される(Wright and Masters<sup>16)</sup>, 1982, p. 147)。なお、 $\pi_{nix}$  は次式で示される Masters<sup>6)</sup> (1982) の Partial Credit モデル確率式(p. 158)である。

$$\pi_{nix} = \exp \sum_{j=0}^x (\beta_n - \delta_{ij}) / \sum_{k=0}^m [\exp \sum_{j=0}^k (\beta_n - \delta_{ij})] \quad x = 0, 1, \dots, m_i$$

又、Facets による Partial Credit モデル測定閾値母数推定値出力は、上式に示された項目  $i$  についてのカテゴリ第  $j$  段階での  $\delta_{ij}$  そのものではなく、 $\delta_{ij} = \delta_i + \tau_{ij}$  として項目  $i$  の位置母数推定値  $\delta_i$  と項目  $i$  についてのカテゴリ第  $j$  段階での閾値母数推定値  $\tau_{ij}$  への分離出力であり、そして、 $\delta_i$  は  $\delta_{ij}$  の平均値として求められている(Andrich<sup>4)</sup>, 2005, p. 31、参照)。

Facets の優れた機能として残差ファイルに各被験者と各項目を一对とするすべての組み合わせ上でのデータ Rasch 測定関連指標値が出力されて、各対に関する Rasch 測定上での応答値分散と応答期待値も含まれている。当該データの Partial Credit モデル測定出力による残差ファイルを MINITAB Release 14 (Minitab Inc., 2003) にコピー貼付した上で、その表転換機能で以って Rasch 測定応答値分散と応答期待値それぞれ 15 項目と 692 名から成る一群を得た上で作成したものが表 12、13、及び、図 8、9 である。Andrich<sup>7)</sup> (1988, pp. 371–372)への上記参照により、Rasch 測定項目弁別力が被験者位置母数推定値に対応する関数上での Rasch 測定応答値分散として算出される。従って、表 12 と 13 においては、各項目弁別力最大値を表すものとして、当該データの Partial Credit モデル測定上での閾値母数推定値増大順序性への違反 9 項目群と準拠 6 項目群に集団化した上で、その二群それぞれにおいて Rasch 測定各項目応答値分散の最大値を降順に項目番号を付して並べている。その違反 9 項目群と準拠 6 項目群それぞれについて、図 8 と 9 では、縦横軸に被験者群に関する Rasch 測定応答期待値と位置母数推定値を配して、各項目特性曲線を与えている。

表 12、13 と図 8、9 により、『閾値母数推定値順序性に違反があれば、その項目弁別力は更に大きくなる』(Andrich<sup>7)</sup>, 1988, p. 370)ことがその傾向上で確かに観察される。但し、第二節冒頭揭示表 1 から察知される様に、当該データ Partial Credit モデル測定上での閾値母数推定値順序性準拠 6 項目それぞれについての閾値母数推定値間の差がすべてかなり小さく、『閾値母数推定値間の差が小さくなればなる程、その項目弁別力は順じて大きくなる』(Andrich<sup>7)</sup>, 1988, p. 370)ことにより、表 12 と 13 での両者項目群総体弁別力の上で大きな違いは観察されない。又、閾値母数推定値の数値順次増大性が保持されていることを以ってしても、閾値母数推定値間での小さな差との相対関係により、かなり大きい項目弁別力として具現され得ることが表 13 と図 9 での項目番号 7 に窺われる。その逆傾向として、閾値母数推定値順序性違反項目であるとしても、その項目弁別力が一様にかかなり大きいとは一概に言えないことが表 12 と図 8 での項目番号 14 と 10 に顕示さ

表 12 閾値母数推定値順序性違反 9 項目  
各弁別力最大値

項目番号	8	12	15	4	13	11	9	14	10
弁別力最大値	1.78	1.59	1.55	1.54	1.48	1.46	1.35	1.15	0.99
平均値	1.43								

表 13 閾値母数推定値順序性準拠 6 項目  
各弁別力最大値

項目番号	7	6	5	2	3	1
弁別力最大値	1.54	1.28	1.21	1.19	1.09	1.08
平均値	1.23					

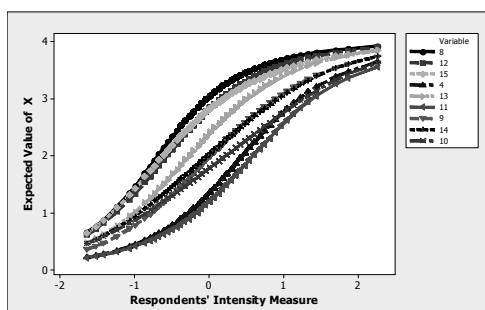


図 8 閾値母数推定値順序性違反 9 項目特性曲線

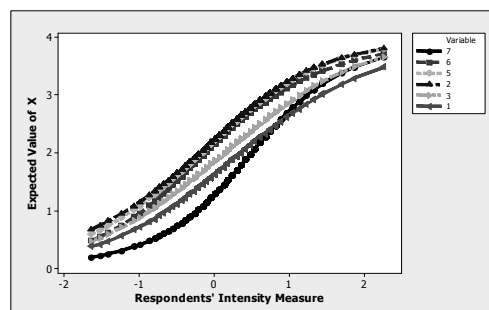


図 9 閾値母数推定値順序性準拠 6 項目特性曲線

れている。又、表 12 と 13 において弁別力最大値として同値 1.54 が付されている項目番号 4 と 7 の素点上での点双列相関係数は第二章冒頭に与えた表 1 によりそれぞれ 0.15 と 0.23 であり、Partial Credit モデル測定と素点の上での指標としての項目弁別力の間に異同のあることが再認識される。

上記の様に、Partial Credit モデル測定上での閾値母数推定値増大順序性についての違反項目群とその準拠項目群は、各項目閾値母数推定値間における差の値小大との相対関係で以って、項目弁別力大小基準の上で一意的に二分され得ないことを顕示するものが下に与える表 14、15、16、及び、図 10 である。上掲表 12 での最下位値 1.15、0.99 が付された項目番号 14、10、及び、上掲表 13 での最上位値 1.54、1.28 が付された項目番号 7、6 の各一對二項目の所属群を入れ換えて構成された弁別力最大値上位 9 項目群データと弁別力最大値下位 6 項目群データのそれぞれ Rating Scale モデル測定上での比較を目的としている。なお、付加情報として表 14 と 15 のそれぞれには素点上での項目弁別力指標として参照された第二節冒頭掲示表 1 における点双列相関係数をも与えている。Partial Credit モデル測定上での弁別力最大値と素点上での点双列相関係数の 15 項目に渡るケンドール順位相関係数  $\tau_{ab}$  は SPSS により 0.127 ( $p$  値=0.517) と出力されている。従って、当該データに関する限りにおいて、両者指標値の間にはほとんど関係がないとみなされ得る。

表 14 と 15 それぞれでの 9 項目群と 6 項目群の構成法により、両群間での項目弁別力最大値平均値の差が当然に前掲表 12 と 13 でのその対応平均値の差よりも大きくなっている。その差の違いにより、図 10 での両群データの Rating Scale モデル測定上での二つの特性曲線の交点での傾斜角の違いが、前節掲示図 6 でのその対応傾斜角の違いと較べて僅かにしても増大している訳である。なお、図 6 と 10 での縦軸には両群データにおける素点上での各データ項目群満点に対する各被験者の各データ項目群得点の百分比としている。これは、Rating Scale モデル測定上での各項目

表 14 弁別力最大値上位群 9 項目の  
弁別力最大値及び点双列相関係数

項目番号	8	12	15	4	7	13	11	9	6
弁別力 最大値	1.78	1.59	1.55	1.54	1.54	1.48	1.46	1.35	1.28
平均値	1.51								
点双列 相関係数	0.25	0.26	0.28	0.15	0.23	0.46	0.18	0.33	0.19
平均値	0.26								

表 15 弁別力最大値下位群 6 項目の  
弁別力最大値及び点双列相関係数

項目番号	5	2	14	3	1	10
弁別力 最大値	1.21	1.19	1.15	1.09	1.08	0.99
平均値	1.12					
点双列 相関係数	0.46	0.28	0.18	0.21	0.26	0.15
平均値	0.26					

に対する各被験者の応答期待値の全項目に渡る総計が項目群素点得点に一致するからである。

表 16 は、図 6 と 10 に関する上記目視上での違いを数値上での違いとして理解するものであり、Partial Credit モデル測定上での弁別力最大値上位 9 項目群データと弁別力最大値下位 6 項目群データそれぞれとこれまでに分析された閾値母数推定値順序性違反 9 項目群データと閾値母数推定値順序性準拠 6 項目群データそれぞれの Rating Scale

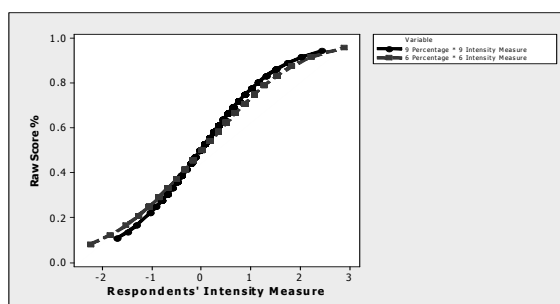


図 10 Partial Credit 測定上での弁別力上位 9 項目群データと弁別力下位 6 項目群データの Rating Scale 測定上でのデータ特性曲線

モデル測定出力上での閾値母数推定値分布幅と分離信頼性係数の比較提示である。その意図により、弁別力高位 9 項目群データ二つと弁別力低位 6 項目群データ二つの組み合わせとしている。

弁別力最大値上位 9 項目群のデータ作成法により、そのデータ項目群全体としての被験者群応答強度位置母数推定値弁別性は閾値母数推定値順序性違反 9 項目群データのそれより大きくなっていると想定される。同様に、弁別力最大値下位 6 項目群データ作成法により、そのデータ項目群全体としての被験者群応答強度位置母数推定値弁別性は閾値母数推定値順序性準拠 6 項目群データのそれより小さくなっていると想定される。組み合わせられたデータ間における Rating Scale モデル測定上での閾値母数推定値分布幅の小さ大が項目群総体としての被験者群応答強度位置母数推定値弁別性の大小に反転呼応することが確認される。Partial Credit モデル測定上での『閾値母数推定値間の差が小さくなればなる程、その項目弁別力は順じて大きくなる』(Andrich<sup>7)</sup>, 1988, p. 370) ことが Rating Scale モデル測定の上でもその項目群総体としての弁別力の想定の下で同様に理解される。

表 16 での各データ項目群分離信頼性係数はいずれも 1.00 に近似する大変に高い値であるが、この指標は被験者数に大きく依存するものであり、各データ項目群が被験者群により単に意味のある程度に分離されていることを示している。例えば、15 項目と被験者数 692 名から成る本稿考察元データについての項目群分離信頼性係数も同じく 1.00 に近似する 0.99 の値が付されているが、

表 16 閾値母数推定値順序性違反 9 項目群データ、弁別力最大値上位 9 項目群データ、  
閾値母数推定値順序性準拠 6 項目群データ、弁別力最大値下位 6 項目群データ  
それぞれの Rating Scale 測定上での閾値母数推定値と被験者群・項目群分離信頼性係数

閾 値	閾値母数推定値順序性 違反 9 項目群データ Rating Scale 測定上での 閾値母数推定値	弁別力最大値上位 9 項目群データ Rating Scale 測定上での 閾値母数推定値	閾値母数推定値順序性 準拠 6 項目群データ Rating Scale 測定上での 閾値母数推定値	弁別力最大値下位 6 項目群データ Rating Scale 測定上での 閾値母数推定値
タウ 1	-0.63	-0.59	-1.00	-1.06
タウ 2	-0.50	-0.31	-0.52	-0.71
タウ 3	0.34	0.10	0.46	0.75
タウ 4	0.78	0.80	1.06	1.03
タウ 4 と タウ 1 の 差	1.41	1.39	2.06	2.09
標準偏差	0.68	0.60	0.93	1.04
各データ被験者群 分離信頼性係数	0.57	0.57	0.54	0.50
各データ項目群 分離信頼性係数	0.99	1.00	0.99	0.97

被験者群一次元性充足程度として表示される 15 項目位置母数推定値不変性成立程度は井澤<sup>1)</sup> (2007)により提示された表 5(p. 19)に高いものではないと顕示されている。表 16 での各データにおける項目群分離信頼性係数に関するその被験者群一次元性充足度についても同様の結果になることに疑いはない。一方、本稿考察元データの被験者群分離信頼性係数は 0.68 であり、筆者の経験則によりその値はデータ素点上での項目群に関する Cronbach のアルファ係数にほぼ近似するものとなり、当該データのそのアルファ係数は SPSS により 0.639 と出力されている。アルファ係数はテスト項目群の素点上での内的一貫性充足度を測る信頼性係数であるが、項目群一次元性充足度を測る上での精度の高い指標ではない。従って、Rating Scale モデル測定上での被験者群分離信頼性係数についても同様に解釈されて、当該データに関する項目群一次元性充足度として表示される被験者群位置母数推定値不変性成立程度は非常に低いと井澤<sup>1)</sup>(2007)での表 4(p. 19)に顕示されている。元データ全体 15 項目による被験者群分離信頼性係数が 0.68 であるから、上表 16 での各データ項目数減少で以って各被験者群分離信頼性係数が 0.68 より低い値になっていることは容易に理解される。

表 16 において筆者に理解矛盾を来たすことは各一対データ間での閾値母数推定値分布幅小大、つまり、項目群全体としての弁別力総体大小、と各データ被験者群分離信頼性係数大小との相対関係である。前掲表 10 と 11 での受験者群分離信頼性係数比較上で、表 10 における受験者群分離信頼性係数が 0 であり、それは困難度推定値中間範囲集中 5 項目群テストにより 211 名受験者群能力がほとんど分離されていないということであり、その 211 名受験者群能力推定値分布幅が狭小である故に、5 項目群総体としての受験者群能力弁別性が大変に強いと理解された。それは、表 10 と 11 に示された受験者群分離信頼性係数小大と項目群弁別力総体大小の反転理解であるが、その反転関係が表 16 には観察されないのである。表 16 での四つのデータについて相対応するいずれの組み合わせ一対上でも、閾値母数推定値分布幅小で以って想定される項目群弁別力総体大が被験者群分離信頼性係数小とはなっていない訳である。表 17 に、前節掲示表 6 の再提示を含めて上表 16 における各データ被験者群分離信頼性係数と各データ Rating Scale モデル測定上での被



験者群位置母数推定値分布幅を与えている。弁別力最大値上位 9 項目群データと弁別力最大値下位 6 項目群データの比較上で、前者データの後者データに上回る項目群弁別力総体が前者データの後者データに下回る被験者群位置母数推定値分布幅と反転呼応していることが表 17 に観察される。従って、前節掲示表 6 についてと同様に、Rating Scale モデル測定上での項目群弁別力総体が大きくなるにつれて被験者群位置母数推定値幅は小さくすると理解される。然しながら、弁別力最大値上位 9 項目群データが弁別力最大値下位 6 項目群データに被験者群分離信頼性において優っていると表 17 に示されている。Partial Credit モデル測定上での閾値母数推定値順序性違反・準拠データの比較上でも同様である。ちなみに、弁別力最大値上位 9 項目群データと弁別力最大値下位 6 項目群データそれぞれについての被験者群分離信頼性係数は以下の様に算出されており、各数値に間違いはない。

$$0.57 = (\text{能力推定値分散} - \text{受験者平均能力推定値誤差分散}) / \text{能力推定値分散}$$

$$\hat{=} (0.247 - 0.105) / 0.247 ; \quad 0.50 \hat{=} (0.368 - 0.183) / 0.368$$

表 17 表 16 での各データ Rating Scale モデル測定上での被験者群分離信頼性係数と各データ被験者群位置母数推定値分布幅

	閾値母数推定値順序性違反 9 項目群データによる被験者群位置母数推定値	弁別力最大値上位 9 項目群データによる被験者群位置母数推定値		閾値母数推定値順序性準拠 6 項目群データによる被験者群位置母数推定値	弁別力最大値下位 6 項目群データによる被験者群位置母数推定値
有効被験者数	692	692		690	691
各データ被験者群分離信頼性係数	0.57	0.57		0.54	0.50
最小値	-2.02	-1.70		-2.87	-2.25
平均値	0.31	0.27		0.11	0.18
最大値	2.46	2.43		2.91	2.90
最大値と最小値の差	4.48	4.13		5.78	5.15
標準偏差	0.50	0.50		0.62	0.61

被験者群位置母数推定値分散については表 17 により弁別力最大値上位 9 項目群データが弁別力最大値下位 6 項目群データに下回っていることが示されているけれども、項目数大小が被験者群平均位置母数推定値誤差分散小大に影響していることが各データ被験者群分離信頼性係数についての上記算出式に窺われる。従って、その項目数大小の影響を本稿考察元データについて最大限に排除した上で、Rating Scale モデル測定上での項目群弁別力総体と被験者群分離信頼性係数との相対関係一端を観察するものが表 18 である。表 14 での項目番号 9 を除外して、更に、項目番号 6 の所属群変更により、弁別力最大値上位群 7 項目と弁別力最大値下位群 7 項目に再編成されている。

表 18 での両群データそれぞれにおける各項目弁別力最大値は両群データの統合による 14 項目データの Partial Credit モデル測定上での値であり、前掲表 14 と 15 での各項目についての値と同値として出力されている。両群データそれぞれの Rating Scale モデル測定上での閾値母数推定値分布幅において弁別力最大値上位群 7 項目データが弁別力最大値下位群 7 項目データに下回っている。それは、本稿これまでの理解通りに、7 項目総体としての被験者群応答強度位置母数推定

表 18 弁別力最大値上位群 7 項目データと弁別力最大値下位群 7 項目データの Rating Scale 測定出力比較

弁別力最大値上位群7項目データ							
項目番号	8	12	15	4	7	13	11
弁別力最大値	1.78	1.59	1.55	1.54	1.54	1.48	1.46
弁別力最大値 平均値	1.56						
タウ 1	-0.62						
タウ 2	-0.22						
タウ 3	0.06						
タウ 4	0.78						
タウ4とタウ1の差	1.40						
標準偏差	0.59						
弁別力最大値上位7項目群データによる被験者群位置母数推定値							
有効被験者数	691						
最小値	-2.78						
平均値	0.29						
最大値	2.20						
最大値と最小値の差	4.98						
標準偏差	0.55						
データ被験者群 分離信頼性係数	0.53						

弁別力最大値下位群7項目データ							
項目番号	6	5	2	14	3	1	10
弁別力最大値	1.28	1.21	1.19	1.15	1.09	1.08	0.99
弁別力最大値 平均値	1.14						
タウ 1	-0.99						
タウ 2	-0.67						
タウ 3	0.64						
タウ 4	1.02						
タウ4とタウ1の差	2.01						
標準偏差	0.98						
弁別力最大値下位7項目群データによる被験者群位置母数推定値							
有効被験者数	692						
最小値	-1.94						
平均値	0.19						
最大値	3.02						
最大値と最小値の差	4.96						
標準偏差	0.56						
データ被験者群 分離信頼性係数	0.53						

値弁別性の上で弁別力最大値上位群 7 項目データが弁別力最大値下位群 7 項目データに上回っていることを明示している。両群データの弁別力最大値平均値の上で示されたその項目群弁別力総体としての違いは、前掲表 14 での弁別力最大値上位 9 項目群データと弁別力最大値下位 6 項目群データについての違いとものとはほぼ同じであり、目視上で前掲図 10 に代替されるものである。表 18 で明確に認知されることは、両群データについての被験者群位置母数推定値分布幅が正に近似しており、被験者群分離信頼性係数が同値である。これは、二つのデータの項目数が同じである場合に、Rating Scale モデル測定上でほぼ同値の被験者群分離信頼性係数が付されるならば、その係数の高低如何に係らず両データ間での被験者群位置母数推定値の分布幅がほぼ同程度であるとの示唆一片である。その傾向一般性は断言され得るものではないが、その傾向的特徴が前掲表 17 での項目数同一データ間でも観察される。又、Rating Scale モデル測定上での被験者群分離信頼性係数は項目数に敏感に反応異動することが表 17 と 18 での弁別力最大値に関する折半データ間でのその係数比較により感知される。表 18 での要点は、両群 7 項目データの少ない項目数で以って、被験者群位置母数推定値の分布幅とその分離程度が両群データにおいていずれも同等に大きくはないということである。それは、両群データ間での僅少としてもその閾値母数推定値分布幅の違いと弁別力最大値平均におけるある程度の差によって意識される程には、7 項目全体としての弁別力総体に両群データ間でその比較の上で大した相違はないとの意味である。従って、両群データいずれにおいても、被験者群位置母数推定値分離程度が同等に低い故に、7 項目全体としての被験者群位置母数推定値弁別性総体がほぼ同等に高いと解釈される。これは、更に、表 18 での両群データの基である 14 項目群データ、ないしは、本稿考察 15 項目群元データにおける項目群総体としての被験者群位置母数推定値弁別性が高いことを示唆しており、それが Rating Scale モデ

ル測定上での当該15項目群データに付された閾値母数推定値の狭小な分布幅に呼応していると理解される。

データに含まれる項目すべての弁別力最大値を同値と規定する Rating Scale モデルの本質的特徴を15項目と被験者数692名から成る本稿考察元データで以って参照として示したものが図11である。その Facets 出力残差ファイル上で各項目弁別力最大値として同値1.0ではなく同値1.30が付されている。この値は当該データの Partial Credit モデル測定上での前掲表14と15に示された15項目弁別力最大値平均値としての1.35に近似しており、両モデル測定上での類似性であると推察される。なお、更なる参照として、上掲表18に与えた弁別力最大値上位群7項目データと弁別力最大値下位群7項目データそれぞれの Rating Scale モデル測定上での各項目弁別力最大値としてそれぞれ同値1.44と1.06が Facets により出力されている。従って、Rating Scale モデル測定による Facets 出力上で同値として供与される項目弁別力最大値はデータ固有のものであると推察される。又、表18について上に観察されたことを踏まえれば、Rating Scale モデル測定上で項目弁別力最大値として同値1.00を超える値が付される場合にはそのデータ項目群全体としての被験者群応答強度位置母数推定値弁別性が強大傾向にあるとの当該データ分析一例上での推測である。これは、Partial Credit モデル測定上でもモデル類似性で以って各項目に付される弁別力最大値についても同様であると推量される。項目群総体としてのその弁別力強大傾向が Rating Scale モデル測定上での閾値母数推定値分布幅の狭小傾向に呼応することは上に理解された通りである。

図11での15項目特性曲線平行性と前掲図8と9での15項目特性曲線交差性がそれぞれ閾値母数を全項目に渡って同一と規定する Rating Scale モデルと閾値母数を各項目の個別属性とする Partial Credit モデルの本質的な違いを顕示するものである。それは各モデル測定上での既述された各項目弁別力最大値の同一性と相違性としてそれぞれ

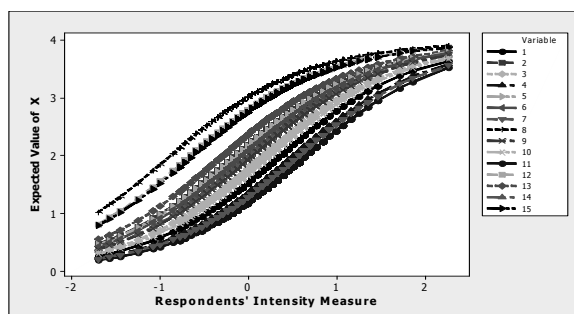


図11 当該15項目群データの Rating Scale モデル測定上での15項目特性曲線

顕現されている。各モデル上での受験者 $n$ による項目 $i$ に対する順序カテゴリー $x$ の選択確率 $\pi_{nix}$ 、その期待値 $E[x_{ni}] = \sum_{x=0}^m x \pi_{nix}$ 、その分散 $V[x_{ni}] = \sum_{x=0}^m x^2 \pi_{nix} - (\sum_{x=0}^m x \pi_{nix})^2$ に基づいて、受験者 $n$ の項目 $i$ への応答に関する標準化残差 $z_{ni} = (x_{ni} - E[x_{ni}]) / \sqrt{V[x_{ni}]}$ として算出される(Wright and Masters<sup>16)</sup>, 1982, pp. 98–99)。その標準化残差 $z_{ni}$ により被験者数を $N$ として $u_i = (\sum_{n=1}^N z_{ni}^2) / N$ が算出されて(Wright and Masters<sup>16)</sup>, 1982, p. 99)、それが Facets 出力上でのモデル適合度指標の一つとして各項目に付される Outfit (Unweighted) Mean Square と呼ばれるものであり、 $N$ に替えて項目数 $L$ を挿入して同様に算出されるものが各受験者についての Outfit Mean Square となる。上記のその算

出過程参照により、そのモデル適合度指標は  $V[x_{ni}]$  に依拠していることが理解される。従って、Rating Scale 測定 Facets 出力上で各項目に付される Outfit  $\chi^2$  平均平方統計量がそのモデル規定としての項目特性曲線平行性からのデータ項目群乖離度についての指標となる。これに関係する参照事項として、Wright<sup>17)</sup> (1996)により、『Rasch 測定上での項目群 Infit 統計量と項目応答理論上での項目弁別力母数推定値それぞれとの相関係数が-0.8 を下回することはめったにない』(p. 6)と指摘されている。

Rating Scale モデルと Partial Credit モデルそれぞれの測定上で当該データ 15 項目に付された上記 Outfit  $\chi^2$  平均平方統計量を示すものが表 19 であり、その表には前掲表 14 と 15 での Partial Credit モデル測定上での各項目弁別力最大値をも併せて含めている。Rating Scale モデル規定としての項目特性曲線平行性からの乖離傾向が項目番号 8、12、4、7、11 に窺われる。又、その 5 項目が表 14 での弁別力最大値上位群 9 項目に属しているけれども、筆者に事象理解不能として各項目弁別力最大値と各項目 Outfit  $\chi^2$  平均平方統計量の表示異同が項目番号 15 と 13 に顕著に観察される。又、項目特性曲線交差性を属性とする Partial Credit モデル測定上で、項目番号 8 ではなく 4 に Outfit  $\chi^2$  平均平方統計量としてその最大値 1.2 が付されていることも筆者には理解不能である。

表 19 Rating Scale と Partial Credit の測定上で当該データ 15 項目に付された Outfit  $\chi^2$  平均平方統計量

項目番号	8	12	15	4	7	13	11	9	6	5	2	14	3	1	10
弁別力最大値	1.78	1.59	1.55	1.54	1.54	1.48	1.46	1.35	1.28	1.21	1.19	1.15	1.09	1.08	0.99
Rating Scale測定Outfit カイ二乗平均平方統計量	1.3	1.2	1.0	1.3	1.2	0.9	1.2	1.0	1.0	0.8	1.0	1.0	0.9	0.9	0.9
Partial Credit測定Outfit カイ二乗平均平方統計量	1.1	1.1	1.0	1.2	1.0	0.8	1.1	0.9	1.1	0.8	1.0	1.1	1.1	1.0	1.1

Rating Scale モデル測定上での前掲図 11、並びに、Partial Credit モデル測定上での図 8、9 のそれぞれに対応する被験者特性曲線について記されている参考文献を筆者は手にしていない。その理解一端を目的として作成したものが図 12、13、並びに、14、15 である。その作成法は、Rating Scale モデル測定上での図 11、及び、Partial Credit モデル測定上での図 8、9 と同様である。当該データの各モデル測定による Facets 出力残差ファイル上で各被験者と各項目の一対に付される応答期待値の MINITAB 表転換機能による表作成上での行配置 692 名の列配置 15 項目に対する各応答期待値、及び、その列配置 15 項目の位置母数推定値を併せた上での表構成に基づくものである。その行に配置した 692 名より無作為抽出された 10 名を一組とする総計 10 組のグラフ目視の上で、両モデル測定いずれにおいても重なっている被験者特性曲線は存在するが交差している被験者特性曲線は皆無である。両モデル測定上でのその作成観察した 10 組のグラフより各モデル測定対応同一被験者群 10 名特性曲線に関する二組の選択で以って図 12、13、及び、14、15 を提示している。

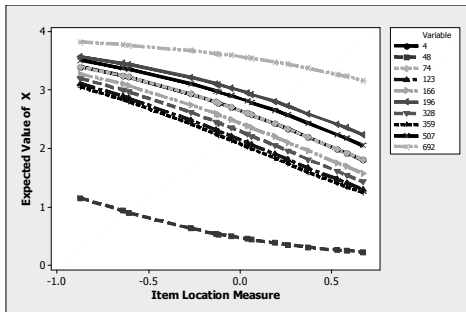


図 12 Rating Scale 測定被験者特性曲線一例

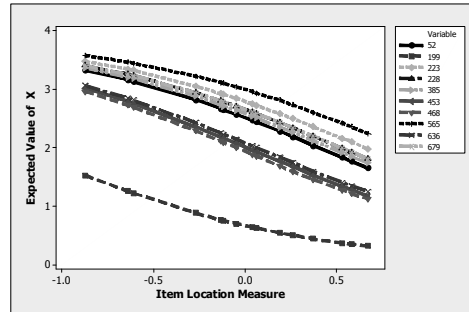


図 13 Rating Scale 測定被験者特性曲線二例

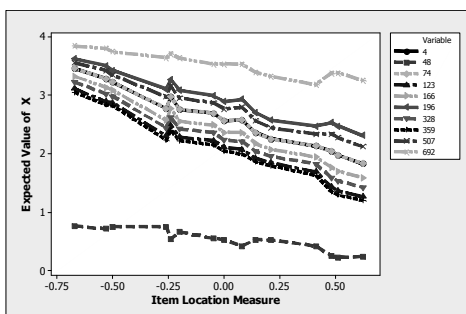


図 14 Partial Credit 測定被験者特性曲線一例

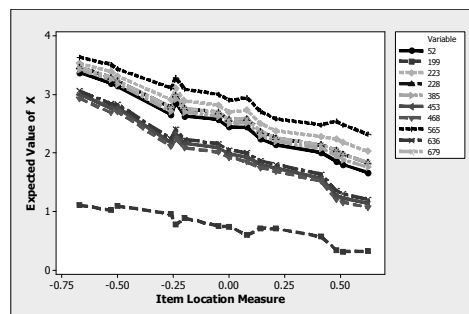


図 15 Partial Credit 測定被験者特性曲線二例

上図一対二組の比較の上で両モデル測定間での被験者特性傾向線同一性が明確に観察され、両モデル測定間での被験者特性曲線上での顕著な違いは平準化状態にあるか否かである。又、各モデル測定における項目特性曲線との目視比較上での興味ある違いが存在する。Rating Scale モデル測定においては、図 11 での 15 項目特性曲線への相応で以って平準化状態にある図 12、13 での被験者特性曲線についても交差が観察されない。一方、Partial Credit モデル測定においては、図 8、9 に平準化状態にある項目特性曲線の交差が観察されるが、図 14、15 での被験者特性曲線に関しては平準化されていない状態の上で交差が観察されない。従って、Rating Scale モデル測定 Facets 出力上では、各項目弁別力最大値同値と同様に各被験者弁別力最大値も同値として規定・設定されていると推測される。但し、Rating Scale モデル測定による Facets 出力上で同値として供与される項目弁別力最大値はデータ固有のものであると推察既述されたが、当該データについてのその各被験者弁別力最大値における同値の値は筆者に知る術がない。Partial Credit モデル測定 Facets 出力上では、各項目弁別力最大値の相違性で以って複数項目特性曲線交差可能性があるけれども、平準化状態ではなく交差状態にもない被験者特性曲線である故に、被験者弁別力はその傾向線上で Rating Scale モデル測定上でのものとほぼ同一であると上図により察知される。それは第二節冒頭掲示表 1 に観察される両モデル測定上での項目群位置母数推定値類似性の上でも道理であるが、当該データに関するその被験者弁別力最大値を筆者には知る術がないことは上記の通りであ

る。当該データの両モデル測定上での項目群分離信頼性係数として同値0.99の値が付されており、被験者群による項目群位置母数推定値の分離程度が大変に高い故に、692名全体としての15項目位置母数推定値弁別性総体が大変に低いとの事象一端だけがいずれのモデル測定上でも筆者に認識される。

本節最後に、前掲表18におけるPartial Creditモデル測定による項目弁別力最大値の上での上位群7項目データと下位群7項目データに関するRating Scaleモデル測定上での閾値母数推定値分布幅と項目群位置母数推定値分布幅との興味ある相対関係を観察する。その目的で以って、下表20には、両群データそれぞれのRating Scaleモデル測定上での項目群位置母数推定値に併せて、両群データそれぞれを0・1データに変換した上でのRasch項目分析モデル測定による各項目困難度母数推定値を与えている。その0・1データへの変換は、元データにおける項目群カテゴリ配点0・1・2・3・4について、前半配点部分0・1を0、後半配点部分3・4を1、そして、中間配点値2を欠損値\*とする配点再構成に基づいている。なお、表20には、表18での両群データそれぞれのPartial Creditモデル測定上での各項目弁別力最大値を標示参照として含めている。

表20 弁別力最大値上位群7項目データ並びに弁別力最大値下位群7項目データのRating Scaleモデル測定項目位置母数推定値と0・1データRasch項目分析モデル測定項目困難度推定値

弁別力最大値上位群7項目データ				弁別力最大値下位群7項目データ			
項目番号	弁別力最大値	Rating Scale 項目位置母数 推定値	0・1データ 項目困難度 推定値	項目番号	弁別力最大値	Rating Scale 項目位置母数 推定値	0・1データ 項目困難度 推定値
8	1.78	-0.76	-1.91	6	1.28	-0.16	-0.42
12	1.59	-0.54	-1.63	5	1.21	-0.23	-0.66
15	1.55	-0.51	-1.84	2	1.19	-0.26	-0.43
4	1.54	0.60	1.78	14	1.15	-0.05	-0.34
7	1.54	0.65	1.97	3	1.09	0.13	0.28
13	1.48	-0.18	-0.57	1	1.08	0.36	0.96
11	1.46	0.73	2.20	10	0.99	0.22	0.60
最小値	1.46	-0.76	-1.91	最小値	0.99	-0.26	-0.66
平均値	1.56	0.00	0.00	平均値	1.14	0.00	0.00
最大値	1.78	0.73	2.20	最大値	1.28	0.36	0.96
最大値と最小値の差	0.32	1.49	4.11	最大値と最小値の差	0.29	0.62	1.62
標準偏差	0.11	0.64	1.91	標準偏差	0.10	0.24	0.62

0・1変換データのRasch項目分析モデル測定上での項目困難度母数推定値を提示している理由は前掲図7で以って確認された様に、困難度推定値中間範囲集中項目群テストが困難度推定値全体範囲網羅項目群テストに受験者群能力弁別性の点において上回る(Andrich<sup>7)</sup>, 1988, pp. 370-371)ことを観察するためである。なお、表20での項目困難度母数推定値とRating Scaleモデル測定上での項目位置母数推定値の両群データ項目総数14個に渡るピアソン相関係数はSPSSにより0.992と出力されており、両者推定値群はその数値間相対関係のみの上ではほぼ同じ情報を有している。前掲表18での両群データRating Scaleモデル測定上での閾値母数推定値分布幅の違いにより弁別力最大値上位群7項目データが項目群全体としての被験者群位置母数推定値弁別性総体において少なくとも同等以上に僅かにしても弁別力最大値下位群7項目データに上回ることが観察された。それは、両群データPartial Creditモデル測定上での弁別力最大値平均値の差とも呼応していると既述された通りである。従って、両群データ項目困難度推定値分布幅比較の上で、弁別力最大値

下位群 7 項目データが Rasch 項目分析モデル測定上で少なくとも同等以上に弁別力最大値上位群 7 項目データよりも広がっているのが道理である。つまり、弁別力最大値上位群 7 項目が困難度推定値中間範囲集中項目群テストを形成して、弁別力最大値下位群 7 項目が困難度推定値全体範囲網羅項目群テストを形成する傾向が観察されて然るべきである。然しながら、大変に興味ある事として、その事象の逆傾向がかなり鮮明に表 20 において観察される。その正確で納得性のある解釈理解は遺憾ながら筆者には不能である。まずは、両群データ変換による 0・1 データの Rasch 項目分析モデル測定上での項目群弁別力総体比較大小についてのその逆傾向事象をそのまま認めることにする。その上で、両群データの Rating Scale モデル測定上での項目群位置母数推定値がすべて絶対値 1 未満の値であることが観察される。従って、Rating Scale モデル測定の上では両群データに内在する同等な高い項目群弁別力総体で以って、両群データの間には被験者群位置母数推定値弁別性に差異はないと考える以外に筆者に術はない。それが、両群データ間での図 10 に代替される同様に傾斜角のかなり鋭いデータ特性曲線と表 18 における被験者群分離信頼性の低い同一程度に示されている。要するに、その被験者群分離信頼性係数からの示唆の上で既述された様に、本稿考察 15 項目群元データにおける項目群総体としての被験者群位置母数推定値弁別性が高いと判断される。その強い項目群弁別力総体、及び、項目群と被験者群いずれについても位置母数推定値分布幅の狭小性が当該データ Rating Scale モデル測定上での各閾値母数推定値として狭小な分布幅を持つ  $\tau_1 = -0.68$ 、 $\tau_2 = -0.48$ 、 $\tau_3 = 0.37$ 、 $\tau_4 = 0.79$  の出力に相互に関連していると推察される。

## 5. おわりに

Rating Scale モデルと Partial Credit モデルは一人の被験者の一つの項目に対する一つの順序カテゴリ  $x$  への応答確率の上では同一物である(Andrich<sup>4)</sup>, 2005, pp. 30–31; Masters<sup>6)</sup>, 1982, p. 163)。然しながら、前節揭示図 11 と 8、9 との対比で示された様に、各閾値母数を全項目に渡って同一と規定する Rating Scale モデル上での項目特性曲線平行性と各閾値母数を各項目の個別属性とする Partial Credit モデル上での項目特性曲線交差性に両モデル間での本質的な違いが顕現されている。項目特性曲線の交差は、被験者群応答強度位置母数推定値尺度上での Partial Credit モデルによるカテゴリ  $x$  の期待値大小がその項目特定曲線交差点を境にして逆転することを意味している。すなわち、特性曲線交差状態にある項目群に関しては、被験者群応答強度性が Partial Credit モデル上でのカテゴリ  $x$  の期待値増大として示されるカテゴリ順次増大性と単調増加関係にないということである。従って、その特性曲線交差項目群については、順序尺度質問紙調査上で想定されているカテゴリ  $x$  の順次増大性とその項目群への被験者群応答強度性に一義的な解釈がなされ得ない。これは、質問紙調査項目群カテゴリ順序段階設定上での順位数値増大性に反するものであり、0・1 データ Rasch 項目分析モデル、並びに、Rating Scale モデルの上で規定される

項目弁別力等位性ないしは項目特性曲線平行性とは本質的に異質である。その意味する処は、母数分離定理(separability theorem)を基底理論とする Rasch モデル(Rasch<sup>18)</sup>, 1960)における「固有客観性」(specific objectivity)(Rasch<sup>19)</sup>, 1977, p. 66)が、Partial Credit モデルにおいて隣接カテゴリー潜在生起境界閾値全数からの観点の上では認められないということである。それが、筆者に本稿考察動機を与えた Andrich<sup>7)</sup> (1988)の論文標題として‘A general form of Rasch’s extended logistic model for partial credit scoring’ が付されている理由である。Andrich<sup>7)</sup> (1988)のその用語 ‘logistic’ に、Partial Credit モデルは Rasch モデルの拡張ではあるが、Rasch<sup>18)</sup> (1960)の理論とは異なる数理形態を持つとの含意(p. 364)が示されている。なお、Partial Credit モデルにおいても、隣接カテゴリー潜在生起各閾値の上では、母数分離定理が充足されている(平越裕之<sup>20)</sup>, 2008)ことに留意される。

最後に、『より少ない前提の上で説明され得ることはより多い前提の上で説明不能となる』との Wright<sup>21)</sup> (1998)引用の Ockam’s Razor (p. 642)が意識される。それについての Wright<sup>21)</sup> (1998, p. 642)によるモデル適用への言及で以って本稿の締めとする。それは、研究推察上での Partial Credit モデル適用における項目間異同尺度設定必要性を質問紙調査研究者に問うものである。

『同一のカテゴリー構成から成る項目群を持つ質問紙調査において Partial Credit モデルを適用することは、Rating Scale モデルによるものとは異なる推察に導く上での項目間異同尺度設定への統計的及び実質的な強い根拠を必要とする。そのような根拠が存在しないならば、実体的には差異のない無意味な識別となる。』

#### 【参考文献】

- 1) 井澤廣行 2007. 「Rasch 評定尺度モデル解題とそのモデル上での閾値順序性」 『流通科学大学論集 — 人間・社会・自然編』第20巻、第1号、pp. 1-25.
- 2) D. Andrich. 1978. A rating formulation for ordered response categories. *Psychometrika*, 43, 4, pp. 561-573.
- 3) D. Andrich. 2004. Understanding resistance to the data-model relationship in Rasch’s paradigm: A reflection for the next generation. In Smith, Jr., E. V. & Smith, R. M. (Eds.), *Introduction to Rasch Measurement* (pp. 167-200). Maple Grove, Minnesota: JAM Press.
- 4) D. Andrich. 2005. The Rasch model explained. In Alagumalai, S, Curtis, D. D., & Hungi, N (Eds.), *Applied Rasch measurement: A book of exemplars* (pp. 27-59). The Netherlands: Springer.
- 5) J. M. Linacre. 1989-2001. *A user’s guide to FACETS: Rasch measurement computer program*. Chicago: Winsteps.com.
- 6) G. N. Masters. 1982. A Rasch model for partial credit scoring. *Psychometrika*, 47, 2, pp. 149-174.
- 7) D. Andrich. 1988. A general form of Rasch’s extended logistic model for partial credit scoring. *Applied Measurement in Education*, 1, 4, pp. 363-378.
- 8) J. M. Linacre. 1999. Category disordering vs. step (threshold) disordering. *Rasch Measurement Transactions*, 13, 1, pp. 675-678.
- 9) J. M. Linacre. 2004. Optimizing rating scale category effectiveness. In Smith, Jr., E. V. & Smith, R. M. (Eds.),



*Introduction to Rasch Measurement* (pp. 258–278). Maple Grove, MN: JAM Press.

- 10) G. Henning. 1988. The influence of test and sample dimensionality on latent trait person ability and item difficulty calibrations. *Language Testing*, 5, 1, pp. 83–98.
- 11) 井澤廣行 2003. 「PROX、UCON、そして母数不変性検証」『流通科学大学論集 — 人間・社会・自然編』第 16 巻、第 1 号、pp. 1–20.
- 12) 井澤廣行 2005. 「Rasch 測定に関する Henning (1988) 論考補遺」『流通科学大学論集 — 人間・社会・自然編』第 17 巻、第 3 号、pp. 1–14.
- 13) J. M. Linacre. 2001. Category, step and threshold: Definitions and disordering. *Rasch Measurement Transactions*, 15, 1, p. 794.
- 14) 井澤廣行 2006. 「Rasch モデル母数推定値同等性検証のためのカイ二乗検定有用性」『流通科学大学論集 — 人間・社会・自然編』第 18 巻、第 3 号、pp. 1–14.
- 15) 静 哲人 2007. 『基礎から深く理解するラッシュモデリング』大阪：関西大学出版部
- 16) B. D. Wright & G. N. Masters. 1982. *Rating scale analysis*. Chicago: MESA Press.
- 17) B. D. Wright. 1996. Comparing Rasch measurement and factor analysis. *Structural Equation Modeling*, 3, 1, pp. 3–24.
- 18) G. Rasch. 1960. *Probabilistic models for some intelligence and attainment tests*. The Danish Institute for Educational Research. (Reprinted in 1980 by the University of Chicago Press with a Foreword and Afterword by B. D. Wright)
- 19) G. Rasch. 1977. On specific objectivity: An attempt at formalizing the request for generality and validity of scientific statements. In Blegvad, M. (Ed.), *The Danish Yearbook of Philosophy*, (pp. 58–94). Copenhagen: Munksgaard.
- 20) 平越裕之 2008 年 5 月 21 日 私的教示。
- 21) B. D. Wright. 1998. Model selection: Rating scale or partial credit? *Rasch Measurement Transactions*, 12, 3, pp. 641–642.