

ロールシャッハテストの最適折半法信頼性

東京大学教育心理学研究室 松 井 仁

Optimum Split-Half Reliabilities for the Rorschach Test

Hitoshi Matsui

A new technique for optimizing split-half reliability estimate was developed. Using this method, internal reliability coefficients for some major Rorschach variables were calculated. The optimum item (card) combinations for each variable was found to provide useful information for insight into item (card) characteristics.

目 次

I. 問 題	
A. ロールシャッハテストの折半法信頼性	
B. 項目特性の測定と反応数の問題	
II. 調 査	
III. 変 数	
IV. 分析と結果	
V. 考 察	
A. 最適折半法信頼性係数について	
B. 最適折半法に基づく項目特性について	
注 一 覧	
文 献	

I. 問 題

A. ロールシャッハテストの折半法信頼性

インクプロットテストの中で最も普及しているといえるロールシャッハテスト（以下、片口（1987）にならないロ・テストと略する）は、松井（1992）に従って刺激図版を項目と見立てる¹⁾と、10項目から構成されていることになる。しかも、この項目選択に関しては、伝統という名の重荷を背負っているとしか言いようのない揺るぎなさがあり、検査者がその目的や状況に応じて検査者の判断で項目を選択したり提示順序を変更することは許されていない。つまり、検査者はこの10項目の組み合わせを絶対的なものとして利用することになる²⁾。よって、ロ・テストに関心を寄せる研究者の興味も、この固定的に構成された10項目の枠内で得られる反応の中に、どれほど被検者特性を反映した情報が含まれているのかといった問題に向けられることになる。

この問題の検討を深める上で求められる論点は多種多

様にわたることとなる（松井，1989）ののだが、本論ではその数ある論点の中から信頼性の問題について資料を提出し、議論を進めることとする。特に、ロ・テストの信頼性の検討として、この10項目の枠内での変数³⁾の内的整合性に注目する。内的整合性とはしばしば項目の1次元性や等質性と表現されることがある。これは、ロ・テストについて言えば「各変数に対してこの10枚の刺激図版（項目）は共通な被検者特性を測定しているか」という問題に言い換えることができる。これは、ロ・テストの集計法が基本的に各項目の変数の項目得点を単純加算してテスト得点とする方法に由来している以上、尺度を構成する上で必須の検討課題と言えるのである。また、この議論の中ではロ・テストの信頼性と共に、各項目がこの検査の測定においてどのように機能しているのかという、項目特性の検討にも触れることになる。

一般に、項目数が固定された項目群で構成される検査の内的整合性の検討手段として、折半法 (split-half method) による分析が用いられる。例えばY-G性格検査やMMPIといった質問紙法タイプの人格検査を標準化する過程では、変数ごとに折半法を含めた内的整合性の信頼性の検討が実行されている。これに対して、ロ・テストに関するこの視点からの基礎研究は次に述べるように非常に少ない。ロ・テストの分析に計量的指標が利用されているにもかかわらず、その信頼性の検討が十分に行き届いていないということは、心理測定論的視点からのロ・テストの価値を引き下げることになり、この検査を実証的な研究に応用したり心理診断のための手法として位置付ける上で大きな障害になっている。

海外の研究では、この課題に対していくつかの基礎研究が見られる。Hertz (1934) や Thorton & Guilford (1936) の研究、Wirt & McReynolds (1953) の研究に

においては「ロ・テストにおいて十分な折半法信頼性が得られた」とされている。これに対して、Vernon (1933)の研究（おそらくこれがロ・テストの折半法信頼性に関して最初の研究であろう）やStein (1960), Stein (1962)においては、否定的な結論が導かれている。しかし、いずれの研究も項目を折半する際の項目選択の方法が、偶数・奇数番目で分けるという折半方法であったり、研究者の経験に基づく折半方法であったりと、各々の研究ごとに恣意的に決められている。よって、これらの研究の結果だけからでは折半法信頼性に関する議論が尽くされたとは考えられない。

このような状況に対してWagner, Alexander, Roos & Prospero (1985)が折半法信頼性の新しい検討方法を提案した。彼らは、これをOptimum Split-Half Reliabilities (以下、本研究では「最適折半法信頼性」と訳す)と称している。そして、投影法検査などの項目特性について未知な領域が残されている検査に対して、その検討を加えるべく従来の折半法の計算方法に工夫を加えたものと説明している。具体的には、可能な5項目ずつの項目の組み合わせ全てについて信頼性係数を計算し、その最大値を最適折半法信頼性係数とするのである。彼らはこの手法の紹介につづいて、ロ・テストデータに対する分析例を紹介した(Wagner, Alexander, Roos & Adair, 1986)。その結果、各変数ごとにこれまでの研究よりも高い信頼性が確認されたことを報告した。この研究の成果に託す彼らの意気込みは、「投影法テストは思ったより信頼されうるものである (Projective techniques are more reliable than we think.)」という表題にも表われていると察することができる。また、彼らはこの手法の導入の利点として、最適折半法信頼性が得られる時の項目の組合せ（これを最適折半法と呼ぶことにする）から項目特性の情報が得られることを指摘した。ただし、後の考察で論じるように、これについては解釈が非常に困難で実用的でないという問題をはらんでいる。実際、Wagner et al. (1986)で示された項目特性に関する議論も、各変数で測定している個々の被検者特性との関連で論じられるところまでは及んでいない。そのため、現実には実際の検査場面に生かせるほどに分析結果が消化されずに議論が断ち切れてしまっている。しかし、このような問題点をはらみながらも、項目特性に関する計量的な議論を実証的な調査結果より可能にすることの突破口を開いたという点で、この手法は高く評価できるものと言える。

B. 項目特性の測定と反応数の問題

通常のロ・テストの実施方法では1項目に対して複数

の反応を許容している。この方法では、同一項目に対して次々と生じた複数の反応の関係を独立なものとして扱うことができない。ここで独立であるとは、ある反応が生じることが他の反応が生じたことに影響されない状況を指す。しかし経験的に言って、ロ・テストでの同一項目内の反応はそれまでの反応に強く影響されており、独立とは言えない。このような過程で得られたデータは、相互に独立な項目特性と被検者特性のもとで発生したデータとは言えない。つまり、項目特性と被検者特性に加わる第3の要因として、それ以前の被検者の反応も反応を規定する要因となっていると考えられる。このような反応の連鎖を継列効果と呼ぶことにする。

ここで、項目特性と被検者特性そのものの情報を純粋に得るためには、この継列効果の要因を除去するための情報が必要である。しかし、現実にはひとたび反応数を制限しない方法でデータを得てしまうと、データからこの要因を除去する情報は得られない。つまり、ロ・テストは実用上の意図からも継列分析を重視するが、そのための前提条件としてKlopfer & Kelley (1942)らが強調している項目特性に関する情報収集は、反応数を制限しないことの結果として、皮肉にもその継列効果のために損なわれることになっているのである。よって、項目特性に関する情報を高い精度で得るためには、通常の実施方法に従わず、むしろ、実験的な制限のものでデータ収集によらざるをえなくなる。そして、継列効果については、このようなデータから抽出された項目特性の情報に立脚して考察が深められるものと考えられる。

ここで、Wagner et al. (1986)の研究の最大の問題点は、通常の項目反応数を制限しない実施方法によるロ・テストデータを分析対象としていることにあると指摘できる。最適折半法信頼性を適用することは、同時に最適折半法の情報も得、その結果として項目特性に関する検討資料を得ることも分析の目的に含まれている。よって、ここではデータに継列効果の混在していない、統制された条件下で得られたデータによる分析を進めることが望ましい。よって本論では、「1項目1反応」の教示で、データ収集時において項目内の継列効果を被らないデータを得、これを分析対象とすることとした。

II. 調査

松井 (1992) で作成したブックレット式の調査用紙⁴⁾を利用し、ロ・テスト反応を収集した。

この調査を依頼した被検者は、年齢が18才から25才の大学生と専門学校生であった。検査の実施はいずれも授

業時間内に行ない、筆者による教示でおよそ1時間で終了した。調査は、1989年6月から1992年7月にかけて実施した。その結果、合計626人の反応を得た。被検者の年齢の平均は21.7才、標準偏差は2.3才であった。ただし、626人の中には時間内にブックレットに編集された最後の項目まで反応できなかった者が12人いた。本研究では、これらの被検者の情報は分析対象から除外し、結果として614人のデータが分析対象となった。

III. 変数

ロ・テストでの反応に対する変数は多種多様である。よって、全ての変数について分析を進めることは不可能であり、次の3つの条件により分析対象とする変数を絞りこみ、スコアリングを実行した。具体的な変数選択の基準は以下の通りである。

- ①ピアソンの積率相関係数を計算する上で、十分な反応頻度がロ・テスト10項目について得られている変数を分析対象とする。
- ②ロ・テストのいずれのスコアリングシステムに関しても共通に導入されている変数を分析対象とする。
- ③集団法による実施方法でも判定に不確定な要素が含まれにくい変数を分析対象とする。つまり、面接法による情報がないと評定できない変数は除外する。

ただし、実際場面での被検者に対する分析に役立つ情報を得ることが目的であるため、ロ・テスト結果を分析する際に重要な役割を持つ変数はできるだけ分析対象として残すこととした。

これらの検討の結果、本論で分析対象として選択した変数の一覧を表1に示す。ここで、上の①の理由により分析対象から除外した変数として、反応領域については特殊部分反応 (Dd) や空白反応 (S) の変数、反応決定因として色彩反応の各カテゴリー (FC・CF・C 単独のもの)、濃淡反応の各カテゴリー (Fc・cF・c 単独のもの) などがあげられる。

分析対象として本論で取り上げた変数は次のとおりである。まず、反応数に関する変数 (R) を分析対象とした。また、反応領域変数に関して全体反応の変数 (W) と普通部分反応の変数 (D) を分析対象とした。

形態水準変数に関しては、通常面接法による情報(特に質問段階での情報)がないと国内で利用頻度の高い片口式の形態水準評定と同等の確信をもって値を評定できないため、この変数は上記の③の理由により分析対象から除外されるべき変数かもしれない。しかし、この変数はロ・テストの解釈過程において被検者の現実検討力を

表1 分析対象とした変数一覧

変数の記号	変数名	評定基準と値	内容
R	反応数	$R_i = \begin{cases} 1 & \text{反応有り} \\ 0 & \text{反応なし} \end{cases}$	
W	全体反応	$W_i = \begin{cases} 1 & \text{反応が全体反応である} \\ 0 & \text{全体反応でない} \end{cases}$	
D	普通部分反応	$D_i = \begin{cases} 1 & \text{反応が普通部分反応である} \\ 0 & \text{普通部分反応でない} \end{cases}$	
FD	Form Definiteness (修正FD) 良い形態性	$FD_i = \begin{cases} 4 & \text{明細化の進んだ反応内容を有する反応} \\ 3 & \text{形態優位な具体的な内容} \\ 2 & \text{形態の具体性にかける反応内容} \\ 1 & \text{非形態的な反応内容} \\ 0 & \text{反応無し} \end{cases}$	
M (graded)	人間運動反応 (段階付け)	$M_i = \begin{cases} 3 & \text{活動的な人間運動反応} \\ 2 & \text{生活空間の増加を示す運動} \\ 1 & \text{通常の運動} \\ 0 & \text{抑制的な運動・表情のみ} \\ 0 & \text{人間運動反応でない} \end{cases}$	
M	人間運動反応	$M_i = \begin{cases} 1 & \text{人間運動反応} \\ 0 & \text{人間運動反応でない} \end{cases}$	
FM (graded)	動物運動反応 (段階付け)	$FM_i = \begin{cases} 3 & \text{活動的な動物運動反応} \\ 2 & \text{通常の運動} \\ 1 & \text{抑制的な運動} \\ 0 & \text{動物運動反応でない} \end{cases}$	
FM	動物運動反応	$FM_i = \begin{cases} 1 & \text{動物運動反応} \\ 0 & \text{動物運動反応でない} \end{cases}$	
m (graded)	無生物運動反応 (段階付け)	$m_i = \begin{cases} 3 & \text{引力に抗する力による無生物運動反応} \\ 2 & \text{引力による運動など} \\ 1 & \text{無生物運動反応でない} \\ 0 & \text{無生物運動反応である} \end{cases}$	
m	無生物運動反応	$m_i = \begin{cases} 1 & \text{無生物運動反応である} \\ 0 & \text{無生物運動反応でない} \end{cases}$	
Movement (graded)	運動反応 (段階付け)	$Movement_i = \begin{cases} 3 & \text{活動的で空間の増大を示す運動反応} \\ 2 & \text{通常レベルの運動反応} \\ 1 & \text{空間の縮小を示す運動反応} \\ 0 & \text{上記3つに当てはまらない反応} \end{cases}$	
Movement	運動反応	$Movement_i = \begin{cases} 1 & \text{運動反応である} \\ 0 & \text{運動反応でない} \end{cases}$	
変数 Movement は、変数M, FM, mの和になっている。つまり、各項目に対する反応に対して、変数M, FM, mの値のうち最も高い値が変数 Movement の値になる。			
Move1 (graded)	筋肉運動反応 (段階付け)	$Move1_i = \begin{cases} 3 & \text{活動的で生活空間の増大を示す筋肉運動反応} \\ 2 & \text{通常レベルの筋肉運動反応} \\ 1 & \text{抑制的な筋肉運動反応} \\ 0 & \text{筋肉運動反応でない} \end{cases}$	
Move1	筋肉運動反応	$Move1_i = \begin{cases} 1 & \text{筋肉運動反応である} \\ 0 & \text{筋肉運動反応でない} \end{cases}$	
変数 Move1 は、変数M, FM の和になっている。			
Move2 (graded)	運動反応 (段階付け)	$Move2_i = \begin{cases} 3 & \text{M以外の運動反応で活動性の高いもの} \\ 2 & \text{通常レベルの活動性のM以外の運動反応} \\ 1 & \text{抑制的なM以外の運動反応} \\ 0 & \text{上記3つに当てはまらない反応} \end{cases}$	
Move2	運動反応	$Move2_i = \begin{cases} 1 & \text{M以外の運動反応である} \\ 0 & \text{上記に当てはまらない反応} \end{cases}$	
変数 Move2 は、変数FM, mの和になっている。			
COLOR	色彩反応	$COLOR_i = \begin{cases} 3 & \text{非形態の色彩反応} \\ 2 & \text{ロールシャッハテストでのCに対応} \\ 1 & \text{CFに対応する反応} \\ 0 & \text{形態優位な色彩反応} \\ 0 & \text{FCに対応} \\ 0 & \text{色彩反応でない} \end{cases}$	
SHADE	濃淡反応	$SHADE_i = \begin{cases} 1 & \text{濃淡反応である} \\ 0 & \text{濃淡反応でない} \end{cases}$	
H	人間反応	$H_i = \begin{cases} 1 & \text{反応内容が人間である} \\ 0 & \text{人間でない} \end{cases}$	
A	動物反応	$A_i = \begin{cases} 1 & \text{反応内容が動物である} \\ 0 & \text{動物でない} \end{cases}$	

(注) 被検者に項目を提示したにもかかわらず、項目に対する反応が得られなかった場合は、いずれの変数においても0点が与えられる。

評価する手掛りとして最も重要な役割を果たしている。よって、この変数の値が得られる前提条件と考えられる HIT⁵⁾の変数 Form Definiteness (良い形体性の変数) を 5 段階評定に修正したものを代用として採用した。ただし、欠測値⁶⁾の処理については、その場合の項目得点を 0 点として扱うこととした。

反応決定因変数に関しては、運動反応の変数である人間運動反応 (M)・動物運動反応 (FM)・無生物運動反応 (m) は、単独ではそれぞれ反応頻度が低く、本来なら①の条件により除外の対象となる。しかし、運動反応の各変数は、体験型の内向性の分析で見られるようにロ・テストの計量的な分析過程で重要な役割を果たしている (黒田, 1988)。よって、これらの変数は除外しないこととした。また、ロ・テスト結果から心理治療の予後を予測する上で有効な指標として知られる RPRS (Klopfers, 1951) では、これらの変数を利用するときの補助変数として、空間における運動の量の評価に基づく多段階評定の変数も活用されているため、分析対象に含めた。また、運動反応に対する分析において反応頻度を確保する目的で、3つの運動反応を込みにした変数 Movement, 変数 M と変数 FM を込みにした筋肉運動反応についての変数 Move1, 変数 FM と変数 m を込みにした潜在的体験型の向性を示す FM+m の記号で知られる指標に対応する変数 Move2 を分析対象とした。ここで、変数 Movement は HIT の同名の変数に相当することになる。また、これらの変数それぞれについて多段階評定の場合については、各変数の記号に graded と記して、二値評定変数と区別した。

色彩反応変数と濃淡反応変数についても、体験型の外向性の分析上重要な指標であることから分析対象とし、各カテゴリーを併合してそれぞれ変数 COLOR, 変数 SHADE とした。ただし、色彩反応の変数 COLOR に関しては、顕在的体験型の指標として利用される変数 ΣC の重み付けに準じて多段階評価を施した。

反応内容変数については、スコアリングシステム間で記号化の種類が多彩である。よって、上の②で述べたようにスコアリングシステム間で共通した変数として捉えにくいので、基本的に分析の対象外とした。ただし、いずれのスコアリングシステムにおいても人間反応変数 H と動物反応変数 A は採用されているため、これらを分析対象に含めた。

IV. 分析と結果

Wagner et al. (1985) の提案した計算方法に基づき、

全ての変数について10項目を5項目づつ2つに分けるあらゆる組み合わせについて折半法信頼性係数を計算した。この項目の組み合わせの総数は ${}_{10}C_5/2=126$ 通りで、全ての組み合わせについて各項目群の得点の相関係数を求め、スピアマン・ブラウンの公式によって全得点の信頼性係数を推定した。この手続きで求められた信頼性係数の最大値が最適折半法信頼係数となり、これに対応する折半法が各変数の最適折半法になる。各変数に対する計算結果を表2に示した。また、各変数に項目の等質性の指標である α 係数 (2 値評定の変数では KR20) も算出し示した。

信頼性係数の解釈のための基準は統一的な意見があるわけではないが、ここでは項目数も考慮して、仮に0.50をその判断の基準として結果を見た。その結果、変数 R, W, D, FD, Movement, Move1, SHADE については、 α 係数及び最適折半法信頼性係数ともに0.50を上回る値が得られた。変数 M(graded), M, COLOR については、最適折半法信頼性係数は0.50を上回ったが、 α 係数の値はこれに達しなかった。これは、項目間相関が損なわれている項目の組み合わせの存在をうかがわせる結果であり、信頼性の低下が推測される。他の変数については、いずれの指標でも満足できる信頼性係数の値が得られなかった。

表2 最適折半法信頼性係数と最適折半法

変 数	最適折半法 信頼性係数	最適折半法		α 係数
		項目群 1	項目群 2	
R	0.71	(I IV VII VIII* X*)	(II*III*V VI IX*)	0.64
W	0.66	(I III*IV VI IX*)	(II*V VII VIII* X*)	0.62
D	0.62	(I III*IV VIII*IX*)	(II*IV V VII X*)	0.55
FD	0.65	(I V VII IX* X*)	(II*III*IV VI VIII*)	0.58
M(graded)	0.53	(I II*III*VIII*IX*)	(IV V VI VII X*)	0.43
M	0.52	(I III*VI IX* X*)	(II*IV V VII VIII*)	0.43
FM(graded)	0.31	(I IV V VI IX*)	(II*III*VII VIII* X*)	0.20
FM	0.32	(I IV V VI IX*)	(II*III*VII VIII* X*)	0.19
m(graded)	0.24	(I III*IV VIII* X*)	(II*V VI VII IX*)	0.11
m	0.29	(I III*IV VIII* X*)	(II*V VI VII IX*)	0.15
Movement(graded)	0.58	(I II*III*IV VIII*)	(V VI VII IX* X*)	0.51
Movement	0.59	(I II*III*IV VIII*)	(V VI VII IX* X*)	0.51
Move1(graded)	0.61	(I II*III*VI VIII*)	(IV V VII IX* X*)	0.53
Move1	0.59	(I II*III*VI VIII*)	(IV V VII IX* X*)	0.52
Move2(graded)	0.31	(I III*IV V VI)	(II*VII VIII*IX* X*)	0.19
Move2	0.33	(I IV V VI IX*)	(II*III*VII VIII* X*)	0.22
COLOR(graded)	0.54	(I II*III*IV VIII*)	(V VI VII IX* X*)	0.45
SHADE	0.73	(I III*IV VII IX*)	(II*V VI VIII* X*)	0.63
H	0.41	(I III*IV IX* X*)	(II*V VI VII VIII*)	0.32
A	0.37	(I III*IV VII IX*)	(II*V VI VIII* X*)	0.24

(注) *がついている項目番号は、彩色項目であることを示す。

V. 考 察

A. 最適折半法信頼性係数について

一般的に考えて、最適折半法信頼性係数を低下させる要因としては、次のものがあるだろう。

- (1)変数に関して項目が等質でない。言い換えると、測定したい被検者特性にとって項目が1次元的な構造を成していない。もしくは項目間にクラスターが存在して、項目間の相関が低下する項目の組み合わせがある。
- (2)そもそも10項目の選択が不適切である。ただし、本論ではロ・テストの10項目を項目選択として絶対的なものとして捉える前提で議論を出発させているため、この問題を考慮することは議論の対象外である。
- (3)変数が項目の特徴を説明するだけのものであるためにその評定値の変動が小さく、被検者の特性を反映する変動が少ない。つまり、このような変数では被検者特性に関して評定値の持つ情報が少なく、結果的に項目間の相関は低下することが考えられる。

ここで変数ごとに最適折半信頼性係数の計算結果を検討してみる。

反応数変数Rについては、最適折半法信頼性係数、 α 係数ともにまずまずの等質性が示された。

反応領域変数としては、全体反応の変数Wと普通部分反応の変数Dを分析対象としたが、変数Rと同様にまずまずの等質性が示された。

形態水準変数としては、先に述べたように良い形体性についての変数FDを検討対象とした。結果は、 α 係数、最適折半法信頼性係数ともに高い水準が保たれていた。よって、尺度として利用する上で、十分な項目に関する等質性を有する変数であることが確認された。

反応決定因変数としては、運動反応、色彩反応、濃淡反応の各変数を分析対象とした。

まず運動反応についてだが、人間・動物・無生物と3つのカテゴリーに分けて単独に変数として扱った場合では、人間運動反応の変数Mに関してはほどほどの等質性が見られるものの(ただし、 α 係数は充分ではない)、残りの動物や無生物の変数での項目の等質性は極端に低いことが示された。ここで、合成変数の分析結果を見てみる。変数Mと変数FMを併合させた変数Move1は、最適折半法信頼性係数、 α 係数ともにほぼ満足できる値が得られた。運動反応変数個々について検討したときの結果とのこの差は、次の様に考えられる。運動反応をその内容よりM, FM, mと3つに分類してしまうと、各項目の

反応頻度が低下することに加えて、反応の出現パターンに項目特徴(おそらくインクプロットの形態の特徴)が強く反映し、結果的に上記の③のように項目の共通性が低下したと考えられる。しかし、変数を併合して被検者の「筋肉運動を見る特性」の強度に関する測定とすることにより、項目の共通性が高くなり、信頼性係数が向上したものと考えられる。また、変数Move1については、多段階評定(graded)したものと「筋肉運動反応である一ない」という2値評定を施したものの両変数について分析を進めた。結果としてはいずれも高い等質性が示された。多段階評定による影響については、この2変数について最適さの上位10位までの折半法のうち7パターンまで同じ折半パターンが含まれた(上位10位までの折半法の提示は省略した)ことより、項目間の構造に大きな変化が生じなかったと解釈できるだろう。つまり、筋肉運動反応では2値データの場合も多段階評定の場合も、ロ・テスト10項目間の等質性に関する構造はほとんど変化していないことが推測される。また、変数M, FM, mのすべての運動反応を併合した変数Movementでは、Move1より最適折半信頼性が低下する結果が見られた。また、ロ・テストの体験型の分析で利用される動物運動反応の変数FMと無生物運動反応の変数mを併合した変数Move2では、変数Movementに比べて極端に等質性の低下が見られた。これらの結果より、無生物運動反応の評定は、他の人間運動反応と動物運動反応といった筋肉運動反応の変数に比べて何か項目の等質性を阻害する要因を包含していることが推測できる。

以上の議論をまとめれば、まず「筋肉運動を見る特性」の視点から見た人間運動反応変数Mと動物運動反応変数FMを合成した変数Move1は、十分な最適折半法信頼性を満たしており、項目の等質性が確認できたと言える。変数Move1にさらに無生物運動反応変数mを合成した変数Movementも、変数Move1に比べれば信頼性係数の低下が見られるものの、その等質性が確認できたと言える。一方、従来から利用されている変数M, 変数FM, 変数mと変数Move2は、最適折半法信頼性の視点より等質性を検討した結果、問題があることがわかった。

色彩反応変数COLORについては、等質性の指標である α 係数の値が十分な値を示さず、最適折半法信頼性係数も他の変数と比べて満足のいく値でなかった。この結果は、ロ・テスト項目の特徴として単色のみの項目(項目I, IV, V, VI, VII)、赤と単色の項目(項目II, III)、多彩色の項目(項目VIII, IX, X)の3種類があり、この特徴によるクラスターが項目間に存在するために生じたものと解釈できる。

濃淡反応変数 SHADE については、 α 係数、最適折半法信頼性係数ともに高い値を得た。この結果を変数 COLOR の場合と対比させてとらえると、変数 SHADE はすべての項目が濃淡刺激として一様に機能していることが推測される。つまり、色彩反応の場合のような項目の特徴による項目間のクラスターはおそらく存在しなかったのだろう。

反応内容変数については、人間反応と動物反応に関する2つの変数について最適折半法信頼性係数を計算した。結果として、 α 係数、最適折半法信頼性係数ともに低く、項目の等質性を確認できなかった。この結果から、運動反応において人間運動反応、動物運動反応の各単独の変数について考えられたのと同様に、反応内容による項目間クラスターの存在が推測できる。すなわち、項目の特徴として人間に見えやすい形態特徴を持つ項目クラスターと動物に見えやすい形態特徴を持つ項目クラスターが存在し、被検者も項目の特徴を語っているだけということが予想される。つまり、先に示した(3)の原因により、このような信頼性の低下が示されたと推測できよう。

B. 最適折半法に基づく項目特性について

本論での分析方法では、信頼性係数だけでなく最適折半法に関する情報も得ることができる。そして、この折半法の項目パターンの内容を見ることにより、項目特性を解釈する手掛りが得られる。最適折半法とは得点の相関が最大になるような二つの項目グループの作り方のことである。ただここで問題となるのは、因子分析法の結果の解釈でしばしば見られる「等質な項目の集まり」といった視点で、この結果の示唆する意味を概観することができないということがあげられる。そして、この困難さが故に変数で示そうとする被検者特性との関連での解釈より、項目の特徴に依存した解釈になる傾向を避け難いことになる。最適折半法のこのような解釈上の問題点に留意しつつ、以下この分析で得られた項目特性について探ってみる。

各変数で最適折半法共通に見られる特徴は、彩色項目の5項目である(以下、ロ・テスト項目についてはローマ数字を使ってその項目を示す)項目II, III, VIII, IX, Xが2対3で2グループに分けられていることであろう。変数M (graded), FM (graded), FM, Move2ではこれに当てはまらなかったが、そもそもこれらの変数は最適折半法信頼性係数と α 係数の少なくともいずれかの値が低く、その折半法自体について検討を加える意味を持たないものと考えられる。このように、信頼性が

一定水準以上に保たれている変数に関して最適折半法による項目群を解釈すると、ロ・テスト項目では彩色項目群と無彩色項目群というクラスターが形成されていることがうかがえる。そして彩色項目群と無彩色項目群とでは、各変数での被検者の得点パターンが異っているであろう。また、彩色項目群の中でも多色彩項目VIII, IX, Xは、必ず1対2で項目群間に分けられている。このことより、この3項目が彩色項目の中でさらにクラスターを形成していると推察できる。これよりロ・テストの集計法で項目VIII, IX, Xの反応数の合計を利用する方法(クローパー式や片口式では変数VIII+IX+X/Rが相当する)の発想は、最適折半法信頼性による検討結果より被検者特性を測定する上で何らかの意味を持ちうるものと言えよう。

注一覧

- 1) 松井 (1989) 以降の一連のインクプロットテストの研究において、筆者は刺激図版を検査の項目と見立てて議論を進めている。
- 2) ロ・テストの刺激図版については、この10枚を「ロ・テスト」という名のもとに与えられた固定的なものとして捉える視点と、数多くあるインクプロット刺激からたまたま抽出されたものとして捉える視点があり得る。
- 3) 松井 (1993) に従って、変数とはロ・テスト反応を分類・評定する各スコアリングを指すこととする。つまり、具体的に言えば、反応数R, 形態反応F, …といった一つ一つのスコアリングを変数と呼ぶのである。
- 4) 本調査用紙は、インクプロット項目を130mm×90mmに縮小印刷したものを利用している。詳細は、松井 (1992) を参照されたい。
- 5) ホルツマンのインクプロットテスト (Holtzman Inkblot Technique) を略してHITと記す。
- 6) 被検者の反応が「何も思い浮かばなかった」などと、項目反応が得られなかった場合、すなわち、変数Rの項目得点が0点の場合が欠測値としての議論の対象となる。

文献

- Hertz, M. R. 1934 The reliability of the Rorschach ink-blot test. *Journal of Applied Psychology*, 18, 467-477.
- 片口安史 1987 改訂 新・心理診断法 金子書房
- Klopfer, B. 1951 Rorschach prognostic rating scale. *Journal of Projective Technique*, 15, 425-428.
- Klopfer, B., & Kelley, D. M. 1942 The Rorschach technique. New York: World Book.
- 黒田浩司 1988 ロールシャッハ・テストにおける認知心理学的アプローチ ロールシャッハ研究, 30, 25-38
- 松井仁 1989 ロールシャッハテストでの信頼性について 東京大学教育学部心理教育相談室紀要, 11, 75-81
- 松井仁 1992 インクプロットテストへの項目反応モデルの応用 教育心理学研究, 40, 29-36
- 松井仁 1993 ロールシャッハテストとその他の投影法検査(渡部編著「心理検査法入門」第13章 分担執筆) 福村出版
- Stein, H. 1960 Rotarion and reliability of the Rorschach.

- Journal of Projective Technique, 24, 171-181.
- Stein, H. 1962 An analysis of two components entering into Rorschach reliability values. *Journal of Projective Technique*, 26, 474-477.
- Thorton, G. R., & Guilford, J. P. 1936 The reliability and meaning of Erlebnistypus scores in the Rorschach test. *Journal of Abnormal and Social Psychology*, 31, 324-330.
- Vernon, P. E. 1933 The Rorschach inkblot test II. *British Journal of Medical Psychology*, 13, 179-205.
- Wagner, E. E., Alexander, R. A., Roos, G., & Prospero, M. K. 1985 Maximizing split-half reliability estimates for projective techniques. *Journal of Personality Assessment*, 49, 579-581.
- Wagner, E. E., Alexander, R. A., Roos, G., & Adair, H. 1986 Optimum Split-Half Reliabilities for the Rorschach : Projective techniques are more reliable than we think. *Journal of Personality Assessment*, 50, 107-112.
- Wirt, R. D., & McReynolds, P. 1953 The reliability of Rorschach number of responses. *Journal of Projective Technique*, 17, 493-494.