Japanese Association of Educational Psychology

The Annual Report of Educational Psychology in Japan
2000, Vol. 39, 96-108

測 定・評 価 部 門

因子分析法の利用をめぐる問題点を中心にして

柳 井 晴 夫
(大学入試センター研究開発部)

1. はじめに

本稿は1998年7月から1999年6月末までに発表された教育心理学研究・心理学研究，同期間中に刊行された書物，および，1998年9月に行われた日本心理学会62回大会および，1999年8月に行われた日本教育心理学会第41回大会の抄録集に掲載された論文，研究発表，および関連学会の中から測定評価に関連するものを選択し，それぞれの内容を紹介するものである。本稿では測定評価の方法のなかでも，教育心理学・心理学の研究に頻繁に利用されている因子分析法の利用をめぐる諸問題をとりあげるが，とくに因子回転の問題について多くの頁を費やすことになる。

TABLE 1 「心理学研究」および「教育心理学研究」において見られる各種因子分析法利用頻度

	1965-1969	1970-1974	1975-1979	1980-1984	1985-1989	1990-1994	1995-1999
A．掲載総論文数	284	305	382	512	511	524	458
B．因子分析利用論文数(／A)	23(8.1)	42(13.8)	33(8.2)	66(12.9)	95(18.6)	117(22.3)	130(28.4)
C．因子分析各手法(／B)							
1. セントロイド法	11(47.8)	19(45.2)	3(9.1)	1(1.5)	0(0)	1(0.8)	0(0)
2. 主因子法	3(13.0)	10(23.8)	16(48.5)	44(66.7)	68(71.6)	70(59.8)	80(61.5)
3. 主成分分析	1(4.3)	1(2.4)	6(18.2)	8(12.1)	12(12.6)	26(22.2)	36(27.7)
4. 最尤法	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	1(0.8)
5. オーソマックス基準(直交回転)							
5.1 バリマックス法	5(21.7)	21(50.0)	22(66.7)	45(68.2)	67(70.5)	84(71.8)	89(68.5)
5.2 その他	0(0)	0(0)	0(0)	2(3.0)	3(3.2)	1(0.8)	3(2.3)
6. ジェオマックス法	0(0)	1(2.4)	3(9.1)	2(3.0)	1(1.1)	0(0)	1(0.8)
7. プロクラステス法	0(0)	0(0)	2(6.1)	1(1.5)	3(3.2)	7(6.0)	5(3.8)
8. 斜交回転							
8.1 プロマックス法	0(0)	0(0)	3(9.1)	2(3.0)	5(5.3)	15(12.8)	24(18.5)
8.2 オブリミン法	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	2(1.5)
8.3 その他	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	2(1.5)

〈注〉 ( ) 内の数値は百分率を表わす。ただし (／A) は「A．掲載論文数」に，(／B) は「B．因子分析利用論文数」に対する百分率であることを示す。

＊本表は，吉村・山住・柳井 (1998) に掲載された表を改訂したものである。なお，1999年は6月までに発行されたもののみを集計した。

2. 因子分析法の適用をめぐって

2.1 教育心理学研究・心理学研究における利用状況

1965年1月から1999年6月までに発行された教育心理学研究・心理学研究に掲載された論文における因子分析，およびその手法の利用頻度について調べた吉村・山住・柳井 (1998) によると，掲載論文総数に対する因子分析を用いた論文数の割合はTABLE 1のように年を追うにつれ次第に増加し，1990年代前半は22%，後半は30%近くになっている。さらに，因子回転の方法としてはバリマックス回転がその主流で，1970～1974年の間に50%の論文でバリマックス回転が施されて以降，常に50～70%の割合で利用されつづけていることがわかる。この「主因子解→バリマックス回転」が因子分析の定石となっているが，最近ではプロマックス斜交回転法の利用が急速に増加している。しかし，それでも，斜交回転を用いた論文数は直交回転を用いた論文数の3分の1以下にすぎない。また表から，因子の抽出法として近年，主成分分析が主因子法にかわって増加していることが興味深い。この理由は，主因子法における共通性の反復推定，時に生じる「不適解」の存在の繁雑さを避けた為であろうか。あるいは，SASで因子分析を行う場合にはデフォルトが共通性の推定値として1を挿入することにより，その結果とし

NII-Electronic Library Service

Japanese Association of Educational Psychology

て因子分析が主成分分析となったケースも多々あろう。また，後に述べるように，因子分析の前段階として，因子分析にかける変数間に内在する多次元的構造を探索的に把握したい場合には，できる限り多数の因子を抽出するという意味で，主成分分析の回転が有効なことがある。

しかし，因子分析にかける変数を規定する因子があらかじめ仮定されていたり，変数間の相関関係をできる限り少数の因子で説明したい場合には，主成分分析でなく，因子分析モデルを仮定した因子の推定法を適用すべきである。こういった場合の因子負荷量の推定法としては，これまで主因子法が多く使われてきたが，SAS や SPSS といったソフトウェアが利用可能な場合には，最小二乗法，重みつき最小二乗法，最尤法などをもっと利用し，可能であれば，異なった手法の適用が分析結果にどのような相違をもたらすかについて考察してほしい。

なお，こういった点に関し，繁桝・柳井・森編（1999）(以降，繁桝他編と略記)の第4節の質問とその回答が参考になる。

2.2 直交回転と斜交回転

ところで，1998年から1999年にかけての1年間に発行された教育心理学研究（第46巻3・4号，47巻1・2号），心理学研究（69巻4・5・6号，70巻1・2・3号）に掲載された論文中，因子分析を用いたとみられる論文数は19編あった。そのうちバリマックス回転を行った論文は13(66%)，斜交回転はプロマックス回転（4）・オブリミン回転（1）・斜交プロクラステス回転（1）であり，いまだにバリマックス回転が主流であることがわかる。バリマックス回転を行っている13の論文のうちには，能力とかストレスといった上位概念に含まれる変数間の因子分析により，数理能力・言語力・空間把握力，さらには精神的ストレス・身体的ストレスといった下位概念の因子を抽出したケースが少なくない。このようなケースには，変数間の相関係数の多くが正で，主成分分析で得られる第一成分の寄与率が高く，しかも第一成分の負荷量のほとんどが正となることが予想される。このような場合，直交回転であると単純構造がえられにくく，2つ以上の因子にわたって比較的高い正の負荷量が得られることがある。

このようなケースに相当するいくつかの事例を最近発行された「教育心理学研究」・「心理学研究」の中から紹介しよう。

例えば，天貝(1999)の TABLE 1 は高校生における信頼感影響経験に関する項目の因子分析においてバリマックス回転で得られた因子負荷量であるが，第4因子に高く負荷する6項目をのぞく37項目は3つの因子（F1）・（F2）・（F3）にすべて正の負荷量をもっている。とくに，F1 と F2 に高く負荷する31項目のうち，実に17項目はもう一方の因子に0.2以上の負荷量を持っている。これは，因子間の直交性という制約があるために，これ以上に明確な単純構造が得られにくいことを意味するものである。F1 は「受容経験」，F2 は「承認経験」の因子と解釈されており，このように命名された因子が無相関であるとは考えにくい。同一データに斜交回転を適用すれば，因子間相関はおそらく0.3を越える筈である。

兵藤・森野（1999）は阪神・淡路大震災直後と2ケ月後に，女子大生の精神的身体的影響に関する40項目の調査を行い，その結果に基づいて因子分析（バリマックス回転）を行い，「余震不安」「思考力低下」「身体的症状」「情緒的混乱」「攻撃性の上昇」「心因反応」「震災からの回避」という7つの因子を抽出した。本研究は内容的には大変貴重なデータを取り扱ったものであるゆえに，あえて苦言を呈しておこう。まず，上記7因子は内容的にいっても無相関とは考えにくい。TABLE 3 には0.4以上の因子負荷量しか掲載されていないが，おそらく，空欄となっている多数の個所で0.2～0.4の因子負荷量は存在するものと推測される。したがって，斜交回転の適用により因子間相関を求めることで，上記7因子間のほとんどのペアの相関は0.2～0.3の値を持つものと推測される。また，同一項目の2回の因子分析でほぼ同一因子が得られたと報告しているが，第一因子である「余震不安」の因子の寄与率が著しく減少していることは，おそらく他の因子に少なからず影響を与えているはずである。この意味で，2つの因子分析の結果として得られた因子の同等性を論じるにあたっては，因子比較の諸技法（柳井他（1990）第5章参照）の適用，さらには共分散構造分析の多母集団モデル（狩野（1997），豊田（1998）参照）の適用を検討してほしい。

この他，柏木・永久（1999）による「子供を産むこと」の理由についての質問項目に基づくバリマックス回転後の因子分析の結果（同論文 TABLE 1 参照）は，30変数のうち16変数が2つの因子に0.3以上の負荷量をもち，単純構造を有していないことを示している。また，内容的にも「情緒的価値」「条件依存」「自分のための価値」「社会的価値」「子育て支援」といった因子は，「子を生む価値観」といった上位概念の下位概念に相当する5因子であり理論的にも無相関でないことは明らかである。

これらの諸事例にみられるように，因子間相関が明らかにゼロでないと仮定される場合には，直交回転でなく斜交回転を適用すべきである。もちろん，斜交回転を適用して因子間相関がすべて0.2以下であるような場合は直交回転を適用しても斜交回転の結果と大差はない。したがって，因子回転にあたっては，まず，斜交回転を行

NII-Electronic Library Service

Japanese Association of Educational Psychology

い，因子相関が低い場合に限って直交回転をすべきといえよう。筆者自身，新性格検査の尺度構成において，斜交プロマックス回転を適用し，120の項目が10項目ずつ12の尺度にきれいに分離された結果（柳井・柏木・国生，1987）を得て以来，上記のように考えるようになってきた。なお，同様の主張は，豊田（1998）のp.280にも明記されている。

次に回転に関するひとつの数値例を紹介しよう。TABLE 2-1に1979年に実施された共通1次学力試験5教科科目の相関係数行列を示した。最大の相関係数は0.644，最小は0.402と明らかに高い相関である。主成分分析の結果(TABLE 2-2)，第1固有値は3.197，その寄与率は63.9%で，この成分は明らかに総合学力の成分である。第2成分の固有値は1以下であるが，寄与率は12.8%あるので必ずしも成分の解釈を止めることはない。この成分は，文系（プラス方向）・理系（マイナス方向）の両極因子である。

ここで，共通性の初期値をSMCとして，反復推定による因子分析を行うと，明らかに1因子構造になり，2つの因子を求めることは適切でないという結論がみちびかれる。そこで，因子分析モデルを一旦脇において，主成分分析によって得られた主成分負荷量を成分数＝2でバリマックス直交回転をした結果をTABLE 2-3に示す。第1因子は数学・物理に高い負荷，第2因子は国語・外国語・社会に高い負荷がみられた。続いて，TABLE 2-4はプロマックス斜交回転の結果として得られた因子パターン係数・部分相関係数・因子構造係数を示した。因子パターン係数は，社会と英語は単純構造になっていないが，国語，数学，理科は，バリマックス解の結果に比べて明らかに単純構造に近くなっていることがわかる。

ところで，斜交回転の方法としては殆どの論文でプロマックス回転法が利用されているが，歴史的にはオブリミン回転の方が早く開発されている。1997年に因子分析の権威マイケル・ブラウン教授が来日された際，斜交回転の方法としてプロマックス回転法とオブリミン回転のどちらが優れているかという質問をしたことがある。このとき，ブラウン教授はためらわずオブリミン法が優れていると回答し，プロマックス回転法については批判的であったことが印象的であった。プロマックス回転とオブリミン回転は方法論的には全く異質なものであり，データによっては2つの方法の使い分けをする必要があるか否かについて今後研究していく必要があろう。なお，安達（1998）はオブリミン回転を行っているが，得られた因子パターンを見るとプロマックス回転の結果に比べ，単純構造性が満たされにくいような印象をうけた。

2.3 斜交解の寄与率について

次に斜交解の寄与率について説明しよう。因子パターン係数 a_{j1}，a_{j2} は，因子（f_1），因子（f_2）を説明変数，x_j $(j=1, \cdots, 5)$ のそれぞれを基準変数とした重回帰モデル

$$x_j = a_{j1}f_1 + a_{j2}f_2 + \varepsilon_j \quad (j=1, \cdots, 5)$$

としたときの，a_{j1}, a_{j2} の推定値で，r_{12} を2つの因子間の相関係数，$r_{(j)1}, r_{(j)2}$ を x_j と f_1，x_j と f_2 の相関係数とすれば

$$a_{j1} = (r_{(j)1} - r_{(j)2}r_{12})/(1-r_{12}^2),$$
$$a_{j2} = (r_{(j)2} - r_{(j)1}r_{12})/(1-r_{12}^2)$$

となる。また，TABLE 2-3の右端に記述されている共通性の値は x_j を基準変数，f_1，f_2 を説明変数とした重相関係数の平方で，次のように分解される。

$$(r_{(j)1}^2 + r_{(j)2}^2 - 2r_{(j)1}{}^*r_{(j)2}{}^*r_{12}/(1-r_{12}^2))$$
$$= r_{(j)1}^2 + ((r_{(j)2} - r_{(j)1}r_{12})/\sqrt{(1-r_{12}^2)})^2$$
$$= r_{(j)1}^2 + r_{(j)2/1}^2 \quad \cdots(1a)$$

または

$$= r_{(j)2}^2 + ((r_{(j)1} - r_{(j)2}r_{12})/\sqrt{(1-r_{12}^2)})^2$$
$$= r_{(j)2}^2 + r_{(j)1/2}^2 \quad \cdots(1b)$$

(a_{j1}, a_{j2}) は因子パターン係数，$(r_{(j)1}, r_{(j)2})$ は因子構造係数，$(r_{(j)1/2}, r_{(j)2/1})$ は変数 x_j と因子 f_1，変数 x_j と因子 f_2 のそれぞれから因子 f_2，f_1 の影響を一方の変数のみから除去した部分相関係数を示したものである。なお，因子パターン係数と部分相関の間には次の関係式が成立する。

$$r_{(j)1/2} = a_{j1}\sqrt{(1-r_{12}^2)}, \quad r_{(j)2/1} = a_{j2}\sqrt{(1-r_{12}^2)}$$

ここで，SASで利用されている寄与率に次の2つがあることを明記しておく。

A）他の因子の影響を除去したもの（部分相関の平方和として計算される）

B）他の因子の影響を無視したもの（因子構造係数の平方和として計算される）

TABLE 2-4の下段の4つの数値に関し，次の等式が成立することに注意しよう。

$$1.358 + 2.483 = 1.134 + 2.707 = 3.841$$

つまり，部分相関の平方和による第1因子の寄与(1.358)と因子構造係数の平方和として計算された第2因子の寄与（2.483）の和が，部分相関の平方和による第2因子の寄与（1.134）と因子構造の平方和による第2因子の寄与（2.707）の和に等しく，その値は5変数の共通性の平方和3.841に等しくなっている。このことは，共通性の値に等しい重相関係数の平方が（1a），（1b）式のように2通りに分解されることから導かれるものである。上記に示したA）による寄与率の総和は多くの場合1より小さく全体的に寄与率は過小評価される。一方，B）による寄

NII-Electronic Library Service

Japanese Association of Educational Psychology

与率の総和は1を越え，過大評価される。この欠点を除くひとつの寄与率の定義の仕方は，因子 f_1，f_2 を対等なものとみなさず，まず第1因子の寄与はB）の方法，第2因子の寄与の方法はA）の方法を用いるようにすれば，寄与率の総和を100%にすることができる。これは，繰り返しの等しくない2元配置分散分析の平方和の分解におけるType Iの方法（市川・大橋他，1987）と同一の方法になる。

なお，次に共通因子数を3にした場合の因子パターン行列・部分相関・因子構造係数の結果をTABLE 2-5に示した。f_1, f_2, f_3 の順に順序をつけた因子を $RF_1=f_1$, $RF_{2/1}$ を f_2 から f_1 の影響を除去した因子，$RF_{3/1,2}$ を f_3 から，f_1 と f_2 の影響を除去した因子として，その因子寄与を求めると，2.411，1.182，0.697となり，その合計は4.291と，主成分分析から得られる3つの成分の固有値の和に一致する。なお，TABLE 2-5の左側に示した因子パターンの結果は，5教科が，(F_1) 数学と理科，(F_2) 国語と外国語，そして，(F_3) 社会と理科，の3つの成分に分かれることを示し，これらの相関が，$r(F_1, F_2)=0.448$, $r(F_1, F_3)=0.537$, $r(F_2, F_3)=0.535$ と，比較的高くなることが導かれた。この結果は，因子分析モデルの適用からは，導かれないものである。このケースのように，探索的因子分析のひとつの手法として，主成分分析結果の回転が意味を持つ場合があることに注意しておこう。

共通第1次学力試験

TABLE 2-1 5教科科目間相関係数行列

国語				
社会	0.537			
数学	0.402	0.504		
理科	0.478	0.624	0.644	
外国語	0.562	0.587	0.563	0.576
	国語	社会	数学	理科

TABLE 2-2 主成分分析

	PF_1	PF_2	PF_3	平方和
国語	0.736	0.573	0.182	0.902
社会	0.817	0.105	−0.498	0.926
数学	0.779	−0.470	0.269	0.900
理科	0.837	−0.271	−0.178	0.805
外国語	0.826	0.103	0.257	0.758
固有値	3.197	0.644	0.451	4.291

（例1） 共通因子数＝2の場合

TABLE 2-3 バリマックス直交回転

	F_1	F_2	共通性
国語	0.157	0.919	0.869
社会	0.533	0.628	0.678
数学	0.892	0.177	0.828
理科	0.801	0.364	0.773
外国語	0.541	0.632	0.692
平方和	2.038	1.803	3.841

TABLE 2-4 プロマックス斜交回転

因子パターン		部分相関		因子構造係数	
F_1	F_2	F_1	F_2	F_1	F_2
−0.115	0.989	−0.098	0.837	0.410	0.924
0.404	0.534	0.343	0.452	0.688	0.749
0.957	−0.096	0.811	−0.081	0.907	0.412
0.793	0.146	0.672	0.123	0.878	0.567
0.412	0.536	0.349	0.454	0.697	0.755
／	／	1.358	1.134	2.707	2.483

因子間相関＝0.531

TABLE 2-5 （例2）共通因子数＝3の場合
プロマックス斜交回転法

	因子パターン			部分相関			因子構造係数		
	F_1	F_2	F_3	F_1	F_2	F_3	F_1	F_2	F_3
国語	−0.072	0.947	0.062	−0.059	0.779	0.048	0.385	0.948	0.530
社会	−0.042	0.100	0.928	−0.034	0.082	0.720	0.501	0.578	0.959
数学	0.966	−0.062	−0.028	0.794	−0.003	−0.023	0.948	0.413	0.487
理科	0.520	−0.062	0.542	0.428	−0.052	0.421	0.783	0.460	0.787
外国語	0.451	0.567	0.003	0.371	0.467	0.003	0.707	0.771	0.549
平方和	／	／	／	0.956	0.835	0.699	2.411	2.208	2.358

NII-Electronic Library Service

Japanese Association of Educational Psychology

2.4 因子パターンか因子構造か

斜交回転の場合因子負荷量として，因子パターン係数・因子構造係数が与えられる。このうち，どちらを因子分析の結果として論文中に明示すべきであろうか。結論からいえば，因子パターン係数を明示すべきである。この理由について，繁桝他編 (1999) のQ67において，明確な解説がされている。この書の p.137 の FIGURE 1, 2 を以下に再掲しよう。

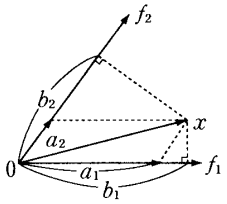

FIGURE 1　2つの因子 (f_1, f_2) が正の相関をもつ場合

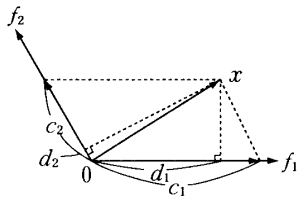

FIGURE 2　2つの因子 (f_1, f_2) が負の相関をもつ場合

(出所：繁桝他編 (1999) Q&Aで知る統計データ解析，サイエンス社 p.137 図1，2より引用)

FIGURE 1, FIGURE 2 においては，a_1, a_2, c_1, c_2 が因子パターン，b_1, b_2, d_1, d_2 が因子構造である。

したがって，f_1 と f_2 という2つの因子間相関を正と仮定した図1の場合は，因子パターンが単純構造，図2の場合は因子構造が単純構造を持っていることがみてとれる。したがって，因子間相関係数が正の場合，因子構造係数に比べ因子パターン係数の方が単純構造性を持ちやすくなることが推察される。なお，相関係数行列を R，因子パターン係数ベクトルを w，因子構造係数ベクトルを a とすると，$a=Rw$，となる。イメージ因子分析におけるアンチ・イメージ変数行列を Z_A とすると

$(1/n)(Z_A)'(Z_A)w=a$,

ただし，$w=D^{(-1/2)}a$，$a=D^{1/2}w$,

$D^{1/2}$, $D^{(-1/2)}$ はそれぞれ，各変数の SMC を1から引いた値の平方根，およびその逆数を対角成分とする対角行列である (Yanai & Mukherjee, 1987) ことから，イメージ変数の因子構造が元の変数の因子パターンに，イメージ変数の因子パターンが元の変数の因子構造にほぼ対応することがわかる。

2.5 その他の問題

重回帰モデルにおいては偏回帰係数の推定値を標準偏差 (標準誤差) で除したものが t 分布にしたがうことを用いてパラメータ=0の仮説を検定できることは周知であろう。

これと同様に，Calis や Amos などの共分散構造分析のプログラムを用いることにより，検証的因子分析における因子負荷量の推定値の標準誤差は計算可能となる。理論的には，主因子法に基づく因子負荷量の標準誤差だけでなく，バリマックス回転後の負荷量の標準誤差の計算法もすでに Archer & Jennrich (1973) で発表されている。Ogasawara (1998, 1999) は因子分析のプロクラステス回転，プロマックス回転で得られる因子パターン係数の標準誤差をはじめ，各種の推定値の標準誤差を求める研究 (Ogasawara (1999)) 等を精力的に行っている。残念なことに，SAS や SPSS 等の因子分析の標準的プログラムには因子負荷量の標準誤差を求めるプログラムが完備されていない。今後，因子分析のプログラムパッケージを作成する機会があれば，是非，標準誤差が計算可能となるように配慮すべきであろう。

3. その他の統計的手法について

3.1 共分散構造分析

共分散構造分析がどのようなデータの解析に有効かについては，共分散構造分析の21事例を紹介した豊田編 (1998) をまず読むべきであろう。紹介された事例の多くは，1998年に立教大学で行われた第26回日本行動計量学会でのリレーセッションで発表されたもので，「タレントの好感度の規定要因」「人事評価の満足を決めるもの」「バニラカップアイスの官能評価モデル」「同調性とリーダーシップの因果モデル」等バラエティに富んだ内容である。また，付録に掲載された指定討論者，さらにそれらに対する豊田氏の回答も読み応えのある力作である。また，共分散構造分析の適用例を扱った解説書として山本・小野寺 (1999) も出版されている。これらを読んだ読者は，共分散構造分析の理論と適用例について詳細に記述されてある，狩野 (1997)，豊田 (1998)，の専門書に進めばよい。

豊田・村石 (1998) においては共分散構造モデルの下位モデルとして記述される遺伝 ACE モデル (豊田, 1997) を YG 検査の12尺度をそれぞれの変数として，男女別の一卵性双生児および二卵性双生児，異性双生児の5つの母集団に適用し，仮定された8つのモデルのうちから最適モデルを探索した。さらに，12の尺度を情緒の適応性，広義の向性といった2つの因子に分類し，一般人と双生児に同一の共分散構造を当てはめた分析を行っている。以上2つの分析から，向性に関する尺度は遺伝的規定性の影響を受ける度合いが強く，情緒の適応性は遺伝的要因の影響は受けず非共有環境の影響を受けているという，興味ある結果を導いている。更に，村石・豊田 (1998) では，数学・理科・国語・社会の学力について同様の分析を行っており．数学・国語は遺伝的要因の影響を受けず，理科・社会は遺伝的要因の影響を強く受けているという

NII-Electronic Library Service

Japanese Association of Educational Psychology

結果を導いている。この2つの解析は我が国における共分散構造分析の第一人者が行った分析として，今後共分散構造分析を学ぶ若い研究者にとって模範となる貴重な研究であり，方法論的には複雑な共分散構造モデルを駆使した優れた分析例であることは確かであるが，第2の論文中においてみられる，数学・国語が遺伝的要因の影響を全く受けておらず，理科・社会は50%以上の遺伝的要因を受けているという教科科目に関して不連続な印象を与える記述には違和感を覚える。一卵性双生児の数に比べ二卵性双生児の数がきわめて少ないという事実は上記の結果に影響を与えていないであろうか。共分散構造分析の結果を補強するために，ペアとなる一卵性・二卵性双生児間で，性格尺度の相関係数・学力検査成績に関する平均値の比較等の基礎集計についても言及すべきであろう。これらの分析例で用いられている多母集団モデルは，共分散構造分析の適用事例にもっと利用されてしかるべきである。田中他（1998）の理系・文系別に行われた共分散構造分析の比較には，やはり多母集団モデルを利用すべきであろう。岡田（1999）は，共分散構造分析の多母集団モデルを用いて，同一項目についての男女間の因子不変性を確認した後，男女間の平均の比較を行っている。

なお，共分散構造分析におけるモデルとデータの適合度を示す統計的指標として，狩野（1997）は特に標本数が少ない場合のカイ二乗検定の利用を推奨しているが，豊田（1998）はカイ二乗検定の利用をあまり推奨していない。筆者はどちらかといえば，後者の立場を推奨したい。

3.2 項目反応理論

過去1～2年，教育心理学研究においては項目反応理論を用いた論文は見られない。項目反応理論を用いた論文としては，大学入試センター研究紀要に掲載された平・小野・前川（1998）が目につく程度である。一方，学会総会の発表として，1998年の日本心理学会，および1999年の日本教育心理学会総会で，日本語基礎能力テストの簡易得点法・テストの得点分布を考慮した項目選択法を扱った平・前川（1998），中村・前川・菊地（1998），一般化部分採点モデルに関する服部(1998)，項目反応理論を利用した項目選択法(藤森，1999)，項目反応理論に基づく等化法(小笠原，1999)がある。項目反応理論に関する論文が少ない点を鑑み，こういった発表は早急に論文としてまとめることを要望したい。

ところで，1996年12月に設立された外国語評価学会〈会長：大友賢二；常盤大学教授〉はその活動として，1998年9月に研究紀要（JLTA Journal）創刊号を出版しただけでなく，1999年7月に第21回言語テスト国際会議を共催した。この国際会議においては，項目反応理論を用いた数多くの研究発表がなされたことを明記しておく。

項目反応理論の発祥地であるアメリカのETS（Educational Testing Service）では，Graded Response Model (Samejima, 1969)，Partial Credit Model (Muraki. E., 1992)等の多肢選択のテストデータを分析する項目反応理論モデルが実際のデータ解析に利用されている。また，センター試験のようにいくつかの大問の下に複数個の設問があり，科目内のすべての項目の局所独立性が成立しにくいような場合には，大問をテストレット（テストの集合）としてとり扱うことのできるように拡張された項目反応理論の研究も行われている。アメリカで開発され，日本でも利用者の多い，TOEFLなどの試験は最近コンピュータテスト（Computer Based Test）化された。このためにも，アメリカに比べかなり遅れている項目反応理論の研究と実践が急務の課題といえよう。このためには，容易に利用可能な項目反応理論のプログラムの普及とその解説書の発行が望まれるところである。なお，最近のアメリカにおける項目反応理論の発展とその関連プログラムに関しては，ETSの村木英二氏（emuraki@ets.org）に尋ねるとよいだろう。なお，ごく最近になって，項目反応理論の最新の成果をとりいれた著書（渡部，野口，1999）が出版されたことを明記しておこう。

3.3 重回帰分析利用上の問題点

重回帰モデル　$y=\beta_1 x_1+\beta_2 x_2+\cdots+\beta_p x_p+\alpha+\varepsilon$，において β_j の推定値 b_j を偏回帰係数，$b_{s(j)}=(s_{xj}/s_y)b_j$ が標準偏回帰係数と呼ばれるものである（s_{xj}，s_y はそれぞれ，x_j，y の標準偏差）。ところで，b_j の値は x_j 以外のすべての変数の値を固定させた場合，x_j を1単位増加させた場合の y の増加分を示すものである。しかし，標準偏回帰係数 $b_{s(j)}$ の値は x_j を1標準得点分増加させた時の y（1標準得点あたり）の増加を意味するもので，意味的には偏回帰係数に比べわかりにくい。また，b_j が固定している場合，標準偏回帰係数 $b_{s(j)}$ は説明変数 x_j の分散（s_{xj}）だけでなく，基準変数 y の分散に影響をうけることがわかる。生物統計学者のGreenland et al.(1986)はこういった理由により，医学における標準偏回帰係数の利用に警鐘をならした。

喜岡・柳井（1997）は上記，および，Newman, T.B. & Browner, W.S. (1991), Greenland, S. et al. (1991) といった計3つの論文の要旨を紹介したあと，心理学や社会学の行動科学においても同様なことがあてはまることを指摘し，重回帰分析を用いる場合，説明変数の尺度の原点や単位が任意でかつ意味が明確でないときには標準偏回帰係数の利用を，原点や単位の意味が明確な場合には偏

NII-Electronic Library Service

Japanese Association of Educational Psychology

回帰係数の利用を薦めている。教育心理学研究で用いられる性格尺度・態度尺度が説明変数となる場合は標準偏回帰係数，医学や疫学で用いられる最大血圧値と血清コレステロール値を用いてある疾患の進行度を予測するような研究の場合には偏回帰係数を用いるべきということになる。例えば，最大血圧値を10下げた場合，最大血圧に与えられた偏回帰係数を10倍した分だけ進行度が改善されることになる。いずれにしても，偏回帰係数と標準偏回帰係数のいずれを選択するか，あるいは回帰分析の使用を止めて他の手法を選択するかは，研究目的・データの収集法・研究結果を実際の現場にどのようにフィードバックすべきかといった様々な要因に影響されるものであることを指摘しておこう。その意味で，判別分析，重回帰分析やパス解析，共分散構造分析において変数の単位や意味に無関係に標準偏回帰係数を利用することは慎むべきであるといえよう。

重回帰分析で多くみられる誤りに，説明変数が2つ以上ある場合の偏回帰係数の解釈がある。説明変数 x_1，x_2 と2つの場合，x_2 に対応する偏回帰係数 b_2 の値がゼロであるからといって，y と x_2 の相関係数が必ずしもゼロではない。y と x_1 の相関係数を r_{01}，x_1 と x_2 の相関係数を r_{12}，y と x_1，x_2 の重相関係数（R），とすれば，

$$R^2=(r_{01})^2+(b_2)^2(1-(r_{12})^2)$$

という関係が成立し，$b_2=0$ の場合には，R と r_{01} の値は等しくなる。つまり，x_2 は y の予測には有用であるが，x_1 と x_2 を用いて同時に y を予測する場合は有用ではなくなる。したがって，岩尾（1999）が述べているように，

（確証度の予測値）＝0.78**（類似度）＋0.06（被覆度）

という予測式が得られ，被覆度の係数0.06が有意でなかったとして，確証度の判断に類似度のみが利用され被覆度は利用されていなかったとする解釈は，類似度と確証度の単相関係数が有意な相関を持たない場合を除いては誤りである。一般的には，上記の結果は，被験者が確証度の判断に類似度を利用している場合には，被覆度の利用は確証度の予測を高めるものとはならない，という解釈が正しいものとなる。同様の誤りは，冷水・冷水(1999)が行っているTABLE 3の数量化1類の重み係数の解釈にもみられる。すなわち，この論文において著者は，手の位置・手の向き・手・腕の動きといった説明変数別に，基準変数である各尺度の関係をまとめているが，数量化1類の各項目の重み係数の解釈も重回帰分析の偏回帰係数の解釈と同様で，手の位置の全てのカテゴリーがある尺度に小さな重みしか持たないとしても，手の位置単独でその尺度と関連を有していることがある。なお，偏回帰係数の解釈をめぐる詳細については，繁桝他（1999）の第3章の該当個所を参照されたい。

3.4 検定力分析について

杉澤（1999）は1992年から1996年までの5年間に発行された「教育心理学研究」に掲載された論文のうち統計的仮説検定の使用されている207論文において，対立仮説として，1）小効果量，2）中効果量，3）大効果量，の3種類を設定し，それぞれの論文で使われている検定の種類・標本数・自由度の情報をもとに検定力を算出した。さらに，論文の扱っている素材・研究方法別に検定力を算出した。その結果，先行研究に比べると検定力は高く，研究方法別にみると調査方法を用いた研究が実験や観察に基づく研究方法に比べ，検定力が高くなっていることが認められた。研究領域別では，障害児を取り扱った研究領域での検定力が低いことが明らかにされた。検定力を高めるためには標本数を多くすることが必要であるが，障害児を取り扱う分野のように標本を集めにくい分野においては標本数を増加させる事は至難の技である。こういった場合，無理に標本数を多くすることなく，既存の類似データの結果を踏まえたメタ分析による結果の統合が有効になることが少なくない，と著者は述べている。いずれにしても，教育心理学研究において検定力を分析した論文はこれまで皆無といってよい。その意味で，本論文は教育心理学研究・心理学研究投稿者に一石を投じた価値ある論文といえよう。

4. 質問紙調査による測定の諸問題

4.1 質問紙尺度の作成と利用に関するシンポジウム

まず1999年の日本教育心理学会第41回総会準備委員会企画のシンポジウムとしてとりあげられた質問紙尺度の作成と利用に関するシンポジウム（天根・吉田他，1999）の内容について紹介しよう。このシンポジウムは大変盛況で200人以上の参加があったことを明記しておく。話題提供者は桜井茂雄・辻平治郎・浦光博の3氏がそれぞれ質問紙研究との関わり，質問紙研究の利点，問題点などが報告された。つづいて指定討論者である南風原朝和・大平英樹・山田洋子・山岸みどり，の4氏が前記3氏の内容をふまえ討論を行った。これらの報告内容・討論内容を筆者の観点からまとめると次のようになる。

1）質問紙尺度の妥当性研究が貧弱である。質問紙尺度の作成が安易になり予備的研究，さらには作成された尺度の妥当性の測定が行われていない。
2）質問紙の回答は状況によって異なるため，質問紙の回答過程の分析がもっと必要である。
3）逆転項目の処理：ネガティブな表現の項目が使用しにくくなっている。ポジティブとネガティブはそれぞれ別の意味をもっている。
4）内省報告も重要な研究である。

NII-Electronic Library Service

Japanese Association of Educational Psychology

5）質問紙は何を測定しているか。意識・態度・性格，いずれも実体でない。
6）最近の調査研究の標本数は大きくなっている。このため，必要以上に有意差が生じやすくなっている。
7）質問紙は本当に日常的行動を測定しているか疑問である。
8）尺度構成のプロセス：概念同士の理論的関連を明らかにすべきである。
9）パーソナリティの基本次元の記述として，質問紙研究は有用であるが，パーソナリティの各特性が特有の行動を生じる理由については説明が十分にできていない。
10）質問紙研究は探索型の研究で仮説検証型の実験で味わえるような研究の醍醐味がすくない。

上記のシンポジウムで取り扱われた内容は広範であったが，質問紙尺度構成の方法論についてふみこんだ討論がなく，内容的には盛り上がりに欠けていたように思われた。ここで，上記のシンポジウムからは離れるが，以下に2つの点から質問紙尺度構成に関する方法論的話題に関連した研究発表をレビューしてみよう。

4.2 信頼性係数について

上記のシンポジウムにおける，質問紙の妥当性研究が貧弱であるという指摘はまさに正鵠を得たものであるが，それでは質問紙の尺度の信頼性の研究が十分であるかといえば，筆者の知る限りでは，必ずしもそうとは言えない。質問紙による尺度構成の手順としては，因子分析(回転を含む)により抽出された各因子に高い負荷量を有する項目をまとめて α 係数を算出するという手続きがかなり定着しているが，村石・豊田（1998）は，尺度に含まれる項目の τ 等価（項目得点が真値を共有している）や平行性（τ 等価に加えて，誤差分散も等しい）を無視して α 係数を求めることは，必ずしも好ましくないことを警告し，因子負荷量を用いた精密な計算式の利用を推奨している。一方，高木・柳井（1998）は α 係数と同時に θ 係数を求め，ある特定の尺度のなかに他の項目と比較してきわめて因子負荷の低い項目が含まれている場合には，θ 係数は α 係数に比べ0.2～0.4増加する事例を報告している。さらに，村上（繁桝他編（1999）のQ55）は尺度の信頼性としては，α 係数と同時に尺度に含めた項目の内容的広がりをチェックする1つの手段として再検査信頼性係数を求める事，さらには理論的に高い相関があることが予測される変数との相関を求めることにより構成概念妥当性を測定することの必要性を指摘している。これまでに，世界各国で開発されたキャリア・セルフエフィカシィ尺度をまとめた廣瀬(1998)のTABLE 1によると，33尺度のうち24尺度で α 係数，8尺度で再検査信頼性係数が記載されており，そのうち4尺度で，α 係数と再検査信頼性係数がともに記載されていたと報告されている。

4.3 順序尺度データの取扱い

尺度構成の基礎となる項目の反応の多くは「はい」「どちらともいえない」「いいえ」，あるいは，「賛成」「やや賛成」「やや反対」「反対」といった順序のついたカテゴリに，1，2，3および1，2，3，4といった得点を与え，それらの加算によって尺度得点を定義するリッカート法が利用されている。このように順序尺度を間隔尺度として取り扱う方法が無反省な数量化であるという指摘はこれまで再三なされてきている。この指摘の真偽を確かめるために，村上（1999）は同一データを代案として想定される数量化3類（あるいは多重対応分析）と非計量主成分分析を適用して比較検討し，その結果，調査項目を上記のように数量化して因子分析（または主成分分析）→バリマックス回転して尺度化することが，無反省な数量化ではないという結論を下している。ここでは詳しい説明は省略するが，その理由のひとつとして村上は因子分析と数量化の最適基準が，前者はr-join，後者はr-meet（Gifi（1990）参照）といったように，相互に異なることを指摘している。なお，村上は順序尺度のデータを数量化3類にかけた場合，実質的に1次元の解しか得られないことが多いと述べているが，この1次元解と主成分分析の第1主成分の関連がどのようになっているかを調べてみることは，興味ある研究課題であろう。

ところで，尺度得点を算出する場合，因子分析により同一の因子に属するとみなされた項目の得点を加算するため，尺度得点は項目の分散の大きさの影響をうける。したがって，分散1の標準得点により導かれた因子間相関と因子に対応して構成される尺度得点間の相関は，尺度に含める項目の分散が均一でない場合には，著しく異なることがある。村上・野上（1999），野上・村上（1999）はこの点に着目し，項目の尺度構成に際し，項目間の相関係数行列でなく，分散共分散行列を求めて因子分析を行うという方法がある程度有効性を持つものであることを示唆している。

5. いくつかの研究テーマの動向

本節では評価測定に関する教育心理学の領域のうち，性格測定と，大学入試データの解析，をとりあげる。

5.1 性格測定の研究

測定評価理論の応用は教育心理学・心理学のあらゆる領域に広がっているが，最近その発展が著しい分野に性

NII-Electronic Library Service

Japanese Association of Educational Psychology

格測定の領域がある。これをうけて，1998年度の第61回の日本心理学会で，日本性格心理学会との共同企画として「性格測定の到達点と課題」というシンポジウムが開催された(杉山・安藤他，1998)。話題提供者の柏木繁男氏(柏木，1999) は性格の5因子論 (FFM : Five Factor Model) の1990年代以降の著しい発展について紹介し，新性格検査(柳井他，1987) が5因子モデルで説明されること (柏木・和田，1996) だけでなく，精神医学の領域で広く利用されているTEG(東大式エゴグラム)が同様に5因子モデルで説明されることを立証している。つづいて，ロールシャッハ検査などによる性格測定に付随する量的分析法の限界を克服する1つの方法として，質的分析法の発展とその限界についての紹介が馬場 (1998) によってなされたことを紹介しておこう。性格5因子論を含む性格研究の成果についての出版は，柏木 (1997) に続き，詫摩 (監) (1998)，辻他 (1999) と続いており，活況を呈している。なお，こういった性格の測定はもちろんのこと，知的能力，興味・意欲の測定を含めた人事評価に関する諸領域を展望した書物 (大沢・芝・二村，2000) が本年の1月に出版されている。

5.2 入試データの解析

入学試験のデータ解析に関連した研究はデータが一般の研究者が入手しにくいため，教育心理学会で発表されることは少ない。共通第1次学力試験が開始された2年後の1980年から毎年1度国立大学入学者選抜研究協議会(入研協)の総合が開催され，入学試験に関連した研究発表が行われている。この会合は，各大学の入学者選抜研究協議会の委員のみしか参加できないクローズドな会合であるが，そこで発表された内容の一部は「大学入試研究ジャーナル」(年1回発行される。平成11年3月に第9号が発行された) という雑誌で公開されている。

ところで，大学入試データ解析に関する研究は1950年代から行われていたが，共通第1次試験が開始された1979年以降増加している。筆者自身は1995年4月から1998年3月にかけて科学研究費補助金により「多変量解析の利用による大学入試データ解析のシステムの開発」という研究を行い，1998年3月にその報告書(柳井，1998) をまとめた。このうち，千野・池田・平野・垂水・岡太・山田 (1998) (題目は巻末の文献リスト参照) の5氏は第25回日本行動計量学会の特別セッション「大学入試データの解析をめぐって」において，それぞれ「入試形態と就職先との関連」「CBTの問題例」「医師国家試験の合格率」「合否入れ替わり率の計算ソフト」「大学進学における都道府県の関連」を発表している (これらをまとめて，柳井・前川 (1999) が刊行した)。さらに，本年 (1999年) の教育心理学総会における自主シンポジウム (柳井・倉元他，1999) において「大学入試の教育心理学」が行われ，楠見孝・馬場久志・倉元直樹・池田央の4氏が講演，村上隆・荒井克弘の2氏が討論で，それぞれ大学生の学力低下が懸念される最近の状況をふまえながら，大学入試に果たす教育心理学の役割について熱心な討論を行った。平成11年度からは国立3大学にアドミッション・オフィスが設置され，情報公開法の制定により平成13年度からは大学が入試に関する情報を可能な限り，受験生に公開することが必至となった状況下において，教育心理学研究者，および教育心理学会においては今後もっと積極的に，大学入試に関する研究発表を行っていく必要があろう。21世紀の入試はどうあるべきかについて展望した書物に吉田編 (1998) がある。

次に，大学入試データ解析に最近頻繁に利用されている共分散比 (Covariance ratio) について解説しておこう。

大学入試センター試験の合計点をX，各大学が行う個別試験の合計点をY，としよう。このとき，センター試験と個別試験は合計点にどの程度寄与しているかを分析したい場合がすくなくない。このとき，次の関係が成立する。

$$(X+Y)\text{の分散}=((X+Y)\text{とXの共分散})+((X+Y)\text{とYの共分散})$$

ここで，$((X+Y)\text{とXの共分散})/(X+Y)$の分散，をXの共分散比，$((X+Y)\text{とYの共分散})/(X+Y)$の分散，をYの共分散比とよぶ。その和が100%となることは明らかであろう。共分散比は，複数個の試験成績の合計点に対するそれぞれの試験成績の寄与の割合，といった場合にも適用可能となる。TABLE 1の共通1次試験データの場合，国語・社会・英語・理科・数学の共分散比はそれぞれ，12.5，16.14，27.13，25.27，18.96となる。この結果，配点が同一の200点満点にもかかわらず，国語の合計点に対する寄与の程度が数学または理科の寄与の程度の半分以下となっていることがわかる。なお，性格検査や態度検査の尺度を作成するために複数個の項目の合計点を尺度得点とした場合，合計点に対する各項目の寄与の程度を共分散比を用いて表わすことが可能となる。

6. 評価・測定に関する国際的研究活動

評価・測定の分野に関する我が国の研究者による国際的研究活動を示すものとして，国際計量心理学会 (Psychometric Society) の発行するサイコメトリカ (Psychometrika) 誌への日本在住の研究者による寄稿状況の概要について触れてみよう。日本人として最初にサイコメトリカに掲載されたものはジェオマックス因子回転法に関する柏木繁男氏の論文 (Kashiwagi, 1965) である。その後、1960年代から70年代にかけてアメリカ大陸に

NII-Electronic Library Service

Japanese Association of Educational Psychology

渡った鮫島史子氏，西里静彦氏，高根芳雄氏がサイコメトリカ誌に数多くの論文を発表し，高根氏は1986年に，西里氏は1996年に，鮫島氏は1997年に名誉ある国際計量心理学会の会長に就任された。こういった3人の活躍に刺激され，上記3名以外で日本に職を持つ研究者が第一著者となったサイコメトリカへの掲載論文は1978年に1編，1980—84年代に2編，85—89年代に5編，90—94年代に4編，95—98年代に3編の計15編あることを明記しておこう。15編の内容は因子分析が9編，(最小二乗法のアルゴリズム，感度分析，SMCの改良，イメージ因子分析，独自性の推定法，独自因子の漸近分散，ブートストラップ法の適用，識別性の問題，AICの適用，3相因子分析の核行列の問題) で最も多く，続いて，多次元尺度構成法2編，最適化法2編，ラッシュモデル1編であった。この他，外国に住む日本人研究者との共著論文の数も1980年代に2編，90年代に3編あることを明記しておこう。

こういったサイコメトリカにおける日本人の研究業績が評価され，2001年の7月15日から20日にかけて，日本で初めての国際計量心理会議 (International Meeting of Psychometric Society) が大阪大学 convention center を会場として開催されることになったことを明記しておこう (詳細については，http://www.ir.rikkyo.ac.jp/imps2001/を参照のこと)。

7. 測定評価に関するコンサルティング活動

前節までに示したように教育心理学の領域では，記述統計・統計的推測の方法・因子分析・共分散構造分析をはじめとする多変量解析の種々の方法が頻繁に使われているが，これらの方法を実際のデータへ適用するにあたり，様々な疑問を抱いている研究者が少なくない。こういった疑問を少しでも解消するために，平成6・7・8年の教育心理学総会において自主企画シンポジウム「教育心理学における統計的方法 DOs and DON'Ts が開催された。ここで取り上げられた質問項目とその回答に加筆されたものをまとめ，平成11年6月に「Q&Aで知る統計データ解析」(繁桝・柳井・森編) という形で出版された。この書物には，記述統計・統計的検定・相関と回帰・因子分析と共分散構造分析・多変量解析・その他の統計的問題と6章にわかれ，全部で115の質問と回答が与えられている。内容的には平易なものから，やや難解なものまでバラエティに富んでいるが，研究者が抱いている疑問のいくつかは本書により回答があたえられよう。ただし，内容がどちらかといえば統計的説明に重点がおかれているため，現実に行われている調査や実験から得られるデータの解析において生じる様々な問題に十分な回答を与えていないという印象をあたえるかもしれない。上記の書物は測定・評価の疑問に専門家が書物を通して回答するというコンサルティング活動とみなすことができる。こういった活動は，1998年の日本心理学会における共分散構造分析の講習会 (講師：豊田秀樹) があり，予定数100名を短時間にオーバーする盛況であった。また日本行動計量学会の主催した計量的手法 (統計学および多変量解析の各種技法) に関する合宿セミナーが1998年3月 (会場：東大検見川グランド施設：企画責任者：繁桝算男)，1999年3月(会場：大阪大学人間学部：企画責任者：狩野裕)に開催され，それぞれ200名近い参加者があった。大阪のセミナーでは，400ページを越す講義録(狩野，1999)が作製されている。これらの講習会・セミナーは，計量的手法を利用したい若手の研究者に対し，計量的手法を専門とする研究者が比較的楽しい雰囲気で講義を進めることにより，成功裡に会をすすめることができたようである。1998年の日本心理学会で行われた日本行動計量学会・日本基礎心理学会共同企画として行われた「心理学方法論の基礎にある数理と計量」というシンポジウム (繁桝，佐藤他，1998) も測定・評価の基礎的考え方に関する啓蒙的シンポジウムであったといえよう。教育心理学会においても，1999年度の総会においては「因果関係の基礎」に関する講習会 (講師：坂元章) があったが，総会の時期に限らず年に2—3回、東京または大阪近辺の大学で測定評価に関する講習会が開催されることが必要であろう。近い将来，教育心理学会総会，または総会とは別の機会において，測定評価に関連した講習会，例えば「因子分析の基礎・応用」「因子分析から共分散構造分析へ」「質問紙尺度の作成をめぐって」等を開催して頂く事を要望したい。

8. 統計的手法の誤用とその改善をめぐって

付言すれば，教育心理学に掲載された論文であっても，統計的手法に関連した誤りが見られる論文は少なくない。因子分析の主因子法，または主成分分析の結果をバリマックス回転して得られた因子負荷量の平方和を固有値と記述してある論文が最近1年間に発行された教育心理学研究の4論文と心理学研究の1論文に見られた。これは明白な誤りであるが，この誤りは，ここ10年来跡を絶たない。すでに述べたように，明らかに直交回転でなく斜交回転を適用すべき場合にはそのように審査員が指示すべきである。さらに，因子数をもうひとつ多く抽出した方がよい，という指摘を与えることにより，興味ある結果が導かれる場合がある。

このような誤りは因子分析に限定されるものではない。心理学研究第67巻 (1996.4〜1997.2) に掲載され，分散分析を行っている32論文を点検して，分散分析における被験者間要因と被験者内要因の扱いに関して生じる誤差項の

NII-Electronic Library Service

Japanese Association of Educational Psychology

取り扱いについて多くの論文で明白な誤りを侵しているという指摘(川野, 1998)がある。さらに，測定・評価に関連する問題の中には，明らかな誤りではないが，改善が望まれる問題が少なくない。その一例を挙げよう。教育心理学研究の実践研究において，あるクラスを実験群，別のクラスを統制群として測定値に関して平均の差を検定することがある。この場合，観測値の独立性が問題となる。この点に関し，1999年度の教育心理学研究総会で行われた「教室というフィールドにおけるデータの収集と解析に関する諸問題(企画者：相川充・小林正幸(東京学芸大))」というシンポジウムはタイムリーな企画といえよう。このシンポジウムにおいて，栗田(1999)は，学級の背後に想定される母集団は独立な要素からなる無限母集団ではなく，要素が学級単位でまとまるような学級の無限集合からなる母集団を設定すべきで，その場合学級のまとまりの指標として級内相関係数の利用を提唱している。

教育心理学研究に掲載された論文のほとんどは何らかの統計的手法を使用しているが，上記にのべたように様々な問題点，あるいは明らかな誤りとみられる部分を残したままで掲載されている論文が少なくない。多くの優れた統計プログラムパッケージの普及によって，安易な統計的手法の適用例が増加し，それに比して誤用が増えているといっても過言ではない。こういった点に鑑み，1つの提案をしたい。すなわち，教育心理学研究の論文審査員を選ぶにあたって，少なくとも1人測定評価の専門家を含めること。もしそれが可能でない場合には，統計的手法の明らかな誤用の有無について，編集委員会が1，2名の専門家を指名して，チェックしてもらう。もちろん，編集委員長，さらには複数名の測定評価の専門家が編集委員である場合は，それらの1人がこの任をおえばよい。そうでない場合に，編集委員長が外部の適任者に依頼することが必要となろう。さらに，こういった誤りを論文から容易に検出できるように，投稿者に，被験者数・平均値・標準偏差を論文中に明記させること(川野, 1998)，因子分析の場合であれば，相関係数行列なども変数の個数が多すぎる場合を除いて，可能な限り論文中に掲載するように義務づけることが必要であるかもしれない。そうすれば審査委員，あるいは編集委員会が委託した専門家が，それを用いて再計算を行うことができるであろう。

(謝辞) 本文の全文を読んで，有益なコメントを頂いた前田忠彦(統計数理研究所)およびTABLE 1の作成にあたり資料の収集などでご協力頂いた山住賢司(同志社大学)の両氏に深謝する。

引用文献

安達智子 1998 セールス職者の職務満足感—共分散構造分析を用いた因果モデルの検討 心研，**69**，3，223-228.

相川 充・小林正幸 1999 研究委員会企画シンポジウム 教室というフィールドにおけるデータの収集と解析に関する諸問題 日教心第41回総会発表論文集，20-21.

天貝由美子 1999 一般高校生と非行少年の信頼感に影響を及ぼす経験要因 教心研，**47**，229-238.

天根哲治・吉田寿夫他 1999 準備委員会企画シンポジウムI 教育心理学研究における質問紙尺度の作成および利用に関する諸問題 日教心第41回総会発表論文集，2-3.

Archer, C.O., & Jennrich, R.I. 1973 Standard Errors for rotated factor loadings, *Psychometrika*, **38**, 4, 581-591.

馬場禮子 1998 ロールシャッハ法による人間—状況関係の把握 日心第62回大会発表論文集，S19.

千野直仁 1998 入試形態，就職先と入学後の成績との関連について 日本行動計量学会第26回大会発表論文抄録集，315-318.

Gifi 1990 *Nonlinear Multivariate Analysis,* Wiley, New York.

Greenland, S., Scllesselman, J.J., & Criqui, M.H. 1986 The fallacy of employing standardized regression coefficients and correlations as measures of effect, *American J. of Epidemiplogy,* **123** (2), 203-208.

Greenland. S, Maclure. M., & Scllesselman 1991 Standardized regression coefficients: A further critique and review of some alternatives. *Epidemiplogy,* **2** (5), 387-393.

服部 環 1998 一般化部分採点モデルの母数等化—最小x自乗法と特性曲線法— 日心第62回大会発表論文集，417.

平野光昭 1998 医師国家試験の合格率を高めるための総合卒業試験等のデータ解析システム 日本行動計量学会第26回大会発表論文抄録集，319-322.

廣瀬英子 1998 進路に関する自己効力研究の発展と課題 教心研，**46**，343-355.

藤森 進 1999 項目反応理論の適用に際しての項目分析の意義 日教心第41回総会発表論文集，289.

兵藤恵子・森野礼一 1999 阪神・淡路大震災による精神的身体的影響に関する調査—女子大学生における地震直後，2ヶ月後，9ヶ月後の状態— 心研，**70**，2，

NII-Electronic Library Service

Japanese Association of Educational Psychology

104-111.
市川伸一・大橋靖雄他　1987　SASによるデータ解析入門（第2版）東京大学出版会
池田　央　1998　CBTの問題例と作問形式のタクソノミー　日本行動計量学会第26回大会発表論文抄録集，307-308.
岩男卓実　1999　カテゴリに基づく帰納推論における専門性の影響　教心研，**47**，59-67.
Kashiwagi S. 1965 Geometric vector orthogonal rotation method in multiple factor analysis, *Psychometrika*, **30**, 515-530.
狩野　裕　1997　グラフィカル多変量解析　目で見る共分散構造分析　現代数学社
狩野　裕(編)　1999　日本行動計量学会第2回「春の合宿セミナー」資料集　日本行動計量学会
柏木恵子・永久ひさ子　1999　女性における子どもの価値　教心研，**47**，170-179.
柏木繁男・和田さゆり　1996　5因子モデル（FFM）による内田クレペリンテストの評価について　心研，**66**，1，24-32.
柏木繁男　1997　性格の表現と理論―特性5因子論からのアプローチ　有斐閣
柏木繁男　1999　性格特性5因子論（FFM）による東大式エゴグラム（TEG）の評価　心研，**69**，6，468-477.
川野卓二　1998　「心理学研究」における分散分析法(1)　日心第62回大会発表論文集，426.
栗田佳代子　1999　学級を単位としたデータを収集した標本と母集団の関係について　日教心会第41回総会発表論文集，20-21.および配布資料
喜岡恵子・柳井晴夫　1998　重回帰分析における偏回帰係数と標準偏回帰係数をめぐって　人事試験研究，**163**，13-20.
村石幸正・豊田秀樹　1998　古典的テスト理論と遺伝因子分析モデルによる標準学力検査の分析　教心研，**46**，395-402.
村上　隆・野上康子　1999　心理学的個人差測定尺度構成のための主成分分析の使用について(1)　日教心第41回総会発表論文集，557.
村上　隆　1999　小講演12　カテゴリ反応の因子分析―Likert尺度は無反省は数量化か　日教心第41回総会発表論文集，93.
Muraki, E. 1992 A generalized partial credit model, Application of EM algorithm, *Applied Psychological Measurement*, **16**, 159-176.
中村知靖・前川眞一・菊地賢一　1998　テスト得点分布を考慮した項目選択法　日心第62回大会発表論文集，419.
Newman, T.B., & Browner, W.S. 1991 In defense of standardized regression coefficicients, *Epidemiology*, **2**(5), 383-386.
野上康子・村上　隆　1999　心理学的個人差測定尺度構成のための主成分分析の使用について(2)　日教心第41回総会発表論文集，558.
Ogasawara, H. 1998 Standard errors for rotation matrices with an application to the promax solution. *British Journal of Mathematical and Statistical Psychology*, **51**, 163-178.
Ogasawara, H. 1999 Standard errors for procrustes solutions. *Japanese Psychological Research*, **41**(2), 121-130.
Ogasawara, H. 1999 Standard errors for matrix correlations. *Multivariate Behavioral Research*, **34**(1), 103-122.
小笠原春彦　1999　IRTにおけるロジット交換を利用した等化と等化数の標準誤差　日教心第41回総会発表論文集，202.
岡太彬訓　1998　大学進学における都道府県間の関連―東京と京都を中心とする2大大学進学圏間の非対称性の変化―　日本行動計量学会第26回大会発表論文抄録集，309-310.
岡田佳子　1999　中学生の学校ストレスに関する研究(1)―学校ストレッサーとストレス反応について―　日教心第41回総合発表論文集，262.
大沢武志，芝祐順，二村英幸編著　2000　人事アセスメントハンドブック　金子書房
Samejima, F 1969 Estimation of latent ability using a response pattern of graded scores, *Psychometric Monograph*, 17.
繁桝算男・佐藤隆夫他　1998　日本行動計量学会・日本基礎心理学会共同企画シンポジウム　心理学方法論の基礎にある数理と計量　日心第62回大会発表論文集，S7-S10.
繁桝算男・柳井晴夫・森　敏明(編著)　1999　Q&Aで知る統計データ解析　心理学セミナーテキストライブラリー3　サイエンス社
冷水來生・冷水啓子　1998　日本手話の動作成分に内在する意味について　教心研，**46**，306-316.
杉澤武俊　1999　教育心理学研究における統計的検定の検定力　教心研，**47**，150-159.
杉山憲司・安藤寿康他　1999　日本性格心理学会共同企画シンポジウム　性格測定の到達点と課題　日心第62

NII-Electronic Library Service

Japanese Association of Educational Psychology

回大会発表論文集，S16-S19.
高木廣文・柳井晴夫 1998 生活習慣尺度の信頼性と因子構造の検討 統計数理，**46**，1，39-64.
詫摩武俊監修 1998 性格心理学ハンドブック 福村出版
平 直樹・前川眞一 1998 日本語基礎能力テストの簡易得点化法 日心第62回大会発表論文集，425.
平 直樹・前川眞一他 1998 日本語基礎能力テストの項目プールの作成 大学入試センター研究紀要，28.
田中美由紀・山崎瑞紀・柳井晴夫・鈴木規夫 1998 個人の資質と大学の専門分野での適応に関する因果モデルの検討 教心研，**46**，262-270.
垂水共之・山本義郎 1998 合否入替わり率解析ソフトウェアについて 日本行動計量学会第26回大会発表論文抄録集，303-306.
豊田秀樹 1997 共分散構造分析における行動遺伝モデルの新展開 心研，**67**，464-477.
豊田秀樹(編) 1998 共分散構造分析「事例編」—構造方程式モデリング— 北大路書房
豊田秀樹 1998 共分散構造分析入門—構造方程式モデル 朝倉書店
豊田秀樹・村石幸正 1998 双生児と一般児による遺伝因子分析—Y-G性格検査への適用— 教心研，**46**，255-261.
辻 平治郎(編) 1998 5因子性格検査の理論と実際 北大路書房
上田尚一 1998 高校／大学進学マトリックスの分析 日本行動計量学会第26回大会発表論文抄録集，311-314.
渡辺直登，野口裕之編著 1999 組織心理測定論 項目反応理論のフロンティア 白桃書房
山本嘉一郎・小野寺孝義(編著) 1999 Amosによる共分散構造分析と解析事例 ナカニシヤ出版
柳井晴夫・柏木繁男・国生理枝子 1987 プロマックス回転法による新性格検査の作成について(1) 心理学研究，**58**，158-165.
Yanai, H., & Mukherjee, B.N. 1987 A generalized method of image analysis from an intercorrelation matrix which way be singular, *Psychometrika,* **52**, 555-564.
柳井晴夫他 1990 因子分析—その理論と方法 朝倉書店
柳井晴夫(編) 1998 多変量データ解析の利用による大学入試データ解析システムの開発 平成7—9年度科学研究費補助金基盤研究(A)報告書
柳井晴夫・倉元直樹他 1999 自主シンポジウム 10 大学入試の教育心理学—学校教育の近未来を考える— 日教心第41回総会発表論文集，42-43.
柳井晴夫，前川眞一編著 1999 大学入試データの解析—理論と応用— 現代数学社
山田文康 1998 入試制度の評価—静岡大学情報学部について— 日本行動計量学会第26回大会発表論文抄録集，323-324.
吉田辰雄(編) 1998 どうする入試改革 21世紀にむけた入試改革の動向 文化書房博文社
吉村 聡・山住賢司・柳井晴夫 1998 心理学における因子分析法—「心理学研究」と「教育心理学研究」に見る利用頻度の変遷— 早稲田心理学年報，**30**，1，63-67.

NII-Electronic Library Service