個人差測定尺度のプロトタイプ理論と主成分のプロトタイプ変換

村　上　　隆

本研究は，複数の質問項目への応答を通じて，対象者の個人差を量的に測定するという方法論について論じたものである。

いわゆるLikert型の質問項目に対する反応を単純加算することによって定義される尺度は，心理学的な検査（inventory）や態度尺度（attitudes scale）にしばしば用いられる。その際，項目の選択や，複数の下位尺度への分類には，古典的テスト理論や，探索的，あるいは確認的な因子分析の手法が用いられる。しかしながら，真の得点または因子と，独自因子やランダム誤差を截然と区別するこれらの理論の分析枠組みは，項目水準のデータには，ほとんど場合十分にフィットしない。

本研究は，項目水準のデータによりふさわしい枠組みとして，項目反応を規定する相互に独立な潜在的要因が存在し，個々の項目反応にはそのうちの幾つかが関わるとする家族的類似（family resemblance）の理論（村上，1989，1993）をもとに，研究者は想定した下位尺度の中心となるプロトタイプ項目を指定できるという特徴を付加した，尺度構成のプロトタイプ理論を提案する。さらに，その理論にもとづいて主成分分析の解を変換する方法であるプロトタイプ変換と，それにもとづいた下位尺度の構成方法を示す。最後に，プロトタイプ変換の実データへの適用と，これらの手続きの適用に関する若干の議論がなされる。

1　質問紙法による個人差測定尺度

1.1　Likert 型の尺度

複数のステートメントを多数の調査参加者に提示し，個々の参加者による「賛成－反対」，あるいは，「当てはまる－当てはまらない」といった反応を，3～7段階の数値としてコード化するような質問項目は，その創始者の名前をとってLikertタイプの項目と呼ばれる。たとえば，表1はHill（1987）による親和動機（affiliation motivation）の尺度 IOS（interpersonal orientation scale）のすべての質問項目の日本語版（中村，2000）であり，これには，「全く当てはまらない」(1）から，「完全に当てはまる」(5）までの5段階で反応することが求められる。

一連のLikert型の質問項目に対する反応は，それらを合計して，個人ごとに（しばしば複数の）得点を求める形で使われることが多い。一組の質問項目と得点の計算方法をあわせたものは，尺度（scale）と呼ばれ，ことに心理学の領域では，これまでにおびただしい数の尺度が開発されてきた（たとえば，堀・山本，2001 など)。尺度得点と他の変数との相関をとったり，対象者の属する群の間で尺度得点の平均値を比較したりすることは，心理学のみならず，社会学や政治学における実証的研究における常套手段となっている。

1.2　尺度と下位尺度

その場合，1つの尺度を構成する項目群は，さらに複数の部分集合（subsets）に分けられ，それぞれの部分集合に含まれる項目ごとに加算されることも多い。これによって，単一の尺度から，複数の尺度得点が得られることになる。実際，表1のIOSでは，「ポジティブな刺激（Positive stimulation; P)」，「情緒的支持（Emotional support; E)」，「注意（Attention; A)」，「社会的比較（Social comparison; S)」という4つの

表1　IOS（日本語版）の質問項目（下位尺度別）

P	3	他の人のそばにいるのが好きなのは，おもにふれあいから暖かい幸福感を得ることが出来るからである
	6	自分は多くの人が認める以上に，他人とのふれあいから満足を得ていると思う。
	10	他の人のそばにいてその人を理解することが，自分にとって最も興味深いことの一つである。
	11	他の人と一緒にいることから，私は多くの人が得る以上の満足を得ているように思う。
	13	誰かと親しくなれたときには，自分はなにか価値のあることを成し遂げたと感じる。
	20	かなり大勢の人と非常に親密な友情を結ぶことが出来るならば，自分は満足するだろうと思う。
	24	他の人と親しくなり一対一で話せることは，私にとって最も満足度が高く気に入った気晴らしの一つであると思う。
	25	好きな人となら誰であれ新しく友達になれば，とても満足に思うだろう。
	26	私が考え得る最高の楽しみの一つは，人々を観察してどんな人なのかを知ることである。
E	1	つらい時の最大の慰めの一つは，他の人と一緒にいることである
	4	なにかよくないことが起きたり混乱したときは，信頼できる親友と一緒にいたいと願うことが多い。
	9	自分にとって非常に重要なことをうまくはやれなかったとき，他の人のそばにいるだけで気持ちを持ち直すことが出来る。
	15	なにか苦しいことを経験しなければならないとき，その間中誰かがそばにいてくれると，苦しさが少なくなるといつも思う。
	17	悲しみに沈んだり，落ち込んだ気分になったならば，その気分を良くするために他の人のそばにいようといつも努める。
	23	なにかで気持ちが動揺したときには，そばに誰かがいてほしいと非常に強く願うことが多い。
A	5	私に強く好意をもち，夢中になってくれているような人を大概私も好きになる。
	8	自分が注目の的となりそうなときは，他の人のそばにいたい。
	16	私の人となりや私のすることを高く評価してくれる人たちのそばにいたいと強く思うことが多い。
	19	大概は，私のことを重要で刺激的な人間だと考えてくれる人のそばにいたい。
	21	私に注目し，私の人となりを理解してくれる人がそばにいてほしいと強く願うことが多い。
	22	私のことを十分にほめてくれないような人とは一緒にいたくない。
S	2	何か活動をするときはそのより方を知りたいので，一人よりは他の人と一緒に参加したい
	7	自分がどの位うまくやれているのか確信がないときは，比べることが出来るので他の人のそばにいたいと強く思う。
	12	仕事の場面や人と接する際などに自分に期待されていることがはっきりしないときは，その手がかりを得るために誰が他の人を当てにすることが出来るようにしたいといつも思う。
	14	この先どうなるのかよくわからないときは，自分と同じ経験をしている人のそばにいたいと願うことが多い。
	18	自分が人に比べてどの位できているかを知るために，他の人に注目することが多い。

得点が得られる。このような部分集合は下位尺度（subscale）と呼ばれることも多い。

下位尺度にまとめられるのは，相互に質問内容の似た，相互に相関の高い一群の項目である。質問内容が似ているというのは，項目作成時に研究者がもっているいわばアプリオリな想定であり，相関が高いというのは得られたデータの分析を通じて明らかになる事実である。これらは，必ずしも相互に一致するとは限らない。それでは，こうした手続きの根拠，あるいは原理は何なのであろうか。

1.3 尺度構成の基本原理

第1に，多くの質問項目を単独で扱っていては，結果が極めて煩瑣なものとなるため，変数を減らして整理するということがあげられる。この場合，相互相関の高い項目を加算することは，変数の減少よる情報のロスを最小限にとどめることになる。これは，統計的にはデータ縮約（data reduction）に相当する。

第2に，複数の項目を加算することによって，測定の信頼性が上昇するということが考えられる。つまり，個々の項目への反応は，さまざまな要因の影響により不安定であるが，複数の項目反応を加算すれば，誤差の影響は打ち消しあって相対的に（下位）尺度得点は安定したものになるという考え方である。すなわち，測定の信頼性（reliability of measurement）という観点である。

第3に，もう少し踏み込んで，各下位尺度は個々の質問項目の内容を包括する一種の上位概念に対応していると見ることもできる。すなわち，下位尺度によって，研究者は，個々の項目の水準を越えた一般的な概念を定義することになる。こうして，質問項目を内容，あるいは相互相関にもとづいて分類することには，概念形成（concept formation）という意味をもつ可能性があることになる。測定論の観点からすれば，この原理は測定の妥当性（validity of measurement）と関わる。

現実の尺度構成の場面で問題がおこるのは，この3つの目的が，必ずしも完全に同じ規準の最適化として実現されるわけではないからである。すなわち，第1のデータ縮約は，データの効率的な要約を目指すから，少数の次元で，項目反応の分散の多くの部分を説明することが目標となる。ここでの規準は尺度によって説明される分散の大きさ，あるいは割合である。第2の尺度の信頼性の観点では，個々の尺度の内的整合性が問題にされる。代表的な指標は後述のCronbachのα係数である。以上の2つの観点では，基本的に項目間相関の高さが問題になる。一方，第3の概念形成は，測定の内容が問題になるから，尺度それ自体についての明確な基準は存在しない。あえて数値的な規準をあげれば，尺度外の変数との相関や尺度値の群間の平均差が理論的な予測と整合していることであろう。しかし，むしろこの原理では，質問内容にもとづく尺度の解釈可能性が重要であり，項目間の意味的類似が問題になる。

本研究は，基本的にはこの枠組みを念頭におきつつ，尺度構成のための適切なデータ分析の方法について論じようとするものであるが，問題の所在をより明確にするために，従来の考え方の数理的基礎について，次の2つの節で簡単に要約しておこう。

2　個人差測定のモデル

2.1　測定の基本モデルと信頼性係数

古典的テスト理論（classical test theory）は，項目反応の加算という形で定義される尺度得点の信頼性に関する数理的枠組みを与えた。Likert型項目反応の和で定義される尺度得点は，テスト理論とは別の態度測定（attitudes measurement）の理論として生まれたものであるが，背景となる理論もデータ分析上の手続きも，現時点では古典的テスト理論とほとんど融合している。したがって，ここでも古典的テスト理論の定式化から出発しよう。なお，以下の議論は，主に池田（1973）に基づく。

テスト得点の基本モデルと呼ばれるものは次の形をとる。

$$X_i = T_i + E_i \quad i = 1, ..., n \tag{1}$$

ここで，X_i は，個人 i の実測得点，T_i は個体 i の真の得点，E_i はランダム誤差を示す。実際の調査によって得られるのは，左辺の実測得点だけであり，右辺の2項はその存在と性質が仮定されているにすぎず，現実に値を求めることはできない。しかしながら，真の得点と誤差の間の共分散の期待値が0であるという仮定，

$$\mathrm{Cov}(X, T) = 0 \tag{2}$$

から，次の式が導かれる。

$$\sigma_X^2 = \sigma_T^2 + \sigma_E^2 \tag{3}$$

ここで，$\sigma_X^2, \sigma_T^2, \sigma_E^2$ は，それぞれ，X, T, E の分散（の期待値）である。測定の信頼性は，測定値の変動中に占める誤差の小ささ，あるいは，真の得点の分散の大きさを指し，それを数値的に表現する信頼性係数 ρ は，次のように定義される。

$$\rho \equiv 1 - \frac{\sigma_E^2}{\sigma_X^2} = \frac{\sigma_T^2}{\sigma_X^2} \tag{4}$$

以上は，あくまでも理念的なモデルに過ぎず，現実の尺度得点の信頼性を求める方法としては機能しない。信頼性係数の値を実際のデータから推測するには，項目水準での議論が必要である。

2.2　項目水準のモデルと内部一貫性

項目反応と尺度得点の関係については，個人 i の項目 j への反応（項目得点）を x_{ij} とするとき，これに対しても，基本モデル，

$$x_{ij} = t_{ij} + e_{ij} \quad i = 1, ..., n \quad j = 1, ..., p \tag{5}$$

が成り立ち，さらに，真の得点とランダム誤差の間の共分散と，異なる項目のランダム誤差間の共分散が0であるという仮定が成り立つとし，その上に，全ての項目の真の得点が相互に等しく，ランダム誤差の分散も全項目を通して一定であるとすると，σ_T^2 と σ_E^2 は，それぞれ次のように書くことができる。

$$\sigma_T^2 = p^2\sigma_t^2 \tag{6}$$

$$\sigma_E^2 = p\sigma_e^2 \tag{7}$$

ここで，σ_t^2, σ_e^2 は，それぞれ，項目の真の得点とランダム誤差の分散，p は項目数である。ここで，尺度（テスト）得点項目反応の合計点として定義される尺度（テスト）得点について，真の得点の分散も，ランダム誤差の分散もともに項目数の増加関数であるが，ランダム誤差の分散 σ_E^2 が，単に項目数に比例して増加するにすぎないのに対し，真の得点の分散 σ_T^2 は，項目数の2乗に比例して増加する。このことは，（下位）尺度得点を定義する項目が多いほど，（下位）尺度の実測得点の分散 $\sigma_X^2 (= \sigma_T^2 + \sigma_E^2)$ の中に占めるランダム誤差の分散の割合が相対的に縮小して信頼性が上昇することを意味している。

ここで仮定した，全ての項目の真の得点が相互に等しく，ランダム誤差の分散が全項目を通じて等しいという条件の下では，項目間の相関係数が全て一定となる。その値を r とすると，次が成り立つ。

$$r = \frac{\sigma_t^2}{\sigma_x^2} = 1 - \frac{\sigma_e^2}{\sigma_x^2} \tag{8}$$

ここで，σ_x^2 は項目得点の分散である（これも，上記の仮定の下では，全項目を通じて一定値となる）。式（3），（4）に，（6），（7）を代入し，（8）を用いて次の Spearman-Brown の公式が導かれる。

$$\rho = \frac{pr}{1+(p-1)r} \tag{9}$$

2.3 Cronbach の α 係数と項目分析

Spearman-Brown の公式の導出のために設定された，項目反応に関する仮定は強すぎて，現実のデータでは成り立たない。たとえば，項目間相関は一定値にはならないが，これを平均値 $\bar{r}$ で置き換えた次の係数は，内部一貫性にもとづく信頼性係数の推定値として知られるアルファ係数である。

$$\alpha = \frac{p\bar{r}}{1+(p-1)\bar{r}} \tag{10}$$

これは，図1に見られるように，p と $\bar{r}$ の単調増加関数である。こうして，尺度得点の信頼性を高めるためには，

- 項目数を増やす
- 相互相関の高い（内容的によく似た）項目を用いる

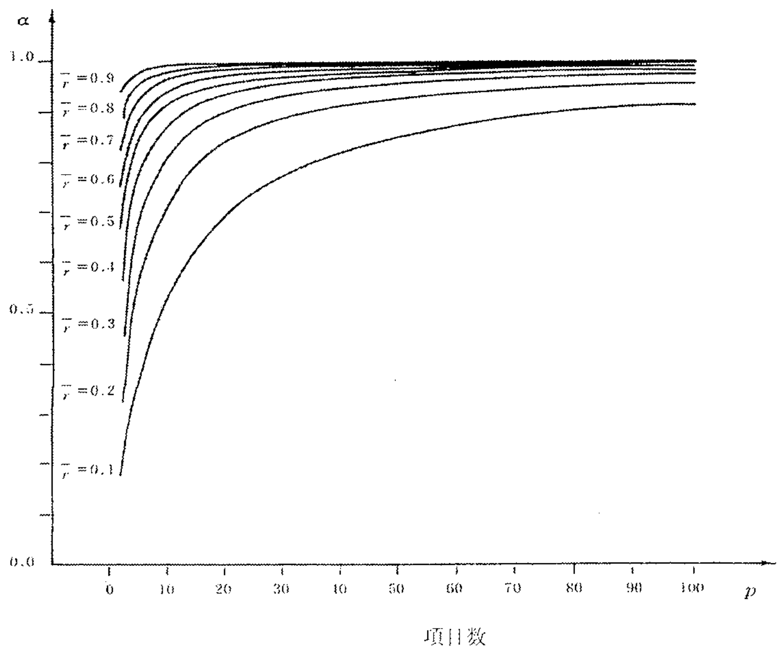

図1 項目数の関数としての α 係数。項目間相関の平均値ごとに示す。

という2つの方法があることが明らかになる。項目水準の議論により，実際のデータと尺度得点の信頼性との関係が明らかになった。

もっとも，(10) 式は，項目得点をすべて標準化してから加算して尺度（テスト）得点とするというやや特殊なやり方を前提としており，素点のままで加算する通常の尺度（テスト）得点に対応する α 係数とは異なる（SPSS 等で normalized alpha として出力されるものである）。しかし，実際にはこの2つの係数の値が，著しく異なることは少ないようである。

こうした理論的前提の下で，尺度の改善のために行われることは，α 係数の低下を招いているような項目，すなわち他の項目との相関が低く，$\bar{r}$ を低下させているような削除することである。そうした項目は，誤差分散の割合が大きく，尺度得点に多くのランダム誤差を持ち込むことになると解されるからである。実際には，その項目は他の項目とは異なる内容をもち，本来測定を目指した概念の測定のために有用ではあるが，意味的に類似した項目がその尺度に含まれていないために相関が低い可能性もあるのであるが，そうした可能性は古典的テスト理論では考慮されていない。

以上の議論において，今後，本研究で検討しようとする問題が，多少とも明らかになったように思われる。

(1) 古典的テスト理論においては，測定誤差は真の得点と截然と区別できるものとして想定されている。現実には，真の得点と誤差は交じり合っている可能性がある。

(2) もし，ランダム誤差を完全に取り除くことができたとしても，テスト得点（尺度得点）が，研究者が意図した次元と正しく対応しているかどうかは，全く不明である。真の得点という言葉に惑わされがちであるが，T_i の示しているものは，当該の尺度のために選択された項目に依存して決まる尺度固有の得点にすぎない。

ただし，この節の議論は，単一の尺度得点に関するものに限定されていた。1節で述べたように，1つの尺度として集められた項目群からは，複数の下位尺度得点が得られるのが普通である。さらに，前述のように，単

一の概念を測定するつもりで書かれた（集められた）項目の中にも，複数の次元に対応した内容が含まれている可能性もある。次に，そうした多次元の尺度を前提とした方法について述べよう。

3 因子分析とその尺度構成への利用

3.1 因子分析の基本モデルと基本方程式

複数の尺度得点を求めるためには，相互に相関が高く，内容的に類似した項目を集めてグルーピングする必要がある。前述のように，この相関が高いことと，内容的に類似していることとは必ずしも両立しない可能性もあるが，ここではまず，内容については無視し，相関の高いことだけにもとづいて，項目を分類するための統計的手法である因子分析（factor analysis）について考えてみよう。

因子分析の基本モデルは，次のような形をとる。

$$z_{ij} = \sum_{l=1}^{m} f_{il} a_{jl} + e_{ij} \quad i = 1, ..., n \quad j = 1, ..., p \tag{11}$$

ここで，z_{ij}は，個体iの項目jにおける標準化された得点，f_{il}は，個体iの因子lにおける標準得点（因子得点と呼ばれる），a_{jl}は項目jの因子lへの標準偏回帰係数で，因子負荷量と呼ばれるもの，e_{ij}はランダム誤差，mは因子の数である。古典的テスト理論の場合と同様，基本モデルにはいくつかの仮定がなされる。重要なものは次である。

［仮定　Ⅰ］異なる項目のランダム誤差は相互に独立である。

［仮定　Ⅱ］因子得点とランダム誤差はすべて相互に独立である。

これらは，古典的テスト理論の基本モデルにおける仮定と，真の得点が因子得点に置き換わった点を除けば同じである。さらに，解を求めるにあたっては，次の仮定も必要になる場合がある。

［仮定　Ⅲ］因子得点は相互に無相関である。

いずれにしても，これらすべての条件が満たされているとすれば，次の因子分析の基本方程式が導かれる。

$$r_{jk} = \sum_{l=1}^{m} a_{jl} a_{kl} + \delta_{jk} \sigma_{ej}^2 \quad j = 1, ..., p \quad k = 1, ..., p \tag{12}$$

ここで，δ_{jk} は，$j = k$ のとき 1, $j \neq k$ のとき 0 となる変数であり，σ_{ej}^2 は項目 j の分散である。要するに，(12) は，異なる項目間の相関係数は，対応する因子負荷量の積和で表現できること，$j = k$ の場合は，$r_{jj} = 1$，すなわち同一の変数間の相関係数は，標準化された変数の分散である 1 に等しいことを考えれば，

$$\sum_{l=1}^{m} a_{jl}^2 + \sigma_{ej}^2 = 1 \tag{13}$$

であることを意味している。なお，(13) 式は，古典的テスト理論の場合と同様，因子分析においても，因子（正確には共通因子）とランダム誤差（因子分析の本来のモデルでは特殊因子と誤差）が截然と区別できることを仮定している。

項目間の相関係数は，すべてデータとして与えられるので，これらを用いて (12) を解き，因子負荷量と項目の誤差分散を求めることが因子分析のアルゴリズムの中心的部分となる。実際には（$q = p-1$ と設定しない限り），(12) の左辺と右辺は完全には一致しないから，相関係数全体に最もよく当てはまるような解を求めることになる。そのための方法は，「あてはまり」の定義により，いろいろと考えられるが，ここではその問題は重要ではない。どのような解法をとるにせよ，因子負荷量は，そのままでは項目の分類には使えない場合が多く，さらに解を変換する必要がある。

3.2　単純構造にむけての因子負荷量の回転

変換の目標となるのは，通常，単純構造（simple structure）である。

Thurstoneによって導入された単純構造の基準は，いささか込み入っているが，ほぼ次のように要約できるであろう（芝，1979）。

1. 個々の項目の負荷量は，1つの因子に対してのみ大きく，他は0である。

2. 項目全体を通してみたとき，大きな値は因子に均等に散らばる。

たとえば，図2は10個の項目から得られる因子負荷量行列を模式的に示したものである。回転なしでは，（a）のような状態で，1つの項目について，正負入り混じった値が複数の因子にわたって存在しているため，項目を分類することは簡単にはいかない。単純構造を目指した回転の結果，基本モデル（12）の値は全く変えないまま，因子負荷行列は（b）のようになる。

単純構造を目指した変換には，大別して直交回転と斜交回転がある。回転の方法にも多くの種類があるが，その詳細についてはここでは触れない。ただし，直交回転と斜交回転の，下位尺度構成のための方法としての優劣についてだけ述べておこう。直交回転とは，先の［仮定 Ⅲ］を維持したままの回転であり，斜交回転とは［仮定 Ⅲ］を外して，因子得点の間に

（a）

項目	因子（回転なし）		
	Ⅰ	Ⅱ	Ⅲ
1	*	*	
2	*	*	
3	*	*	
4	*	*	
5	*	－*	*
6	*	－*	*
7	*	－*	*
8	*	－*	－*
9	*	－*	－*
10	*	－*	－*

（b）

項目	因子（回転後）		
	Ⅰ	Ⅱ	Ⅲ
1	*		
2	*		
3	*		
4	*		
5		*	
6		*	
7		*	
8			*
9			*
10			*

（空白はゼロに近い数値を示す）

図2

単純構造の達成。回転により，あてはまりはかわらないまま（a）が（b）に変わる。

相関があることを許容する方法である。ただし，その結果として（12）は次のようにやや複雑になる。

$$r_{jk}=\sum_{l=1}^{m}\sum_{l'=1}^{m}a_{jl}a_{kl'}\phi_{ll'}+\delta_{jk}\sigma_{e_j}^{2}\quad j=1,\ldots,p\quad k=1,\ldots,p \tag{14}$$

ここで，$\phi_{ll'}$ は因子間の相関係数を示す。

項目を分類して下位尺度を構成する方法としては，斜交回転が薦められる。その理由は，第1に，斜交回転の方が，より明確な単純構造を示すことであり，第2に，斜交回転で得られる因子間相関が，下位尺度の選択にあたって有利に働くからである。なぜなら，下位尺度は全体としては最初に設定された上位概念を測定するものと考えられるから，相互にある程度の正の相関を持つという，一種の階層構造をなすことが期待されるからである（柳井・村上，1999）。

ところが，直交回転では，その構造がある意味で歪められる。また，斜交回転によって他の因子と無相関，あるいは逆相関であるような因子が見出された場合には，それらに対応する項目を尺度に含めることの可否を問題にすることができるであろう。直交回転では，このことに気がつかれないままに終わる可能性がある。

3.3　因子分析による下位尺度の構成

いずれにせよ，単純構造がおおよそ実現したとすれば，大部分の項目は，どれかの因子に一義的に対応づけられる。すなわち，項目は m 個のグループに分けられ，同じグループに属する項目得点の単純和として，m 個の下位尺度が定義できる。同一の下位尺度に含まれる項目群は，一般に相互に相関が高いから，下位尺度は α 係数で測られる信頼性が高くなる。

また，同一の下位尺度に属する項目内容の共通性に着目すれば，下位尺度が測定している上位概念についてもある程度見当をつけることができる。前述のように，個人差測定尺度の構成にあたっては，項目はある程度一般

性をもった上位概念（たとえば，親和動機）の測定を目指して書かれ（あるいは，集められ），それらが，より詳細に区別される下位概念に対応する下位尺度へと分類される。単純構造への変換を含む因子分析という手法は，仮に研究者が，測定しようとしている概念の下位概念について，極めて漠然とした考えしかもっていなかった場合であっても，ほとんど自動的に下位概念を形成してくれるという便利な装置という側面をもっているわけである。

そこで，因子分析によって適切な因子数が決められ，各因子が適度な単純構造を持つとともに，相互に緩やかに正に相関することが確認できた場

表2　IOS（原版）の主成分パタン行列（Hill, 1987）

		Ⅰ	Ⅱ	Ⅲ	Ⅳ
3	ふれあいから暖かい幸福感を得ることが出来る	**0.35**	0.34	−0.13	0.09
6	他人とのふれあいから満足を得ている	**0.29**	0.19	0.11	0.01
10	他の人を理解することが，最も興味深いことの一つ。	**0.63**	0.18	−0.14	−0.04
11	他の人と一緒にいることから満足を得ている	**0.42**	0.32	−0.01	−0.03
13	誰かと親しくなるのは価値のあること	**0.57**	0.00	0.03	0.11
20	大勢の人と非常に親密な友情を結ぶことができれば満足	**0.40**	0.01	0.21	0.06
24	他の人と一対一で話せることは気に入った気晴らし	**0.71**	0.05	−0.01	−0.03
25	好きな人と新しく友達になればとても満足	**0.56**	−0.07	0.08	0.11
26	最高の楽しみは人々を観察してどんな人なのかを知ること	**0.57**	−0.03	0.01	−0.07
1	つらい時の最大の慰めは他の人と一緒にいることである	−0.01	**0.70**	−0.05	0.07
4	混乱したときは信頼できる親友と一緒にいたいと願う	0.15	**0.46**	0.02	0.05
9	他の人のそばにいるだけで気持ちを持ち直すことが出来る	0.09	**0.56**	−0.01	0.00
15	誰かがそばにいてくれると苦しさが少なくなる	0.13	**0.53**	0.07	0.03
17	気分を良くするために他の人のそばにいようと努める	−0.03	**0.77**	0.12	−0.03
23	動揺したときにはそばに誰かがいてほしい	−0.02	**0.72**	0.13	−0.03
5	私に好意をもってくれているような人を私も好きになる	−0.07	0.12	**0.58**	−0.01
8	注目の的となりそうなときは他の人のそばにいたい。	0.00	−0.04	**0.56**	0.12
16	私を高く評価してくれる人たちのそばにいたい	−0.03	0.13	**0.75**	0.01
19	私を重要で刺激的な人間だと考えてくれる人のそばにいたい	0.08	0.00	**0.71**	0.01
21	私の人となりを理解してくれる人がそばにいてほしい	0.15	−0.04	**0.64**	0.06
22	ほめてくれないような人とは一緒にいたくない	0.00	−0.02	**0.44**	0.00
2	活動をするときは他の人と一緒に参加したい	−0.08	0.32	−0.08	**0.43**
7	比べることが出来るので他の人のそばにいたい	0.01	−0.06	−0.01	**0.74**
12	期待されていることがはっきりしないときは他の人を当てにする	0.08	−0.02	0.12	**0.48**
14	自分と同じ経験をしている人のそばにいたいと願う	0.22	0.10	0.09	**0.30**
18	どの位できているかを知るために，他の人に注目する	0.00	−0.06	0.24	**0.58**

表 3　尺度間相関（Hill, 1989）

	刺激	支持	注意	比較
刺激		0.55	0.28	0.31
支持	0.58		0.27	0.40
注意	0.39	0.30		0.45
比較	0.47	0.39	0.52	

（下三角部分は男子，上三角部分は女子）

合，それぞれの因子に高い（たとえば，0.5 以上の）負荷量をもつ項目を選択することにより，下位尺度を定義するという方法は，多くの研究で頻繁に採用されてきた。なお，どの因子にも高く負荷しない項目については，いずれの下位尺度に含めても，α 係数を低下させるから，「残余項目」などと称して削除されることが多い。

実際，Hill（1987）は，表 1 に示した IOS の 26 の項目を（もちろん，英文の原文で），1,078 名の心理学入門コースの学生に実施した結果について，主成分分析（principal component analysis; この方法と 3 節で説明した狭義の因子分析との関係については後述）を斜交回転した因子負荷量行列（因子パタン）を示している。表 2 に，IOS 日本語版の項目（表記を短縮した）とともにそれを再現した。確かに表 1 の 4 つの下位尺度への分類は，これを見る限り正当であるとみることができる。Hill（1987）には，因子得点間相関は示されていないが，単純和によって定義された下位尺度得点間の相関係数を男女別に示しており，表 3 にそれらを再録した。

4　尺度構成への因子分析の適用の問題点

4.1　因子分析による尺度の再現性

このように便利な道具として活用されている因子分析ではあるが，現実場面ではいろいろと困った状況に遭遇することがある。たとえば，中村（2000）らは，日本の大学生 1,645 名に対し，IOS の日本語版を実施した。

表4　日本語版の因子パタン

	I	II	III	IV	h^2
3 ふれあいから暖かい幸福感を得ることが出来る	0.11	**0.54**	0.26	−0.21	0.36
6 他人とのふれあいから満足を得ている	0.30	**0.52**	−0.04	−0.17	0.39
10 他の人を理解することが，最も興味深いことの一つ。	**0.76**	0.09	0.01	0.04	0.69
11 他の人と一緒にいることから満足を得ている	**0.51**	0.49	−0.09	−0.06	0.54
13 誰かと親しくなるのは価値のあること	**0.38**	−0.04	0.37	0.12	0.36
20 大勢の人と非常に親密な友情を結ぶことができれば満足	0.02	0.12	**0.57**	−0.06	0.48
24 他の人と一対一で話せることは気に入った気晴らし	0.27	0.18	0.46	−0.14	0.35
25 好きな人と新しく友達になればとても満足	0.07	0.11	**0.62**	−0.20	0.19
26 最高の楽しみは人々を観察してどんな人なのかを知ること	**0.76**	−0.22	0.09	0.13	0.79
1 つらい時の最大の慰めは他の人と一緒にいることである	−0.03	**0.77**	−0.02	−0.04	0.72
4 混乱したときは信頼できる親友と一緒にいたいと願う	−0.05	**0.64**	0.19	0.00	0.43
9 他の人のそばにいるだけで気持ちを持ち直すことが出来る	0.18	**0.56**	−0.23	0.22	0.52
15 誰かがそばにいてくれると苦しさが少なくなる	−0.14	**0.60**	0.16	0.23	0.66
17 気分を良くするために他の人のそばにいようと努める	−0.01	**0.63**	−0.04	0.27	0.56
23 動揺したときにはそばに誰かがいてほしい	−0.08	**0.64**	0.19	0.09	0.56
5 私に好意をもってくれているような人を私も好きになる	−0.14	0.24	**0.37**	−0.01	0.28
8 注目の的となりそうなときは他の人のそばにいたい。	0.02	0.15	−0.12	**0.59**	0.34
16 私を高く評価してくれる人たちのそばにいたい	−0.01	0.04	**0.61**	0.24	0.44
19 私を重要で刺激的な人間だと考えてくれる人のそばにいたい	0.09	−0.15	**0.69**	0.14	0.55
21 私の人となりを理解してくれる人がそばにいてほしい	−0.02	0.11	**0.67**	0.08	0.42
22 ほめてくれないような人とは一緒にいたくない	−0.08	−0.16	0.28	**0.34**	0.20
2 活動をするときは他の人と一緒に参加したい	−0.08	**0.41**	−0.03	0.27	0.15
7 比べることが出来るので他の人のそばにいたい	0.01	0.13	−0.04	**0.67**	0.44
12 期待されていることがはっきりしないときは他の人を当てにする	0.16	0.04	−0.03	**0.63**	0.30
14 自分と同じ経験をしている人のそばにいたいと願う	0.03	0.13	0.22	**0.50**	0.45
18 どの位できているかを知るために，他の人に注目する	0.03	−0.10	0.20	**0.61**	0.41

やはり主成分分析によって得られた因子負荷量を Promax 法によって斜交回転したものを，表4に示した。表2とはかなりの違いが見られる。

また，相関行列の固有値を降順に，その順位に対してプロットした scree plot を図3に示したが，必ずしも4が適切な因子数ではないことがわかる。

こうしたことが発生する理由としては，さまざまなものが考えられる。まず，常にサンプルは有限の個人からなるものである以上，サンプリング誤差の影響が考えられる。ただ，2つの標本はともに 1,000 以上の大きさをもっており，通常の統計学的常識からすれば，こうした大きな差異を

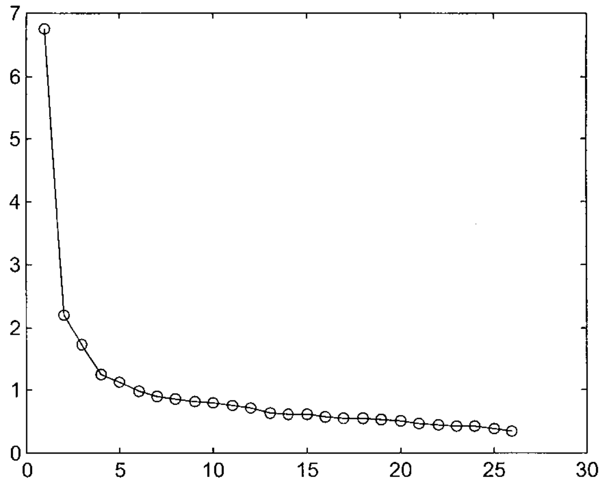

図3　IOS（日本語版）の scree plot

生じることは考えにくいであろう。

ただし，この種の分析にあたっては，以下のような点も考慮しなければならない。第1に，前述のように，この尺度に限らず，個人差測定尺度は，全体としては1つの個人差の次元の測定を目指すため，1因子性が高い傾向がある。すなわち，相関行列の第1固有値が著しく大きくなり，2番目以降の固有値は概して小さくなりがちである。このことは，前述の項目の階層的構造に対応しており，これは図3から示唆されるとおり，IOS においても成り立っている。この場合，いわゆる「固有値の推移」から因子数を決定することが難しくなるだけでなく，オリジナル版の分析データでは表面に出ていた変動が，日本語版では捨てられた成分の方に入ってしまうといったことがありえよう。

第2に，質問文の言語的な意味の問題が考えられる。特に翻訳の場合には，原語の意味を忠実に再現する翻訳を行うことは難しく，複数の質問内容の相互の類似性が，言語の場合とは異なってくることもありえよう。

第3に，複数の理由による負荷量の布置の変化が，比較的微小なものであった場合でも，項目を下位尺度に分類した結果には，大きな変化が起こることがありうることを指摘する必要がある。仮に，ある項目の2つの因子への負荷量の差がわずかであっても，2つの下位尺度の一方に一義的に

（いわばディジタルに）分類しなければならない状況では，僅かの差が結果を大きく変えるように見えることはありうる。

こうした点からすれば，単純構造への回転を含む因子分析を，完全に探索的に用い，その結果をそのまま採用することには，大きな危険があることは明らかであろう。

4.2　現実場面での対処

そうは言っても，実際の研究の過程では，単に「結果が不一致である」とか，「意図した下位概念が得られなかった」といって済ませるわけにはいかないのが現実であり，何らかの対策をとらざるを得ない。その場合，より進んだ数理的手法に頼りたくなるのは，もっともである。

近年，アルゴリズムの発達とソフトウェアの普及，それに伴う適用例の蓄積によってしばしば用いられるようになった確認的因子分析（confirmatory factor analysis）や，その上位手法としての構造方程式モデリング（structural equation modeling; SEM，共分散構造分析）はそうした方法の候補である。こうした方法では，あらかじめ理論的に定められた項目の下位概念への分類を仮説的構造として分析を行うことができる。IOS の例で言えば，Hill（1987）のオリジナルな下位尺度の定義を仮説的構造として採用できるわけである。

しかしながら，こうした進んだ方法も，必ずしも意図した結果をもたらさないことが多い。すなわち，こうした項目水準のデータに対するモデルの当てはまりは通常の基準を大きく下回る。その問題を解決するために，多くの項目を削除し，1つの因子あたりの項目数を2～3項目に絞り込むことがしばしば行われた。もし，因子モデルへの当てはまりが完全に近いものであれば，このような大幅な削減を行っても因子の性質は変化しないはずである。しかしながら，このようなやり方で得られた潜在変量間の相関関係は，項目の選択によってかなり変動し，そこから得られる因果モデルは再現性に乏しいことが多いようである。そもそも，項目が相互に交換

可能な程度に因子モデルへの当てはまりが良いのであれば，項目を削除する必要はなかったはずであるから，これは当然の結果である。

なお，モデルへの当てはまりを改善するために，事後的に誤差項の間に相関を仮定するというやり方も時々見かけるが，著者には，あたかも天動説を維持するための周転円のように思われる。

実は，探索的因子分析を用いて項目分類をする場合にも，アルファ係数の上昇だけを基準にすると，内容的に良いと思われる項目が削除の対象になりやすいという感触がある。ただ，実務家は意外にこの点には，こだわりをもたないことが多いのではあるが。

この問題に対処するために，最近では，parcel と呼ばれる項目反応の単純和（つまり，本研究で言う尺度値である）を分析の対象とすることが推奨されている（たとえば，Kline, 2005）が，それでは parcel に含めるべき項目の選択という問題が発生する。つまり，ここで扱っているような問題を，潜在方程式モデルの枠組みだけで扱うことは困難であると言わざるを得ない。

また一方では，項目数が多いと，確認的因子分析における仮説的構造が，実務家によって自信を持って構成できないという問題もありそうである。現実問題として，多くの質問項目が演繹的に導出できるほど，行動科学の理論は精緻なものとはなっていないのが普通であると思われるからである。

4.3 問題の所在

こうした問題が生じる理由は，以下の点にあると考える。

(1) 心理測定モデルが，測定の対象となる概念の構造にうまく対応していない。特に真の得点（あるいは共通因子）とランダム誤差（＋独自因子）の截然たる区別が可能であるという想定は，少なくとも項目水準では現実と合わない。

(2) 一方，実務家のもつ理論は，質問項目を導出し，意味づけるには不完全であり，しばしば，断片的である。

(3) 数理的方法が，そうした断片的情報を扱うには不向きである。すなわち，探索的方法と確認的方法は，完全に切り離されており，不十分な情報を取り入れにくい。データ縮約で必要な情報が切り捨てられる。

(4) 根底にある科学観の問題。データから学ぶ姿勢の欠如，あるいは，データに頼りすぎる態度の両極端が見られる。

こうした問題点を解決して，より尺度構成の現場の実態に合った理論と方法を開発するために，まずは従来の理論と方法の根底にあった考え方を明確にすることから始めよう。

5　個人差測定尺度のプロトタイプ理論

5.1　尺度に関する従来の想定とその問題点

ここではまず，古典的テスト理論と因子分析を通じて，仮定されてきたことを6つの命題にまとめ，それぞれについて若干のコメントを付していこう。

1. 多数の項目に対する反応は，比較的少数の直接観察不可能な潜在変数の関数である。

観察可能な行動が，直接見えない要因によって規定されているというのは，心理学においては，極めて広汎な分野でなされている仮定である。ここで「潜在（latent）」という語には，可能性はあるがまだ実現していないといった意味合いは含まれておらず，実在してはいるが直接観察不可能という意味で用いられている。質問項目よりも少ない数の潜在変量によって，項目反応の変動（分散）の一定の部分が説明できるという想定は，データ縮約という観点からしても，概念形成という観点からしても，まずは動かせない前提のように思える。しかし，本研究はこの前提を疑うところから始めたい。

2. 大部分の項目は，1つの潜在変数にのみ関わる。すなわち，大部分の項目反応の意味は単純である。

これは，事実であるとすれば，潜在変数のデータ縮約と概念形成の機能を大幅に強めるものである。実際，因子分析にこの性質を持たせるために，因子負荷量の単純構造への変換が有用だったと考えられる。これまた，動かせない前提であるように見えるが，この点についても疑ってかかること必要であると考える。

3. 項目反応は，潜在変数以外にもランダム誤差による影響を受けるが，それらは，潜在変数とも，他の項目のランダム誤差とも独立である。すなわち，個人差を反映した systematic な反応とランダムな変動とは截然と区別できる。

項目反応の個人差の大半が，少数の潜在変数によって測られる個人差よって説明されるとすれば，さらに，1 つの項目反応は 1 つの潜在変数によって説明されるとすれば，問題は極めて単純になる。この場合，複数のデータの因子分析結果が食い違うといったことはほとんど起こらず，扱われている個人差についても，比較的（学問的にも，実用的にも）受けいれやすいものとなろう。しかし，多くのデータ分析の経験が示しているように，この想定は現実的でない。

4. 調査票に含められた項目群は，意図した概念の測定のために完全である。項目間には，潜在変数を反映する度合い（負荷量）に関する差異はあるとしても，それ以外の点で違いはない。

4'. 測定の信頼性を犠牲にするなら，どの項目も尺度の「意味」を変えることなく削除することができる。

これは，ここまでの想定の必然的な帰結であり，それらが正しいとすれば，この議論は論理的に正しい。現実の尺度構成においては，最終的に残される項目よりもずっと多くの項目が候補として作成される。実際，表 1 の IOS の場合においても，Hill は 41 項目からなる最初の版から，（恐らく）因子分析によって，最終版の 26 の項目を選択している。こうした項目選択にあたっては，負荷量の高低が選択の基準となる。負荷量の高い項目は，(8) 式から示唆されるように，得点の変動に含まれる誤差分散の割

合が小さく，したがって他の項目との相関が高い傾向がある。したがって，(10) 式の α 係数で測られる尺度得点の信頼性を高める効果をもつ。

尺度全体として項目数が多くなりすぎるために，項目を削減する場合の手続きについても同様であり，α 係数の低下を最大限抑えるように，やはり負荷の低い項目が削除されることが多いようである。こうした手続きにおいては，同一の因子にある程度の負荷をもつ項目間で，その意味に相違があるといった可能性については，全く考慮されていないように見える。

現実には，質問項目が加除され（入れ替えられ）れば，因子，あるいは尺度の測定内容は微妙に，時には，かなり変化するのは，経験的には常識である。

5. 潜在変数は実在する。

5'. 潜在変数は測定を意図した構成概念そのものである。すなわち，測定尺度の妥当性は完全である。

最初の命題の正当性は，「実在」という言葉の意味にかかっている。項目間に相関があること，すなわち，ある質問項目に対して肯定的に反応する個人が，別の項目に対しても肯定的に反応する傾向があるという事実は，それらの項目反応の背後に「何か」の存在を仮定する根拠にはなるであろう。これが潜在変数である。ただ，それらがさらに，個体の内部に潜在変数（上の値）に対応する何らかの生理学的メカニズムや，遺伝子レベルでの何らかの特徴，あるいは，何らかの具体的・組織的な経験の蓄積といったものの存在へと一般化されるとすれば，項目反応からの推測としては行き過ぎであろう。

第2の命題は，論理的誤りであると考えられる。しかし世上，データから導き出された潜在変数と，論理的な産物である構成概念とを，単に直接観測可能でないものという意味で同一視しているような議論があるのも事実である。ただし，潜在変数の「意味」が，質問の内容とかかわりをもつという事実は否定できないから，これをただ全く無根拠な命題として片付けるもの適当でないであろう。

潜在変数と構成概念の関係を検討していくのがvalidationの過程であるとするならば，こうした議論の可能性と限界について，ある程度明らかにしておく必要はあろう。

6. 構成概念の定義は，先験的に与えられるべきである。データは，その先験的定義にもとづく仮説の検証のためだけに用いられるべきである。

こうした言説は，論理実証主義に基礎付けられた科学論の本流からすれば，まったく正統的な命題なのであろう。概して分析家は，この種の信念を持っている場合が多い。因子分析の数理的に厳密な解説には，多くの場合，本論文でいう「概念形成」の機能を謳わないものが多いのも，それが理由であろう。

ただし，尺度構成の手順の中で，現実に因子分析が果たしている機能からすれば，上のような命題は信じがたいものである。しかし，データが多変量正規分布することを前提とするような因子分析のモデルは，Likert型項目のような，本来の意味での数値（いわゆる，間隔尺度以上の変数上の値）であるかどうかも疑わしいようなデータに適用されることは想定されていなかったとも言える。その点では，現実場面での実用性を優先する実務家に対して，それを理論的に基礎付ける努力が不足していたとも考えられる。

5.2　家族的類似にもとづくモデル

前項で列挙した「原理」は，概して後のものほど非現実的である。もちろん，このようなことがそのまま信じられているとは考えにくいが，現実の「本質的」な側面を捉えて単純化するというモデルの本性からして，現実がモデルを通じてしか見えなくなるという弊害は付き物であるとも言えよう。

ここでは，前項の命題を逐一否定する形で，質問項目の分析によりふさわしいと思われる新たな原理を提出していきたい。

1. 個々の項目は，単一の潜在変数ではなく，極めて多くの観測不可能な

要因の影響によって変動する。

従来の考え方と（前節の 1）と，この考え方との相違は，図 4 に表されている。図 4 の右側のようなモデルがここでの提案である。これは，今までの測定モデルとは著しく隔たった考え方であろう。このようなモデルでは，まずデータ縮約の役目を果たさない。また，概念形成の能力も疑わしくなる。何よりも，変数の数よりもパラメータの数の多いモデルでは，解を一意的に定めることができないのではないか。

いずれももっともな反論であるが，まず，潜在変数を一意的に定めようとする意図はもっていない。潜在変数という用語を避けあえて「観測不可能な要因」としたのは，潜在的変数には離散的な変数，あるいは 2 値変数も含まれる可能性を含めるためである。ただし，個々の要因の性質も数もデータから定めることはできないし，その必要もないと考えている。この想定から導かれる，以下の含意が重要である。

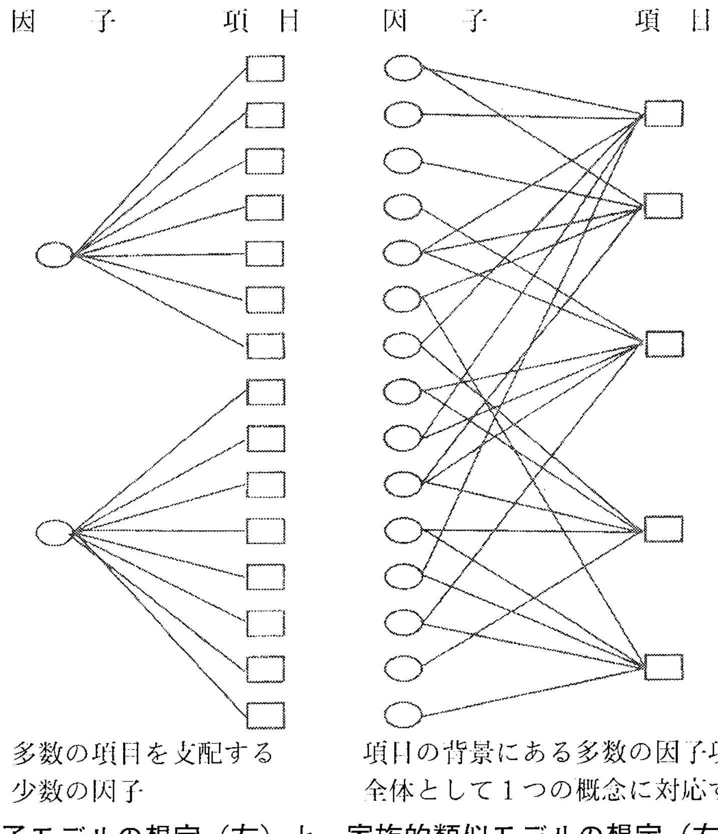

図 4　通常の因子モデルの想定（左）と，家族的類似モデルの想定（右）（村上，1989）

2. 多くの観測不可能な要因は，複数の項目によって共有されるが，仮に同じ下位概念に分類される項目であれ，そのすべてによって共有されるとは限らない。

観測不可能な要因は，基本的に相互に独立である，あるいは少なくとも1次独立であると考える。項目間相関の高低は，2つの項目がどの程度，要因を共有しているかによって定まると想定するわけである。こうした考え方は，まったく新しいものではない。たとえば，Cox（1939）は，この方法で生成した相関行列にもとづいた因子分析のシミュレーションを行っている。

村上（1989，1993）は，こうした想定から導かれる測定モデルを，Wittgenstein の言う家族的類似（family resemblance）にもとづく概念形成の議論にもとづき，家族的類似のモデルと呼んだ。ただし，共有されている要因の個数について推定することは，現時点では不可能という外はない。

3. 多くの項目によって共有された複数の観測不可能な要因の合成されたものが，因子を構成する。したがって，変動の少なくとも一部は，選択された項目次第で潜在変数に分類されることもあれば，ランダム誤差（固有因子を含む）に分類されることもある。

これは，前の2つの命題からの自然な帰結である。たまたま，単一の項目にのみ影響を与えている要因は，独自因子，あるいはランダム誤差に分類されざるを得ない。しかしながら，もし，この要因に影響を受ける項目が調査票に加えられ，分析の対象となるならば，その要因による変動は共通因子に属することになる。

すなわち，項目得点の変動は，真の得点と誤差，あるいは共通因子由来の成分と，独自因子由来の成分＋ランダム誤差に截然と2分できるわけではない。

4. 調査票に含められた項目は，測定を目指す概念，下位概念をすべてカバーするものであるとは限らない。

4'. 因子の意味は，調査に用いられた項目によって作られる観測不可能な要因の分布によって，異なったものとなりうる。

第１の命題は，ある程度の一般性をもつ概念を測定しようとする尺度の開発者，実務家，あるいは，内容領域の研究者にとっては，わざわざ明示的に述べるまでもない自明の事柄であろう。むしろ，分析家の方がこの点については，質問項目の「完備性」を根拠もなく信頼する傾向があるのではなかろうか。ともかく，（未だに思いつかれてすらいないかもしれないにせよ，）測定の目的にかなった項目はまだ他にありうることを常に意識する必要がある。

第２の命題は，ここまでの命題からの自然な帰結である。このように考えることによって，因子は（たまたま）調査票に含まれた項目によって，内容的に異なったものとなりうることがわかる。当然，単純和によって定義される尺度得点も，どのような項目を下位尺度に含めるかによって，実質的に別のものになる可能性も示唆される。

5. 他の情報が無い限り，質問項目だけで定義される潜在変数は，実在する保証はない。しかしながら，潜在変数はデータによる裏づけがある以上，単なる理論上の構築物でもない。

5'. 潜在変数と構成概念とは区別されるべきである。データから導かれた潜在変数と理論上の産物である構成概念が対応することを示すためには，妥当性検討のための尺度外の情報が必要である。

この点は，伝統的な測定の妥当性の議論の前提であり，特に新しいことを述べているわけではない（たとえば，村上，2003)。残念ながら，進んだ数理的方法によって，こうした常識が隠されてしまう傾向があるように思われる。

むしろ，質問項目への反応が，何らかの潜在的次元上の個人差の反映とみることが，理論的に，あるいは実用的に適切かどうかについては，近年，より本質的な問題が提起されている（Fayers and Hand, 2002)。これは，実証研究の手続き自体にも大きく影響する重要な問題であるが，本論文で

カバーする議論の範囲を超えている。

6. 尺度の定義は先験的にのみ与えられるわけではない。データに基づく理論の形成や修正もありうる。

尺度というものは，理論にもとづくいわばトップダウンのアプローチと，データ分析にもとづくボトムアップアなアプローチの両方から，徐々に定義を固めていくものであると考える。理論の形成過程は双方向的に行われるべきものである。

ただし，これは科学方法論としては，かなり危険なものであるとみなされるかもしれない。もし，探索的データ分析から得られた命題を，そのまま検証された事実と見なしてしまうならば，これは重大な誤謬につながりかねない。ここで主張していることは，そうしたことではなく，仮説の生成過程において経験的事実を参照することは許されるはずだということに過ぎない。仮説は天から降ってくるとは限らないからである。

5.3　プロトタイプモデル

家族的類似を論じていた時点で，村上（1989, 1993）は，探索的因子分析を尺度構成の方法とすることに，特に大きな問題は感じておらず，ただその結果の利用について，前項で述べたような柔軟な対応を求めていたにすぎなかった。しかしながら，質問項目の意味にもとづく実務家の仮説を分析に取り入れるという観点からは，なおなすべきことが残っていた。そのヒントは，1970 年代から認知心理学において展開されてきたプロトタイプ理論（prototype theory）から得られた。

Rosch（1987）によれば，家族的類似にもとづく概念には，その概念の典型となる中心的な成員，すなわちプロトタイプと，より周辺的な成員が存在するという。心理測定における個々の質問項目と測定の対象となる下位概念の間にも，同様な関係が成立すると想定することは自然である（図5）。

もし，尺度開発者があらかじめ下位尺度についてある程度の断片的なア

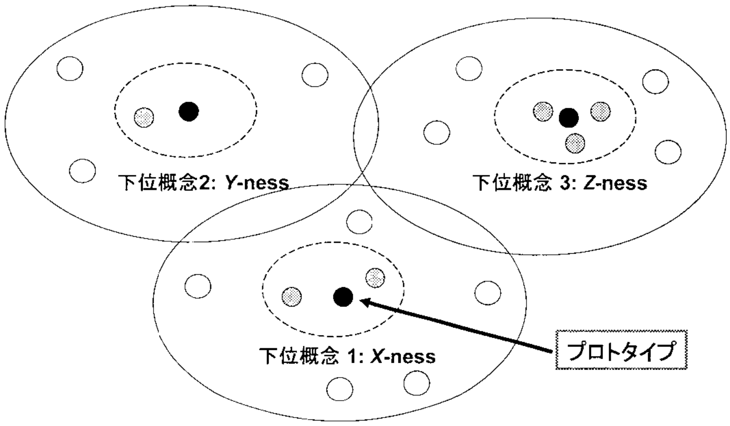

図5　3つの下位概念と，プロトタイプ項目と周辺項目を示す模式図

イディアをもっており，使用するすべての項目について下位尺度への分類ができるわけではないとしても，各下位尺度についてその中心となると想定する項目，すなわちプロトタイプ項目を指定できるとすれば，若干の新たな方法を提案することができる。次の節では，そのような方法について論じよう。

6　プロトタイプ変換

項目水準のデータを，項目の下位尺度への適切な分類を目的として分析する場合，その方法が備えるべき特徴は，次のようになると考えられる。

(1) データ縮約は行うが，真の得点とランダム誤差の区別をあいまいなままにとどめる。

(2) 内容領域の専門家が指定する断片的な理論を分析結果に反映することが容易にできる。

こうした条件を満たす方法として，主成分のプロトタイプ変換（prototype transformations of principal components）を提案する。これは，以下のようなアイディアである。

まず，すべての項目反応を主成分分析し，負荷量行列を求める。主成分分析は狭義の因子分析よりも，ランダム誤差に対するモデルが厳しくなく，

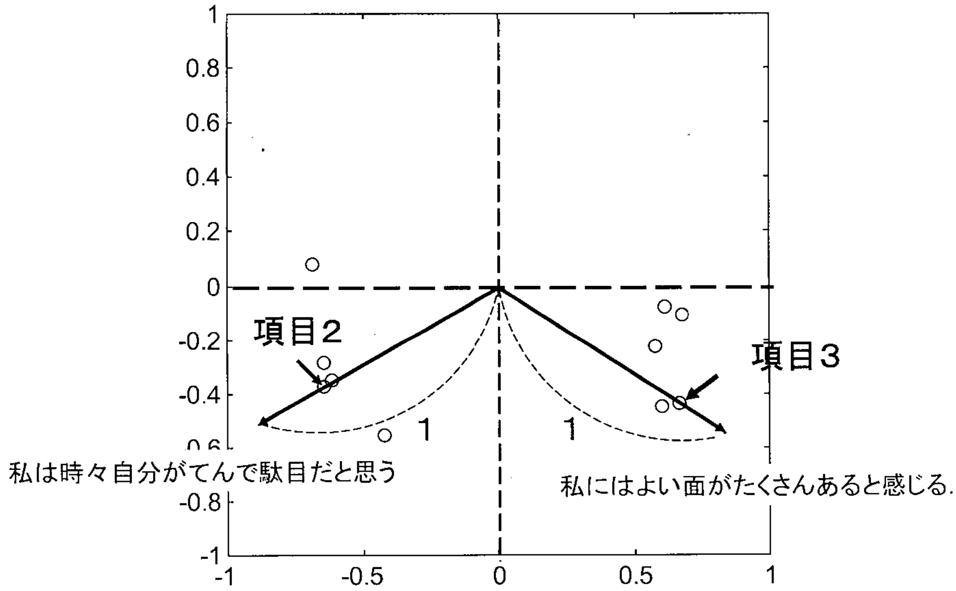

図6　プロトタイプ変換

Rosenberg の「自尊感情尺度」の10項目の主成分分析で得られた2つの主成分負荷量を図示したもの。このうち，項目3と項目2を通る長さ1のベクトル（プロトタイプベクトル）を主成分空間の新たな軸として設定する。構造ベクトルの要素は，各項目からプロトタイプベクトルへの射影の長さであり，各項目とプロトタイプとの類似度と見ることができる。

この分析の目的によりふさわしいと考えられる。この際，主成分の数は，次の段階で選ばれるプロトタイプよりも多くとることも許す。これは，想定されたプロトタイプ次元に対応する変動が，抽出されないままになるリスクを回避するためである。ただし，この数を多くしすぎれば，説明力が低下するので，最終的には慎重な選択が必要になる。次に，実務家によって選ばれた複数のプロトタイプ項目の主成分空間への射影ベクトルを長さ1に延長し，それを新たな主成分とする[1]。

図6は，中村（2000）にあるRosenbergの自尊心尺度の10項目について，この方法の具体例を示したものである。

分析の具体的な手順は以下の通りである。

1　項目間相関行列の固有値と固有ベクトルをすべて求める。

2　スクリー・プロット等を参考にしながら，暫定的に主成分数 m を決める。

3　主成分負荷行列を，

$$\mathbf{A} = \mathbf{K}_m \mathbf{L}_m^{1/2} \tag{15}$$

によって求める。ここで，$\mathbf{L}_m, \mathbf{K}_m$ はそれぞれ，相関行列の最大 m 番目までの固有値を要素とする対角行列と，対応する正規直交な固有ベクトルを列とする行列である。これは，回転されていない直交する主成分と各項目得点との相関行列，あるいは，（同じことであるが）各項目の主成分への標準回帰係数の行列である。

4　指定されたすべてのプロトタイプ項目について，対応する構造ベクトル，すなわち主成分空間内における各項目とプロトタイプ項目との相関係数を要素とするベクトルを，次の式で求める。

$$\mathbf{s}_t = \mathbf{A}\mathbf{a}_t(\mathbf{a}_t{}'\mathbf{a}_t)^{-1/2} \quad t \in T \tag{16}$$

ここで，t は指定されてプロトタイプの項目番号（$1 \leq t \leq p$）であり，$\mathbf{a}_t$ は $\mathbf{A}$ のプロトタイプ項目に対応する行ベクトルを転置した（列）ベクトル，また，T はすべてのプロトタイプの項目番号の集合である。T の要素の数 q については特に制限はないが，m を超えないとするのが常識的であろう。

5　すべての構造ベクトルを並べた $p \times q$ の行列 $\mathbf{S}$ を作る。各行について，最大値を求め，対応する列番号の集合にその項目の番号を所属させる。このようにして，すべての項目を下位集合に分類する。

6　各下位集合に属する項目による尺度得点の α 係数を求める。

7　各階集合の項目について，（できれば尺度構成を行った実務家とともに）検討する。また，各項目を除去した後の α 係数を検討し，著しく α 係数を低下させている項目について，問題点を検討する。

8　必要なら，m を変えて再計算を行う。さらに，（実務家の考えにもとづいて）プロトタイプ項目を入れ替え，構造ベクトルを再計算し，それにもとづいて，項目を分類しなおす。このプロセスは，実務家と分析家の双方が納得するまで繰り返す。

この際，最適解を1つに絞ることを急がない方がよいであろう。たとえば，複数の下位尺度群によって，何らかの外部規準を予測する研究を行ってみることは，解を選択するための示唆を与えることになろう。そうした分析のために，下位尺度得点を標準得点化されたデータ行列から直接求めるための重みベクトルは，次のように計算できる。

$$\mathbf{w}_t = \mathbf{K}_m \mathbf{k}_t (\mathbf{k}_t' \mathbf{L}_m \mathbf{k}_t)^{-1/2} \quad t \in T \tag{17}$$

ここで，$\mathbf{k}_t$ は $\mathbf{K}_m$ の（プロトタイプ項目に対応する）第 t 行を転置したベクトルである。全対象者の得点は，標準得点化された $n \times p$ のデータ行列を $\mathbf{Z}$ とすると，$\mathbf{Zw}_t$ で求めることができる。

7　適用例

先に例示したIOSの日本語版に対してこの方法を適用する。尺度開発者に対してプロトタイプ項目を直接尋ねることは不可能なので，Hill (1987) における各下位概念の定義から推測して，4つの下位概念について，表5で枠で囲んだ各1項目をプロトタイプとした。主成分の次元数 m は4とした。表5は，表4と（構造とパタンの違いがあるにもかかわらず）驚くほどよく似ている。残す主成分の数を5，6と増加させても，結果にはほとんど影響がなかった。

このことは，選択したプロトタイプ項目が，日本語版データの項目のクラスターの布置のほぼ中心に位置していたことを示すように思われる。また，このことは，主成分分析＋プロマックス回転の結果が，単なる気紛れの産物でないことを物語ってもいる。

そこで，表4と表5を改めて検討した上で，「ポジティブな刺激」について，表6のような3つのプロトタイプを指定することにした[2]。これらは，対人関係からくる幸福感（感情的側面，項目3），知的な他者理解（認知的側面，項目10），対人関係の形成（社会的側面，項目20）という，

表5　日本語版のプロトタイプ変換（1）

	刺激	支持	注意	比較	h^2
3 ふれあいから暖かい幸福感を得ることが出来る	0.41	**0.62**	0.30	0.07	0.49
6 他人とのふれあいから満足を得ている	0.50	**0.58**	0.09	0.03	0.45
10 他の人を理解することが，最も興味深いことの一つ。	**0.80**	0.33	0.28	0.14	0.64
11 他の人と一緒にいることから満足を得ている	**0.69**	0.60	0.14	0.13	0.61
13 誰かと親しくなるのは価値のあること	0.49	0.20	**0.52**	0.24	0.38
20 大勢の人と非常に親密な友情を結ぶことができれば満足	0.25	0.29	**0.56**	0.15	0.38
24 他の人と一対一で話せることは気に入った気晴らし	0.48	0.37	**0.51**	0.09	0.44
25 好きな人と新しく友達になればとても満足	0.29	0.29	**0.56**	0.03	0.43
26 最高の楽しみは人々を観察してどんな人なのかを知ること	**0.70**	0.06	0.34	0.14	0.56
1 つらい時の最大の慰めは他の人と一緒にいることである	0.30	**0.74**	0.10	0.22	0.55
4 混乱したときは信頼できる親友と一緒にいたいと願う	0.30	**0.68**	0.29	0.28	0.51
9 他の人のそばにいるだけで気持ちを持ち直すことが出来る	0.38	**0.58**	0.05	0.36	0.43
15 誰かがそばにいてくれると苦しさが少なくなる	0.21	**0.64**	0.34	0.48	0.53
17 気分を良くするために他の人のそばにいようと努める	0.29	**0.65**	0.20	0.48	0.51
23 動揺したときにはそばに誰かがいてほしい	0.27	**0.68**	0.33	0.36	0.53
5 私に好意をもってくれているような人を私も好きになる	0.08	0.30	**0.36**	0.17	0.22
8 注目の的となりそうなときは他の人のそばにいたい。	0.11	0.19	0.18	**0.60**	0.36
16 私を高く評価してくれる人たちのそばにいたい	0.22	0.25	**0.70**	0.43	0.54
19 私を重要で刺激的な人間だと考えてくれる人のそばにいたい	0.25	0.10	**0.73**	0.29	0.53
21 私の人となりを理解してくれる人がそばにいてほしい	0.25	0.32	**0.70**	0.31	0.55
22 ほめてくれないような人とは一緒にいたくない	−0.03	−0.05	**0.36**	0.35	0.22
2 活動をするときは他の人と一緒に参加したい	0.12	**0.41**	0.15	0.39	0.25
7 比べることが出来るので他の人のそばにいたい	0.12	0.21	0.28	**0.70**	0.48
12 期待されていることがはっきりしないときは他の人を当てにする	0.23	0.17	0.31	**0.64**	0.43
14 自分と同じ経験をしている人のそばにいたいと願う	0.20	0.27	0.46	**0.60**	0.43
18 どの位できているかを知るために，他の人に注目する	0.10	0.04	0.44	**0.63**	0.46

下位概念のかなり異なる側面である。計算上は，3つの項目の重心（平均ベクトル）をプロトタイプとした。主成分の数は，4の場合と5の場合を試みた。結果には大きな差はないが，ここでは，やや原版の構造に近い $m = 5$ の場合を表6に示した。

結果的に，ほぼ原版に近い下位尺度が得られたので，これらの α 係数を求めたところ，順に，0.71, 0.71, 0.58, 0.61 となった。相関研究として用いるには最低限の水準に到達している。また，尺度間相関行列は表7のようになった。オリジナルの表2に比べるとやや高めであり，これは，こ

表６　日本語版のプロトタイプ変換（２）

		刺激	支持	注意	比較	h^2
3	ふれあいから暖かい幸福感を得ることが出来る	**0.64**	0.62	0.27	0.09	0.49
6	他人とのふれあいから満足を得ている	**0.60**	0.59	0.02	0.10	0.51
10	他の人を理解することが，最も興味深いことの一つ。	**0.63**	0.32	0.29	0.11	0.43
11	他の人と一緒にいることから満足を得ている	**0.67**	0.61	0.11	0.15	0.55
13	誰かと親しくなるのは価値のあること	**0.51**	0.20	0.52	0.21	0.36
20	大勢の人と非常に親密な友情を結ぶことができれば満足	**0.56**	0.31	0.45	0.25	0.36
24	他の人と一対一で話せることは気に入った気晴らし	**0.64**	0.38	0.48	0.09	0.44
25	好きな人と新しく友達になればとても満足	**0.60**	0.30	0.47	0.11	0.39
26	最高の楽しみは人々を観察してどんな人なのかを知ること	**0.49**	0.06	0.33	0.13	0.40
1	つらい時の最大の慰めは他の人と一緒にいることである	0.51	**0.74**	0.09	0.23	0.55
4	混乱したときは信頼できる親友と一緒にいたいと願う	0.53	**0.68**	0.30	0.25	0.51
9	他の人のそばにいるだけで気持ちを持ち直すことが出来る	0.37	**0.57**	0.09	0.30	0.34
15	誰かがそばにいてくれると苦しさが少なくなる	0.42	**0.62**	0.38	0.40	0.57
17	気分を良くするために他の人のそばにいようと努める	0.39	**0.63**	0.27	0.39	0.51
23	動揺したときにはそばに誰かがいてほしい	0.47	**0.66**	0.39	0.27	0.60
5	私に好意をもってくれているような人を私も好きになる	**0.38**	0.32	0.27	0.25	0.18
8	注目の的となりそうなときは他の人のそばにいたい。	0.12	0.20	0.15	**0.60**	0.36
16	私を高く評価してくれる人たちのそばにいたい	0.43	0.24	**0.73**	0.36	0.57
19	私を重要で刺激的な人間だと考えてくれる人のそばにいたい	0.42	0.09	**0.75**	0.23	0.57
21	私の人となりを理解してくれる人がそばにいてほしい	0.49	0.30	**0.74**	0.23	0.61
22	ほめてくれないような人とは一緒にいたくない	0.04	−0.06	**0.36**	0.32	0.23
2	活動をするときは他の人と一緒に参加したい	0.34	0.44	0.01	**0.51**	0.41
7	比べることが出来るので他の人のそばにいたい	0.18	0.22	0.22	**0.72**	0.52
12	期待されていることがはっきりしないときは他の人を当てにする	0.21	0.17	0.28	**0.63**	0.41
14	自分と同じ経験をしている人のそばにいたいと願う	0.31	0.26	0.45	**0.58**	0.42
18	どの位できているかを知るために，他の人に注目する	0.18	0.05	0.38	**0.65**	0.49

表７　日本語版の尺度間相関

	刺激	支持	注意	比較
刺激		0.58	0.50	0.33
支持			0.43	0.46
注意				0.46
比較				

こで仮定した構造がこのデータにはやや無理があることを意味しているのかもしれない。3つの例外的な項目（6, 16, 22）については，翻訳上の問題も含めて，それぞれの理由を考えることができるであろう。

なお，4つの主成分によって説明される分散の大きさは，11.9（全分散の45.8%）であるが，表6の解では，11.8（45.3%）で，この点に関するロスはほとんどない。単純和による尺度得点でさえ11.3（43.5%）の分散を説明する（下位尺度を独立変数，個別項目を従属変数とする重回帰分析による）ことを考えれば，この下位尺度がデータ全体の構造を歪めているとは言えないであろう。

以上の結果は，探索的分析のブラインドな適用にとどめず，比較的わずかな理論的情報を補うことによって，理論的にはより解釈しやすい下位尺度を正当化できることを示す例となっていると思われる。

プロトタイプ変換は，内容領域の専門家と分析家とが，ディスプレーに向かって直接対話しながら，尺度開発を進めることのできるような方法であると考える。たとえば，分析の結果，十分な項目数がない下位尺度について，新たな質問項目が考えられないかを実務家を中心に再度検討することも可能であろう。また，分析結果から，2つのプロトタイプが融合したり，2つに分ける必要が生じたりすることもありうる。

8　討論と今後の問題

8.1　方法上の問題点

結果の解釈に当たって，変量の主成分への偏回帰係数の行列である主成分パタンと，変量と主成分との相関行列である主成分構造のどちらを主体とすべきであろうか。斜交回転では，パタンの解釈が薦められる場合が多い（柳井・村上，1999）が，プロトタイプ変換に関する限り，構造行列の解釈を中心としたい。

第1の理由は，この方法の目的が変換された主成分そのものを以後の分

析に用いるというよりは，単純和で定義される各下位尺度に，項目を分類することにあるからである。そのための単純なルールとして，最も相関の高いプロトタイプが定義する主成分によって代表される下位概念に分類するというのは，当面，実用性の高い方法であろう。

第2に，主成分パタンの1つの列が他のプロトタイプをどの項目に設定するかによって影響を受けるのに対し，主成分構造の列は，当該のプロトタイプだけで定まることがあげられる。この方法は，尺度開発者と分析家が対話しながら尺度構成をしていく場面での使用を考えており，その際，尺度開発者側が何らかの理由でプロトタイプを変えること，あるいは（1つの下位概念に対して複数のプロトタイプを考慮している場合に，）プロトタイプを付け加えたり，削除したりすることが考えられる。そのとき，検討中の下位尺度以外の係数が，当該の下位尺度の定義の変更にともなって変化することは，（そのこと自体はデータの重要な特徴であることは認めるとしても，）コラボレーションへの集中を妨げることになりかねない。

なお，不完全プロクラステス法（芝，1979）等，同様の目的をもった方法との比較も今後必要であろう。

8.2　測定の妥当性との関係

項目が対象者に伝えようとする質問文の意味は，日常言語の水準にある。一方，研究者が，尺度を用いて測定しようとする概念の意味は，学術用語によって記述される。この間の区別は決して無視されてはならない。これはたとえば，「あなたの自我同一性の達成度を5段階で評定してください」などという質問が，少なくとも一般人を対象とした場合に成立しないことを考えれば明らかであろう。

したがって，質問文の日常言語としての意味をそのまま学術用語に移行し，それを測定尺度の妥当性の根拠とすることはできない。妥当性の根拠は，何らかの尺度外の規準となる変数との関連の情報であるべきである。それもできる限り尺度と同じ質問応答形式とは異なる水準の，できる限り

外的行動水準を表す変数が望ましいであろう。

しかしながら，項目内容を経験豊かな実務家が判断する場合には，少し事情が異なってくると考えるべきである。たとえば，それが多くの個人を継続的に知る立場にあったカウンセラーであるとすれば，彼／彼女には，個人の日常レベルの言語反応と学術用語による概念とを結びつけるための経験が蓄積されていると考えられる。それは，完璧なものとは言えないまでも，妥当性（内容的妥当性）の根拠とすることはできると思われる。

そうした経験を積んだ人物が尺度の作成者でもある場合，そうした人物と分析家がディスプレーをはさんで対話しながら項目分類を行うことができれば，尺度構成の初期の段階において一定の妥当性を持つ可能性をもった項目群から出発することができる点で，非常に有利になると考えられる。

8.3 終わりに

本研究は，今後の研究プロジェクトの企画書のスケッチのようなものであり，文献的にも，経験的データの上でも十分な裏づけをもつものではない。しかしながら，質問紙法による個人差測定に関して，かなり包括的な問題提起を行うことはできたと考えている。これは出発点にすぎないが，多少とも実りある議論の発端になれば幸いである。

注

1 このアイディアは，故水野欽司教授（統計数理研究所，大学入試センター）が，名古屋大学助教授時代（1975年頃）に，著者に示されたものである。すなわち，「因子軸は，最も意味のはっきりした変数を通るように引くべきである」ということである。ただし教授自身は，この方法を具体的に展開されることはなかったようである。

2 言語学的な研究からも，こうした複数のプロトタイプをもつ概念が存在する証拠が得られているらしい。トンガ語には緑と青をひとまとめにした色彩後があるが，そこには他の言語における緑と青に相当する2つのプロトタイプがあるという。Wikipedia 'Prototype thoery' http://en.wikipedia.org/wiki/Prototype_Theory（final revision: 19 November 2007）による。

文 献

Cox, G.M. (1939). The multiple factor theory in terms of common elements. *Psychometrika*, 4, 59-68.

Fayers, P. M. and Hand, D. J. (2002). Causal variables, indicator variables and measurement scales: *An example from quality of life. Journal of Royal Statistical Society, Series A*, 165, 233-253.

Hill, C.A. (1987). Affiliation motivation: People who need people… but different ways. *Journal of Personality and Social Psychology*, 52, 1008-1018.

堀洋道・山本眞理子（編）(2001) 心理測定尺度集Ⅰ　人間の内面を探る　＜自己・個人内過程＞　サイエンス社

池田央（1973）心理学研究法８　テストⅡ　東京大学出版会

Kline, R.B. (2005). *Principles and practice of structural equation modeling.* 2nd ed. New York: Guilford.

村上隆（1989）心理測定の理論と家族的類似の概念 名古屋大学教育学部紀要—教育心理学科—，36，149-156.

村上隆（1993）家族的類似の概念にもとづくテスト得点のモデル　名古屋大学教育学部紀要—教育心理学科—，40，34-52.

村上隆（2003）測定の妥当性　日本教育心理学会（編）教育心理学ハンドブック　有斐閣 159-169.

中村陽吉（2000）対面場面における心理的個人差　測定の対象についての分類を中心として　ブレーン出版

Rosch, E. (1987). Wittgenstein and categorization research in cognitive psychology. In M, Chapman, and R.A. Dixon (Eds.) *Meaning and growth of understanding: Wittgenstein's significance for developmental psychology.* Berlin: Springer.

芝祐順（1979）因子分析法　第２版　東京大学出版会

柳井晴夫・村上隆（1999）因子パターンと因子構造　繁桝算男・柳井晴夫・森敏昭（編著）Ｑ＆Ａで知る統計データ解析　DOs and DON'Ts サイエンス社 136-139