

構造方程式モデルによるグループ間比較方法の検討

——政治的関心の男女差と MGCFA モデル——

宮 野 勝*

A Consideration on Comparing Group Means through SEM: Political Interest and Gender through MGCFA

MIYANO Masaru

We examined which gender has more political interest in Japan through MGCFA (Multigroup Confirmatory Factor Analysis). Our observed variables are six items of political interest from our Web survey in February 2015. They are self-evaluation of political interest, three types of media use (TV, newspaper, and internet) / people (family and friends).

All six items positively correlate with others and the Cronbach's alpha is 0.772. This relation is similar in each gender. However, among items the female means are lower in four, higher in one than male's and about the same in one. This peculiarity makes the signs of correlation between gender and each item different across items.

As we can expect from the peculiar correlations, when we adapted group mean-structure model of one factor MGCFA, both the configural model and metric model fit sufficiently and Scalar Invariance model does not.

Then we estimated Partial Invariance models. Models fixing one of intercepts of six items fit adequately. The four-items model by dropping two peculiar items also fits well. The difference in means between female and male depends on models.

Since we do not have particular standards to fix one of the intercepts or to drop items, we did not find a persuasive way to choose one model.

We conclude that the concept of political interest has some complexity and we cannot be too cautious in constructing and selecting Partial Invariance models of MGCFA.

キーワード：政治的関心, Political Interest, 構造方程式モデル, SEM, MGCFA, グループ間比較, Group Mean, 性別比較, invariance, 切片固定, メディア利用, 政治的会話

* 中央大学文学部教授

【目次】

はじめに：MGCFA モデルによる異なる母集団間の比較

- 1 MGCFA モデルの簡単な説明
- 2 政治的関心の男女間比較
- 3 分析と考察
- 4 結 論

はじめに：MGCFA モデルによる異なる母集団間の比較

意識調査を用いたグループ間の比較研究は、方法論的には、さまざまな議論が続いており、検討の余地が大きい。

とりわけ国際比較において、「測定の等価性」が問題になる。Davidov et al. (2014) は、この問題をレビューし、MGCFA の利用を推奨している。すなわち、近年の国際比較調査における測定の同等性の研究動向を展望した Davidov et al. 2014 “Measurement Equivalence in Cross-National Research” *Annual Review of Sociology* Vol.40: pp.55-75 では、構造方程式モデル (SEM = Structural Equation Model)、特 に MGCFA (Multi-Group Confirmatory Factor Analysis: 直訳では「多集団確証的因子分析」などとなるが、日本では一般的に「多母集団比較」「多母集団同時分析」あるいは「多母集団・平均構造モデル」などと呼ばれている) を用いたグループ間の比較が、有望な方法であるとしている。

これに対し、Comparative Political Studies July 2016 vol.49-8 では、SEM を用いた異文化グループ間の比較に対して、意見が分かれている。すなわち、SEM モデルによる分析例を示しつつ (SEM を使っていない) 世界価値観調査の分析を批判した Aleman and Woods (pp.1039-1067) に対して、Welzel and Inglehart (pp.1068-1094) は、逆に国際比較調査における SEM モデルの利用に対する批判を返している。

国際比較を含め、グループ間比較の方法論に関しては、決着がついたとは言い難い状態にある。特に国際比較はグループ間比較がその目的であるだけに、この問題は重要である。

第 1 に、一般に複雑なモデルになるほど、適用するための条件が厳しくなりがちであり、SEM モデルも、その一形態である MGCFA も、適用するための諸仮定の充足を前提としている。

第 2 に、諸仮定が充足されたか否かの判定方法、どの程度まで充たされなくても結果を採用できるのか (頑健性の程度)、充たされない場合に取りべき方法、などについて、必ずしも決着がついているわけではない。

第 3 に、SEM あるいは MGCFA モデルの下で、複雑で多様なモデルの展開が可能であり、Davidov et al. (2014) が推奨する適用方法が唯一の方法とはいえない。特に、Partial Invariance という方法とその適用の仕方など、検討の余地は大きいと思われる。

第4に、MGCFA モデル以外にも、IRT モデル、Anchoring Vignettes 法なども、有望な国際比較・母集団比較の方法でありうる（本稿では MGCFA に限定し、MGCFA とその他の諸モデルとの比較は、将来の検討課題である）。

本研究では、第1に、MGCFA について、ごく簡単に紹介する。第2に、MGCFA モデルによるグループ間比較の方法を、政治的関心を対象として、実際のデータで男女間の比較に適用を試みる。第3に、その中で、MGCFA の適用方法について改めて検討する。

本研究の大きな特徴として、われわれが用意した「政治的関心」の6つの観測変数は、相互間の相関は高く1次元性を示しているにもかかわらず、比較を予定している集団である性別のダミー変数との相関が、①有意な正の相関、②有意な負の相関、③有意な相関なし、と3種類に分かれた点をあげられる。質問紙を設計する時点では予想していなかった事態であるが、MGCFA を用いる集団間比較という方法について再考する機会としたい。

本論文は、MGCFA という集団間比較、特に国際比較のために提唱されている方法について、男女比較のデータ分析に使用しつつ、利用可能性について考える試論である。

1 MGCFA モデルの簡単な説明

はじめに、仮定などを確認・検討するため、MGCFA を含む SEM の基本モデルと、MGCFA それ自体とについて、簡単に紹介する。（SEM や MGCFA について既知の方は、この部分はスキップして頂きたい。）

1-1 SEM の基本モデルと諸仮定の確認

SEM はその中に多くの複雑なモデルを包み込む、広汎なモデル群ともいえるが、本論文で扱うモデルは、それほど複雑ではない。SEM の基本モデルを、ごく簡単に記述しておく。

「潜在」変数と「観測」変数

SEM では、「潜在」変数と「観測」変数が登場する。たとえば、動物が喜んでいるかどうか、怖がっているかどうかなど、動物の心の中は直接的には観測できない潜在変数である。それに対し、動物の身体各部の状態・動きは測定可能な観測変数である。われわれは観測可能な動物の姿勢・尾の動き方などを通して、動物の心の中を間接的に推定したりする。

同様に、人間の心理や態度も直接的には測定できない潜在変数である。そのため、観測できる行動や回答を通じて、人間の心の中を測定することが考えられてきた。

本論文では、潜在変数として「政治的関心の高低」を取り上げ、直接には測定できないとする。そして、間接的に複数の質問への回答（X1・X2……），たとえば、メディアで政治関連の記事やニュースを見る頻度や、あるいは、他の人と政治について会話する頻度などに関する回

答を通じて測定することを試みる。

測定方程式

複数の観測変数（政治的会話の頻度など）から、1つの潜在変数（政治的関心など）を推定するモデルを考える。あるいは、潜在変数 ξ （共通因子とも呼ばれる）は、複数の観測変数 X_j （ j は変数番号）を通して測定する。これは、因子分析のモデルに相当するが、SEMでは、モデルのより柔軟な設定が可能になる。個々の回答者番号を i とすると、（平均構造がある）SEMの測定方程式は、次のようになる。

$$X_{ij} = \tau_j + \lambda_j \cdot \xi_i + u_{ij} \quad (1)$$

式(1)は、観測変数ベクトルの観点から見ると、 j 本の連立方程式を意味する。式(1)では、潜在変数 ξ は、各観測変数 X_j に、線形式（ τ_j は切片、 λ_j は ξ の係数または因子負荷、 u は誤差変数または独自因子）で作用すると、仮定している。

各回答者は潜在変数 ξ について特定の値 ξ_i を持つ。この値は、各回答者の観測変数 $X_{i1} \cdot X_{i2} \dots$ に影響を与え、観測される。その際に、各観測変数 X_j ごとに（回答者 i によらない）特定の切片 τ_j と係数 λ_j とを持つと仮定している（ここでは、 X_j は連続変数を仮定する）。

以上をまとめると、潜在変数が1つだけの因子分析的な（ただし平均構造のある）SEMモデルの変数は、①潜在変数 ξ_i 、②観測変数 X_{ij} 、③誤差変数 u_{ij} 、の3種類となり、相互の関係を特定したものが測定方程式(1)である。また、推定すべきパラメータは、切片 τ_j 、係数 λ_j 、誤差分散 $\sigma^2_{u_{ij}}$ となるが、モデル次第で、それらに仮定を置いたり、特定の制約を課したりすることができる。

SEMの仮定

しばしば用いられる仮定は、回帰分析に準じている。以下は、豊田（1998）を参照する。

$$\text{仮定 1: (}\xi\text{についての仮定)} \quad E[\xi_i] = 0 \quad V[\xi_i] = 1$$

潜在変数の「平均値や分散の値そのものを解釈するのは難しい」（同：p.52）ため、平均0、分散1に固定する。ただし「比較の対象がはっきりしている場合」などには、「期待値や分散を推定することもある」。たとえば、多母集団で潜在変数の平均値の比較をする際には、グループ1の平均値のみ0とし、他のグループの平均値は自由パラメータにして推定する。

$$\text{仮定 2: (}u\text{についての仮定)} \quad E[u_{ij}] = 0 \quad V[u_{ij}] = \sigma^2_{u_{ij}}$$

誤差変数は、「微小な影響をもつ多くの要因の和と解釈」（豊田：1998, 53 頁）すること

できる。

($X_1 \cdot X_2 \cdots$ の期待値は、仮定 1・2 を採用する場合、 $\tau_1 \cdot \tau_2 \cdots$ となる。)

仮定 3：(ξ と u の関係についての仮定) $E[\xi_i, u_j] = 0$

通常、誤差変数は潜在変数と無相関と仮定する。

仮定 4：(u_i と u_j についての仮定) $E[u_i, u_j] = 0 \quad (i \neq j)$

通常、誤差変数間では無相関と仮定する。ただし、誤差変数間で相関を持つ可能性はある。たとえば、類似した質問形式、同一変数の時系列データ、などからの回答の場合である。SEM では、そのような特定の誤差項間についてのみ、この仮定を外してパスを設ける。

共分散構造

観測変数の「共分散を母数の関数で表現すること」を「構造化」と呼び、「観測変数の共分散行列の要素のすべてを、測定方程式のパラメータで表現したもの」を「共分散構造」という（豊田：1998, 55 頁を参照。少し表現を変えている）。

調査者が入手できるのは、観測変数の共分散行列であり、特定の測定モデルを仮定すると、モデルのパラメータと観測変数の共分散行列との対応関係を考えることができる。SEM においては、連立方程式モデルで表現されたパラメータについての方程式を、観測された共分散行列を代入して解くことになる。

このとき、「識別問題」が発生する。連立方程式は、常に解を求められるとは限らない。「識別可能」なモデルでなければならず、数学的に「不能」な場合でも推定できるが、「不定」な場合は「識別」できず解を求められない。（豊田：1998, 65 頁）

1-2 MGCFA（多母集団・平均構造）モデルと諸仮定の確認

MGCFA は、SEM のモデルの一種である。グループ間の平均値の比較に MGCFA を利用するには、複数の条件をデータが充足している必要がある。①その条件とは何か、②条件の充足をどのように判定するか、③条件の充足は容易か否か、④条件を充たせない場合にどうするか、という点が問題になる。

MGCFA の方法と条件を、小杉・清水（2014：104 頁）などを参考に、簡単に紹介する。

国際比較などの異なる母集団の間で、潜在変数の平均値を比較するために、しばしば取り上げられるのは、次の 3 条件である。

①「因子構造」の同等性 配置不変 Configural Invariance モデル

各グループの因子構造が等しい

②「因子負荷」の等値性 測定不変 Metric Invariance モデル

各測定モデルの因子負荷 λ_j のグループ間等値制約

③「平均構造」の等価性 Scalar Invariance モデル

各測定モデルの切片 τ_j のグループ間等値制約

これらに加え、さらに条件が付けられる場合もあるが、ここでは省略する（小杉・清水：2014, Millsap & Aguilar：2012 参照）。

Davidov et al. (2014) では、上記の条件を順番に確認し、「Scalar Invariance」まで成立すれば、共通因子の平均値を比較できるとする。すなわち、①～③が成立する場合、1つのグループのどの平均値は0にし、その他のグループのどの平均値を推定して差を比べる。

1-3 モデル間の比較方法

小杉・清水（2014：105 頁）は、尤度比検定とモデル選択基準によって、制約を追加していた異なるモデルを比較することを推奨している。

1 尤度比検定

- (1) $\Delta \chi^2$ = 等値制約が①多いモデルと、②少ないモデルの、 χ^2 値の差
- (2) Δdf = モデル①と、モデル②の、自由度の差
- (3) (1)を(2)で割って検定する：(1)が有意でなければ、①を採択する（等値制約を増やしても適合度が悪化しないことを意味する）。

2 モデル選択基準

- (1) CFI・RMSEA・SRMRなどで、各モデルの適合性を確認する。
- (2) AIC, BIC など、モデル間で比較する。

本研究では、1については、 χ^2 値と df を示して参考とするだけにし、2(1)の CFI など、2(2)の AIC などを使用して、モデル適合度の確認とモデル間比較を行う。

2 政治的関心の男女間比較

使用するデータ・変数を紹介し、単純集計など、予備的な分析を示す。

2-1 使用するデータと、「政治的関心」の「観測」変数

使用するデータは、2015 年 2 月の Web 調査である。（調査の詳細は、Miyano (2015), 宮野 (2016) 参照）

政治的関心はどのような質問への回答に表れるだろうか。この調査で、「政治的関心」の表れとして用意した項目群は、次の7変数である（以下、後尾の括弧内のような略称を併用する）。具体的な質問文は、後掲の表-1を参照されたい。

- | | | |
|----|--------------------------------|-------------|
| X1 | 政治関心の度合いの自己評定（7pt）= q01 | （政治関心の自己評定） |
| X2 | テレビのニュースや番組で政治を知る（5pt）= q02_1B | （TV視聴） |
| X3 | 新聞記事で政治を知る（5pt）= q02_2B | （新聞記事） |
| X4 | ネットの情報で政治を知る（5pt）= q02_3B | （ネット情報） |
| X5 | 家族と政治の話をする（7pt）= q21_1 | （家族と政治会話） |
| X6 | 友人と政治の話をする（7pt）= q21_2 | （友人と政治会話） |
| X7 | 職場の人と政治の話をする（7pt）= q21_3 | |

通常は1の自己評定をもって「政治的関心」の測定とみなしているが、本研究では、「政治的関心」は潜在変数とし、1の自己評定も、多くの「観測」変数の中の一つと考えることにした。

これは、政治情報への接触頻度や他者との政治会話頻度を通じて、自己評定を相対化しようとする試みになるだろう。たとえば回答者が「政治に大きな関心がある」と答えたときに、テレビや新聞やネットで政治情報にまったく触れていない、家族・友人・職場の人などと政治について一切会話していない、などすると、関心の大きさを割り引いて理解した方がよいのではないか、という発想に基づく¹⁾。

これらの質問への回答に、政治的関心の度合いが表れると想定したが、本研究で用いたX1～X6という指標の選択が適切であるか否かは、重要な問題であり、議論の余地がある。

1) 質問紙調査における質問を、3つの基準で分類しておく。

(1) α ：内面の「心理」「意見」「態度」などについての質問

β ：外面の「言動」「行動」などについての質問

(2) α 回答する際の基準が「よく」や「あまり」など「主観的基準」

β 回答する際の基準が「週に2日以上」「月に5冊から10冊」など「客観的基準」

(3) α ：回答者が自分自身に関する事項について答える「自己評定」

β ：回答者が他者やヴィネット上の仮想的な人間について答える「他者評定」

今回われわれが使う、政治的関心に関する観測変数にあてはめてみよう。X1「政治的関心の自己評定」は、「態度」についての「主観的な基準」による「自己評定」である。X2～X7のメディア接触と政治会話は、「行動」についての「客観的な基準」による「自己評定」である。

質問紙調査では、細部まで詰めて回答を求めることは少なく、「行動」や「客観的な基準」についても測定の曖昧さは残る。しかし、「態度」や「主観的な基準」での質問に比べると、ある意味での客観性がある。「自己評定」と「他者評定」の正確さについては、一般的には述べられない。（たとえば、本人にしかわからない情報もあるし、本人の癖や性格や心理は周りの人間の方が知っているかもしれない。自己評定の方が自己合理化が働きやすいかもしれない。）

2-2 データ選択

メディア接触頻度に関する 3 変数は、元来 7pt スケールであるが、接触頻度が少ない 3 カテゴリーは回答が少なかったため、これらを統合し、メディア接触頻度を、7pt から 5pt に変換した。

また、「X7 = 職場の人との政治的会話の頻度」という変数は、無職の人は職場での会話の機会がそもそも存在せずに非該当になる。無職の人々を分析対象から外さないために変数 X7 は、分析から除くことにした²⁾。

また、当初はヴィネット分析との比較を企画していたためにヴィネット分析で使用したケースに限定してあり、かつ X1 から X6 に欠損値があるケースは除く分析方法を選んだため、分析するケース数は全回答数 N = 1493 から減少して N = 991 になっている³⁾。

2-3 質問文と予備的分析：単純集計・平均値・標準偏差・相関係数

具体的な質問文と単純集計 (N = 991) を、表-1-1 ～表-1-3 に示す。表内の数値は%で、各行の合計 = 100% である (丸めの誤差のため、単純和が 100 になるとは限らない)。

表-1-1 政治関心の自己評価 (N = 991) の単純集計

問 1 あなたはふだん、政治にどのくらい関心がありますか。1つだけ選んでください。

	1 とても関心がある	2 関心がある	3 どちらかといえば関心がある	4 どちらともいえない	5 どちらかといえば関心はない	6 関心はない	7 全く関心はない
政治関心 (自己評価)	11.5	25.5	37.5	9.0	11.9	2.9	1.6

(注：以下では、変数名を、q01 とする。)

2) 無職の人も少なくないため、「職場での政治の話」は、省くことにした。「職場での政治の話」については、無職の人は「該当者はいない」を選ぶことになり、欠損値のあるケースとして扱うことになる。欠損値のあるケースを省いて分析すると、有職者のみを対象とする分析になってしまう。手法的には欠損値があるケースを分析に含めることも可能であるが、たとえば多重代入などをつかうにしても、無職の人の「職場の人との政治的会話」を想定することの解釈が難しい。有職者のみの分析にするなど、この変数を含めた分析は、今回は見送る。

3) 2015 年 2 月 Web 調査の回収数は、N = 1493 である。ヴィネット分析との比較を考え、Miyano (2015) でヴィネット分析の対象としたケースに限定することになると、N = 1111 になる。使用した変数の中で、欠損値が存在するのは、政治会話頻度である。今回の分析では欠損値は除くことにしたため、Q21-1 と Q22-2 の欠損値を含むケースを除外した結果、N = 1000 になった。さらに、当初は「Q19：あなたは、「政治に関心を持つべきだ」と思いますか」という質問を分析に含める予定だったため、その欠損値分をさらに減らし N = 991 になっている。この限定の仕方には議論の余地があるだろうが、今回はこのデータを分析する。

表-1-2 メディア利用（N = 991）の単純集計

問 2 あなたは、政治の動きを知るために、テレビ・新聞・インターネットを利用することがどの程度ありますか。それぞれについて、1～7から1つだけ選んでください。

	1 毎日、利用する	2 週に3～6日、利用する	3 週に1～2日、利用する	4 数週間に1回、利用する	5 数か月に1回、利用する	6 ほとんど利用しない	7 全く利用しない
テレビのニュースや番組	53.2	19.8	14.0	4.5	0.9	3.5	4.0
新聞の政治に関する記事	36.5	15.0	11.8	5.2	3.2	14.3	13.8
ネットの政治に関する情報	26.8	18.2	20.7	7.7	3.8	13.9	8.9

（注：これら3問は、6～7を5に合併し、5pt スケールとして用いる。以下では、それぞれの変数名を、q02_1B・q02_2B・q02_3B とする。）

表-1-3 政治会話（N = 991）の単純集計

問 21 あなたは政治や政策や選挙について、次の方と話題にすることはどのくらいありますか。最も近いものを1つだけ選んでください。

	1 ほぼ毎日、話題にする	2 週に数回、話題にする	3 週に1回、話題にする	4 数週間に1回、話題にする	5 数か月に1回、話題にする	6 ほとんど話題にしない	7 全く話題にしない
家族	7.6	19.1	15.9	16.6	15.1	17.5	8.2
友人	0.4	4.9	8.9	15.7	18.7	30.9	20.5

（注：以下では、それぞれの変数名を、q21_1・q21_2 とする。）

これらの政治的関心項目の相互の相関係数を表-2に示す。相互の相関は高く、因子分析（抽出は最尤法）でも1因子になり、クロンバックの $\alpha = 0.772$ （どの項目を落としてもこれより下がる）である。

表-2 政治関心関連項目間の相関係数（N = 991）

	q01	q02_1B	q02_2B	q02_3B	q21_1	q21_2
q01 Self_PI	1					
q02_1B TV	0.406	1				
q02_2B Paper	0.429	0.477	1			
q02_3B Net	0.458	0.290	0.255	1		
q21_1 Discuss_Family	0.416	0.372	0.360	0.287	1	
q21_2 Discuss_Friend	0.378	0.270	0.310	0.286	0.493	1

表-3に、政治的関心の各項目について、男女別の平均値と標準偏差とを示す。また、これらの変数と性別との相関係数を示す。

表-3 観測変数の男女別の平均値と標準偏差, 性別との相関係数

2015-02		q01	q02_1B	q02_2B	q02_3B	q21_1	q21_2	N
Female	Mean	3.28	1.95	2.83	2.98	3.8	5.35	570
	S.D.	1.34	1.22	1.5	1.34	1.79	1.45	
Male	Mean	2.61	1.86	2.35	2.35	4.22	5.04	421
	S.D.	1.25	1.12	1.45	1.24	1.76	1.43	
F-M	Mean の差	0.67	0.09	0.48	0.63	-0.42	0.31	
	S.D. の差	0.09	0.10	0.05	0.10	0.03	0.02	
性別との相関係数		.245**	0.036	.157**	.234**	-.118**	.104**	991

(注：平均値が小さい方が，相対的に政治的関心が高く，メディアに接触し，政治的な会話をする.)

2-4 予備的解釈：7つの観測変数に関する男女差の有無

以上の情報をもとに，これらの質問と回答について，予備的な解釈を試みる。

表-2 から，用意した6項目は，相互の相関は高く，また特に q01 の政治的関心の自己評価との相関も高く，因子分析や信頼性尺度の検討からも，いずれも政治的関心の指標として適切であると思われる。

男女別でも，結果はほぼ等しい。相互の相関は高く，最尤法で抽出した因子分析は1因子になり，クロンバックの α は，女性で 0.788，男性で 0.752 で，どの項目を落としてもこれより下がる。通常は，これら6変数で尺度を構成して研究を進めることになる。

しかし，同時に，表-3（観測変数の性別の平均値と，観測変数と性別との相関係数）より，われわれの6つの観測変数（いずれも自己評定である）を，男女差の観点から，次の3種類に分けられる。

A 群：男女差がない（性別との相関が有意でない）項目

X2 TV ニュース視聴

B 群：男性の方が政治関心ありの項目

X1 政治関心の自己評定：男性の方が関心ありと回答

X3 新聞の政治記事：男性の方が読む

X4 ネットの政治情報：男性の方が読む

X6 友人と政治の話：男性の方がする

C 群：女性の方が政治関心ありの項目

X5 家族と政治の話：女性の方がする

A 群・C 群よりも B 群が多いことから，「女性の政治関心は男性よりも低め」に「見える」が，

同時に、「回答は、観測変数（質問項目）に応じて、性差の影響を受けている」と考えられる。

また、A 群・B 群・C 群のすべてが存在することから、X1 ～ X7 の観測変数の男女差は、潜在変数である政治関心の差だけでは説明できないと予想される。

たとえば、英語のテストでは、集団 1 の方が読解問題の平均点が高く、集団 2 の方がリスニングの平均点が高いというようなことが生じうる。あるいは、読解ないしリスニングのみに限定しても、内容によって（評論か小説か、文系の内容か理系の内容か、など）、集団 1 と集団 2 の優劣が逆転することも起こりうる。英語力を測ると言っても、複数の質問を用意する以上、全く同じことを測定しているわけではないためである。

通常の質問紙調査でも同様のことが生じる。測定しようとしている潜在変数は同じでも、用意した観測変数に対して、潜在変数以外の要因が影響を与える余地は大いにある。それだけ、「項目（観測変数）の用意」が重要だということを意味する。

今回用意した観測変数の特殊性として、可能性としてはあるが、

- ①女性は（潜在的政治関心が同程度の男性と比べ）新聞で政治面を読む率・ネットで政治情報を得る率が少し低い。
- ②女性は（潜在的政治関心が同程度の男性と比べ）テレビで政治ニュース・番組を見る率は同じくらいである。
- ③女性は（潜在的政治関心が同程度の男性と比べ）家族とは政治の話をするが、友人とは政治の話はしない傾向がある。

潜在変数の平均値で説明できない部分は、テレビと新聞の利用の違い、家族と友人との会話の違いなど、グループごとに観測変数の平均値に表れている可能性がある。

なお、政治関心の自己評定についても、

- ④女性は（潜在的政治関心が同程度の男性と比べ）政治に関心があると答えない傾向がある、という可能性もある。

そして、上記のような場合、平均値の比較のための Scalar Invariance の成立が懸念される。測定方程式の切片をグループ間で固定して適合的なモデルを構成できるか、MGCFA を適用して検討する。

2-5 モデルの推定

推定には、R（version 3.5.3：2019-03-11）で、Lavaan 0.6-3 package を用いる。Lavaan については、たとえば小杉・清水（2014）を参照されたい。また、適宜、AMOS（vn25）でも同様の分析を試みたので、必要に応じて参照する⁴⁾。

4) Lavaan と AMOS とで、パラメータの推定値には差を見いだせなかったが、適合度の指標などでは、微妙に異なる数値も見られた。たとえば、CFI はすべて一致したが、RMSEA は異なる場合があった。

3 分析と考察

3-1 基本モデルと基本モデル内での比較

MGCFA では、順に複数のモデルを検討するが、最初にそれらの基本となるモデルを考える。最初に適合度の高いモデルを作れなければ、MGCFA に入る前に話は終わる。また最初選んだ、または最後に残った、観測変数の項目群によって、潜在変数の平均値の比較は、大きな影響を受ける可能性がある。特に、集団間で「平均値の多少が逆転する項目（先の B 群・C 群）」の存否とその割合は、重要であるかもしれない。集団間で「平均値の差が大きく異なる」観測変数群では、全切片の等値仮定は難しいのではないかと予想している。

以下では、これらの点を考慮しつつ複数のモデルを用意し、モデルの適合度を比較し、グループ間の潜在変数の平均値の差を検討する。

A：基本モデル 1

潜在変数： ξ = 政治関心

観測変数：X1 = 政治関心自己評価, X2 = TV 視聴, X3 = 新聞記事, X4 = ネット情報,
X5 = 家族と政治会話, X6 = 友人と政治会話

B：基本モデル 2

基本モデル 1 に、次を加える。

共分散①：X2 (TV) - X3 (新聞), 共分散②：X5 (家族) - X6 (友人)

図-1 に、基本モデル 2 を示す。

基本モデルにおいて、政治的関心それ自体は、直接には測定できない潜在変数とする。潜在変数は、6 つの観測変数に影響を与える、ないし反映される、と考える。

ただし各観測変数は、潜在変数のみで決定されるとは限らない。また観測変数ごとの独自性の問題が存在する。測定を線形式を考えると、各観測変数の独自性は、各式の「切片」や、潜在変数 (= 共通因子) にかかる「係数 (= 因子負荷)」に表れるであろう。

今回のモデル選択には影響しないと判断したため、基本的に Lavaan の結果を紹介する。

また AIC は、定義が異なると思われ、全く異なる値が出た。ただし、同系列のモデル間での AIC の差は一致したため、今回のモデル選択には影響しないと判断した。たとえば、M1 から M3 の AIC は、AMOS_25 では、O6・M1 で 153.6、O6・M2 で 53.8、O6・M3 で 65.8 で、モデル間の AIC の差は、Lavaan の結果と、3 桁目まで一致した。

AMOS の AIC は「 χ^2 値 + $2 \times$ (推定パラメータの数 q)」で定義される。この値は、各モデルの自由度 df と観測変数の分散・共分散の数ならびに平均値の数から推測可能であるため、表には Lavaan の AIC を掲載する。

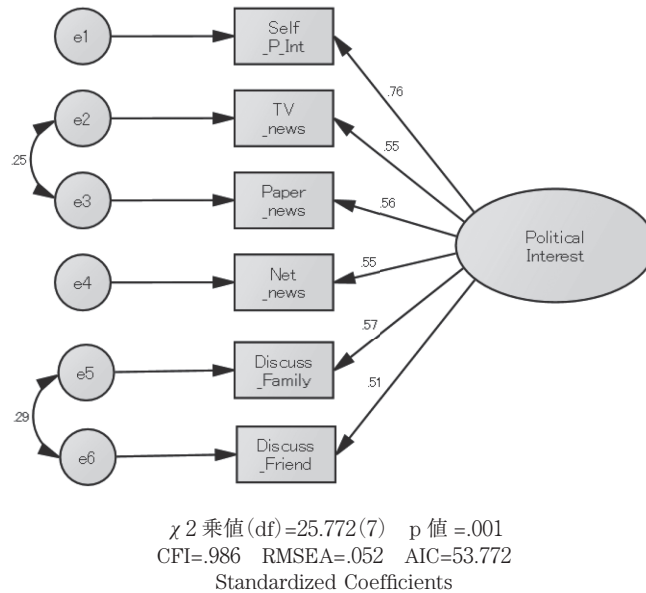


図-1 基本モデル 2 の標準化解（Amos25 の出力より）

各項目の誤差成分（＝固有因子）間の共分散＝0を仮定するか否かは議論の余地がある．基本モデル 1 は誤差項間の共分散なしのモデルであり，各項目の誤差成分（＝固有因子）間の共分散＝0を仮定する．しかし，① X2：TV 視聴と X3：新聞記事，② X5：家族と政治会話と X6：友人と政治会話は，それぞれ，質問に類似性があり，観測変数間の相関も高く，潜在変数以外から何らかの影響を受けていると思われる．このため基本モデル 2 では，それぞれにゼロでない共分散を想定した．

基本モデル 1・2 のモデル評価を試み，どちらかを採用できた場合には，その先の検討を進める（どちらも不適格の場合，検討はそこで中断し，新たな基本モデルを考えるなどの検討をすることになる）．

Lavaan における基本モデル 2 の式を示す．

$$P_Int_2015 = \sim q01 + q02_1B + q02_2B + q21_1 + q21_2$$

$$q02_1B \sim q02_2B$$

$$q21_1 \sim q21_2$$

観測変数 6（O6 と示す）で，性別による集団構造を入れていない（グループが 1 つという意味で GP1 で示す）基本モデルの推定結果を表-4 に示す．この中で，「O6・M3：GP1・平均

構造モデル」は、基本モデル 2 に、平均構造を入れたモデルである（推定するモデルに通し番号をつけ M1 ～ M12 で示す。観測変数の数を優先したため、本文に現れる順ではない）。

表-4 基本モデルの推定結果の適合度：観測変数 6 個

		潜在変数 平均値差	χ^2 乗	df	p	CFI	AIC	BIC	RMSEA	SRMR
O6・M1	GP1・基本 モデル 1		129.7	9	0.000	0.913	19802.4	19861.2	0.116	0.050
O6・M2	GP1・基本 モデル 2		25.80	7	0.001	0.986	19702.5	19771.1	0.052	0.024
O6・M3	GP1・平均 構造モデル		25.80	7	0.001	0.986	19714.5	19812.4	0.052	0.021

M1 と M2 を比較する。「モデル適合度の指標（CFI, RMSEA, SRMR）」の値でみて、M1 はいずれも適合度が低いが、M2 はいずれも適合的である。「モデル比較の指標（AIC, BIC）」の値でみても、M2 の方が優れている。

M2 モデルに、MGCFA の前提となる平均構造を入れた M3 モデルも試みた。適合度の指標はほぼ M2 と等しく適合度は高い。モデル比較の指標では、M1 より優れているが、M2 よりは悪化している。

M2 モデルは、適合度もよく、モデル間比較でも M1 モデルより優れていたため、基本モデルとする。M3 は M2 より AIC・BIC では劣るが、平均構造の分析を目的とするため、以下では、M3 モデルを基本として、集団間比較に進む。

3-2 MGCFA（多母集団平均構造）モデルによる集団間比較

基本モデルがデータに適合したので、MGCFA で集団間比較を試みる。グループは男性と女性の 2 グループである。モデルとして検討するのは、次の M4 ～ M6 の 3 モデルである。これらのモデルでは、それぞれ集団間での重要な仮定を検討することになる。そして M4 が適合すれば M5 へ、M5 が適合すれば M6 へと順に進む。推定結果を表-5 に示す。

O6・2 集団・平均構造モデル

O6・M4：配置不変 = Configural Invariance モデル

O6・M5：測定不変 = Metric Invariance モデル

O6・M6：全切片固定 = Scalar Invariance モデル

表-5 MGCFA による集団間比較：観測変数 6 個のモデル

		潜在変数 平均値差	χ^2 乗	df	p	CFI	AIC	BIC	RMSEA	SRMR
O6・M4	GP2・配置 不変モデル		21.68	14	0.085	0.995	19532.4	19728.3	0.023	0.019
O6・M5	GP2・測定 不変モデル		25.70	19	0.139	0.995	19526.4	19697.9	0.027	0.025
O6・M6	GP2・全切片 固定モデル	〈0.522〉	176.17	24	0.000	0.892	19666.9	19813.8	0.113	0.071

M4 は、2 集団間の「配置不変」モデルである。M4 は、適合度の指標においても、モデル比較の指標でも、M3 のみならず M2 も凌駕している。性別という集団間での「因子構造の同等性」の仮定をおくことは、適格的である。

そこで、MGCFA の手順に従い、次の「測定不変」の M5 モデルに進む。M5 では、「潜在変数の係数」＝「共通因子の因子負荷」が集団間で等しいという強めの等値制約を課して推定する。適合度の指標は、M4 とほぼ等しく適格的である。モデル比較の指標については、AIC も BIC も M4 よりも優れている。性別という集団間で「因子負荷の同等性」を仮定したモデルは適格的である。

さらに手順を進め、次に「切片不変」の Scalar Invariance の M6 モデルに進む。M6 では、測定方程式の各観測変数の切片を集団間で同一の値に固定してモデルの適合性をみる。適合度の指標は、いずれも M4・M5 よりも大きく悪化しており、M6 モデルは適格的とは言えなくなっている。モデル比較の指標についても、AIC も BIC も M4・M5 より悪化しており、M6 モデルは採用できない。性別という集団間で「測定方程式の切片の同等性」を仮定したモデルは適合しなかった。2-4 の「予備的解釈」でも懸念したが、この仮定には無理があるようだ。

しかし、このままでは、集団間で潜在変数の平均値を比較できない。

「Scalar Invariance」は、必ずしも成立しない。不成立の場合、「観測変数の平均・分散・相関などの差が、共通因子以外の要因に影響を受けている」（H_SEM：Ch23, 381）ことになるという。

一般に、「観測変数の平均・分散・相関などの差が、共通因子のみに影響を受けている」という仮定は、アンケート調査に適用するには、強すぎる場合も少なくないのではなかろうか。グループ間（性別、あるいは、文化別など）で、特定の観測変数への回答のグループ差が、共通因子にすべて集約されるという仮定が、どの程度まで一般的に成り立つだろうか。

この仮定を比較の前提にすると、グループ間の差が共通因子にすべて集約されるような集団・観測変数に関してしか適用できなくなる懸念がある。たとえば、男女間でメディアの利用の仕

方が異なる場合、その相違の程度によるが、グループ間でメディア利用の切片固定はできず、共通因子の平均値を比較できないということが生じうる。

3-3 Partial Invariance モデルによる集団間比較

Davidov et al. (2014) では、「切片固定」の条件を充足できない場合、観測変数や比較するグループの一部を落とし、「Partial Invariance」モデルとして比較することを提案している。

そこで次に、「Partial Invariance」モデルを試みる。

今回は、そもそも 2 グループしかなく、グループの一部を落とすわけにはいかないため、観測変数の一部を落とすことを考える。ただし、「Partial Invariance」モデルはより多様な考え方ができる。観測変数やグループを落とすことはせず、切片の一部のみ自由パラメータとするモデルも可能ではある。

そこで、共通因子の平均を推定するために、次の 2 つの方法を試みる。

- (1)切片等値可能な観測変数を探し、「切片を等値できない観測変数は分析から捨てる」。これは、通常、使われている方法であろう。具体的には、平均値が同じ方向にずれている 4 つの観測変数 B 群のみを使うことを考える。すると、「切片の等値制約モデル」が成り立つかもしれない。
- (2)切片等値可能な観測変数のみ等値固定し、「切片を等値できない観測変数は自由パラメータにする」。モデルによるであろうが、一部の切片を固定したモデルが適合すれば、潜在変数の平均値を比較できるかもしれない。切片をすべて自由パラメータとすると潜在変数の平均値を比較できないので、一部の切片のみ固定する。ただしこの方法では、どのパラメータを固定するかで、平均値が大きく変わる可能性がある。

(1)まず Partial Invariance モデルとして、観測変数を落とす前者を試す。複数の変数を残すためには、異質な変数すなわち用意した変数が 1 変数ずつしか含まれない A 群・C 群の観測変数を外し、4 変数が含まれる B 群の変数を残すことにする。

その場合、基本モデル M2 で仮定していた共分散に含まれる 2 変数が落ちるため、共分散なしモデルと、新たに、「新聞記事」と「ネット情報」に共分散を仮定したモデルとを試してみた。表-6 で、それぞれの「切片固定」モデルの結果のみ紹介する。

共分散を導入しない M11 と導入した M12 を比較する。M11 よりも M12 の方が、モデルの適合度でも、モデル間比較の指標でも優れているため、この 2 つのモデルの中では、M12 を採用する。M12 においては、「切片固定」モデルが適合した。すなわち、Partial Invariance が成立し、平均値が比較可能になった。

M12 モデルは、Davidov et al. (2014) が推奨している標準的な方法に従うものであり、通常

表-6 観測変数4個の全切片を固定したモデル間比較

		潜在変数 平均値差	χ^2 乗	df	p	CFI	AIC	BIC	RMSEA	SRMR
O4・M11	男性高・切片 固定・共分散0	<0.676>	22.87	10	0.011	0.977	13269.3	13357.5	0.051	0.035
O4・M12	男性高・切片 固定・共分散1	<0.646>	11.39	8	0.181	0.994	13261.8	13359.8	0.029	0.027

はこの結果を採用してグループ間の比較を行うこととなると思われる。

(2)次に、後者の方法を試す。M6の「全切片固定」モデルが不適合だった理由は、グループ間で値の「差」が大きく異なる（とM5測定不変モデルから推測される）観測変数の切片まですべて含めて「切片等値」を仮定したためであろう。

この問題を回避するため、観測変数群の切片のグループ間の差が、相互に大きくは異ならない観測変数（観測変数のA群・B群・C群に分け、それらの1群内の変数）の切片のみ固定することを考えてみた（M7～M10）。次のように、どの変数を固定するかで、大きく4つの場合に分け、4つのモデルを試みる。

O6・M7：同一方向（男性が関心高）の4切片固定（Partial Invariance）モデル

O6・M8：男性が関心高の1切片固定（Partial Invariance）モデル

O6・M9：関心に差なしの1切片固定（Partial Invariance）モデル

O6・M10：女性が関心高の1切片固定（Partial Invariance）モデル

具体的には、M7では「政治関心の自己評価」「新聞記事」「ネット情報」「友人と政治会話」の切片を固定した。M8では「政治関心の自己評価」、M9では「TV視聴」、M10では「家族と政治会話」、の切片をそれぞれ固定した。表-7に、結果をまとめた。

第1に、M5の測定不変モデル、M8～M10の4つのモデルの間では、モデル適合度の指標（CFI, RMSEA, SRMR）の値がすべて等しく、モデル比較の指標（AIC, BIC）の値も全く等しい。M7の4切片固定モデルは、上記の4つのモデルと比べ、適合度は少し悪化しているが十分に適合の許容範囲であり、AICは少し高いがBICは少し低い。このため、これら5つのモデル（M5, M7～M10）のうち、どれを選択すればよいのか、指標だけでは判断が困難である。

第2に、M7～M10の4つのモデルで、推定される「潜在変数（共通因子）の平均値の差」が、-0.427から+0.664までであり、大きく異なっている⁵⁾。

5) 潜在変数の平均値は、比較した限りでは、LavaanとAMOSとで値は等しかった。ただし、潜在変数からのパスの係数一つ固定する必要がある、表-7では、固定する切片と同じ観測変数について

表-7 観測変数 6 個の一部の切片を固定したモデル間比較

		潜在変数 平均値差	χ^2 乗	df	p	CFI	AIC	BIC	RMSEA	SRMR
O6・M7	男性高の4切片 固定モデル	〈0.659〉	35.73	22	0.032	0.990	19530.4	19687.2	0.035	0.033
O6・M8	男性高の1切片 (Q1) 固定	〈0.664〉	25.70	19	0.139	0.995	19526.4	19697.9	0.027	0.025
O6・M9	性差なし1切片 (TV) 固定	〈0.086〉	25.70	19	0.139	0.995	19526.4	19697.9	0.027	0.025
O6・ M10	女性高の1切片 (家族) 固定	〈-0.427〉	25.70	19	0.139	0.995	19526.4	19697.9	0.027	0.025

モデルの適合度の点でもモデル選択の観点からも、いずれも適切と思われるモデルになったが、適合度の高いモデル間で、潜在変数の平均値が大きく異なり、どのモデルを選択すべきか、迷うことになった。

あえてこれらの中で選ぶとすると、用意した多くの変数（6 変数中で 4 変数）を固定した M7 モデル、全切片を自由パラメータとした測定不変モデルでグループ間の差が最小だった切片のみ固定した M9 モデル、のいずれかになろうか。

仮に、用意した 6 つの変数のうち、2 つは指標として不適切であると判断するならば、M7 モデルを選ぶことになる。また、その場合、他の 2 つの変数の切片は自由パラメータとして推定される。

観測変数の中で、もともと性別との相関に有意差が無く、測定不変モデルにおいても性別の平均値の差が最も小さかった、「テレビ視聴の頻度」変数を固定すべきであると判断するならば、モデル M9 を選ぶことになる。

確かに、方法(1)のように、グループ間の差が、ほぼ「等しい」観測変数群（グループ間の差がある観測変数で d なら他の観測変数でもほぼ d）を集めれば、「切片等値」モデルがデータに適合し、Partial Invariance が成り立つかもしれない。

しかし、第 1 に、「ありうる観測変数全体」を見渡しているどうか、という大きな問題がある。相関の在り方が異なり不適合と判断された観測変数を外すことの理論的な理由だけでなく、他のありうる観測変数を除いた理論的な理由を示すことが求められよう。またこのことは、第 2 に、

のパスの係数を 1 に固定する場合を示した。Bollen (1989: 366-7) では、このような指定を用いている。なお筆者は Lavaan で同一の変数の切片とパスのみを固定する方法を見いだせていないため、M9・M10 モデルについては、潜在変数の平均値のみ AMOS の値を用いた。O6 の 1 切片固定モデルでは、切片を固定する変数と係数を 1 に固定する変数を一致させるか否かで、値は変わる。

モデルの成立-不成立に、ワーディングや翻訳なども大きな効果を及ぼす可能性を示唆する。

Millsap & Aguilar (2012 : 383-384) は、Scalar Invariance が成り立たない場合の、2つの困難に言及している。第1に、モデル不適合を起こしている発生場所の特定の難しさである。第2に、モデル不適合を発生した項目を除いたモデルで、当初の目的を達成できるか否かの判断である。そして「文献では、ガイダンスがほとんどない」、また、「いくつかの不適合を許容するか、簡単なルールはない」としている⁶⁾。

4 結 論

国際比較のための方法を探す一環として、推奨されている方法の一つである MGCFA の適用可能性を検討した。

Davidov et al. (2014) などによる、① Scalar Invariance を充たせば比較できる、②充たさない場合は、グループまたは観測変数を落として Partial Invariance として比較する、という考え方が、どの程度まで実際のデータ分析で妥当な結論を生むか、を考察した。例として、具体的な様相を相対的に理解可能である「国内の性別」という2グループ間での政治的関心を取りあげ、比較を試みた。

今回の暫定的な結論を示す。

1) MGCFA は、適用に際して、モデルの細部の適切性など、考慮すべき点が少なくない。

今回のデータでは、観測変数内の相関は高く、1次元であり、同一の潜在変数の測定とみなせる。しかし、観測変数と「グループを表すダミー変数」との相関係数の正負が一定ではなかった。それが原因ではないかと推測しているが、測定不変モデルにおいて自由パラメータとして推定した各観測変数 X_j について、グループ間の切片の差 d_j も、相互にほぼ等しい観測変数群と、相互の差が大きい観測変数とに分かれた。MGCFA の利用に際して、このような場合への対応方法は十分には検討されていないのではなかろうか。

2) MGCFA の国際比較データへの適用例（たとえば Jilke et al. : 2015）を見ると、Scalar

6) Millsap & Aguilar (2012 : 389-390) は、3つの代案がありうるとしている。

第1に、一般的な共通因子の他に、一部の変数にのみ影響する group 因子を用意する2因子モデルである。

第2に、順序カテゴリ用の因子モデルである。

第3に、IRT モデルである。順序カテゴリ用の因子モデルはある種の IRT モデルと等価だが、IRT はそれに尽きるものではないという。

しかし、第1の提案は魅力的だが、そのような因子を見つけることは、容易でないかもしれない。第2・第3の提案は、本研究では、今回は検討の範囲外である。また、管見では、たとえば Davidov et al. (2014) や Jilke et al. (2015) などを見る限り、IRT などによって問題が解決したというよりも、別の方法として試されている段階のように思われる。

Invariance は達成が難しい場合も少なくないかもしれない。また達成された場合でも、モデルや観測変数が適切に選択されているか吟味することは不可欠になる。

今回の M12 モデルで使われていた 4 つの観測変数のみが最初に用意されていたならば、Scalar Invariance が達成されたということで、平均値差をそのまま採用したと思われる。今回は、他に 2 つの観測変数があったため、先の 4 変数のみで検討を終わらせて良いのかという問題に気づくことができた。この点は、事前にも事後にも潜在変数の定義や観測変数群を吟味することの望ましさを示唆している。

3) Scalar Invariance を達成できずに、Partial Invariance として使う場合も、注意が必要である。

3-1) 「一部のグループを分析から除外する」とき、たとえば国際比較で最初に用意した国の一部を省いて分析するのでは、本来の目的が達成できない場合も生じうる。

3-2) 「一部の観測変数を分析から除外する」とき、当初適切であるとして選ばれている変数や、あるいは、本来含まれるべき変数が、分析（やプリテスト）の結果などで除外されている場合に、その理由を理論的に説明できるかどうか、重要になる。

一般に、一部の観測変数を除外する場合、類似した傾向がある観測変数群が残され、他と異なる観測変数が除外されることになる。しかし、一部のグループや観測変数を分析から外す方法は、比較できるものに偏りが出るおそれがある。そもそも比較が難しいものにこそ、差異が隠れている可能性がある。

今回の分析では、一案としては、従来の適用法に従い、観測変数を落として 4 つにしたモデル M12 を選ぶことが考えられる（類似する M7 もありうる）。他方、観測変数の中で測定不変モデルにおいて性別の平均値の差が最も小さい（あるいは、もともと性別との相関に有意差が無かった）「テレビ視聴の頻度」を固定した（しかし用意した観測変数を一切落とさない）モデル M9 を選ぶことも考えられる。

しかし決め手となる論拠は不充分と判断した。M12（または M7）と M9 のモデル間で、推定された男女間の平均値の差も大きく、結論は出さないことを選んだ。政治的関心の定義や内容の問題に立ち返って、（英語テストの内容を細分化するように）概念を細分化していくなどが必要になるのかもしれない⁷⁾。

今回の検討は試論であり、方法的にも実質的にも検討しきれない問題が多く残っている。

7) この点も、さらなる検討の余地はある。「政治」的関心といってもより多様な関心のあり方も考えられるためである。たとえば、国政には関心があるが地方政治には関心がないとか、候補者には関心があるが政策には関心がないとか、そのような場合の政治的関心の高低をどのように判断するかなど、概念の再検討も必要になる。この点では Coffe (2013) が参考になる。

謝辞：本論文は、中央大学 2017 年度特定課題研究の成果の一部である。2018 年 2 月に収集した Web データは分析が進んでいないため、2015 年 2 月に収集した Web データを用いた。

参考文献

- Aleman, Jose and Dwayne Woods 2016 “Value Orientations from the World Values Survey: How Comparable are they cross-nationally?” *Comparative Political Studies* July 2016 vol.49-8: 1039-1067.
- Bollen, Kenneth A. 1989 ‘Structural Equations with Latent Variables.’ NY John Wiley & Sons.
- Coffe, Hilde 2013 “Women Stay Local, Men Go National and Global?: Gender Differences in Political Interest.” *Sex Roles* September 2013 Vol.69 5/6, 323-338.
- Davidov et al. 2014 “Measurement Equivalence in Cross-national Research.” *Annual Review of Sociology* vol.40: 55-75.
- Jilke, Sebastian, Bart Meuleman, Steven Van de Walle 2015 “We Need to Compare, but How? : Measurement Equivalence in Comparative Public.” *Administration. Public Administration Review* January-February 2015 36-48.
- 小杉考司・清水裕士 2014 『M-plus と R による構造方程式モデリング入門』北大路書房 特に、第 8 章 多母集団同時分析 103-116 頁。
- Millsap, R.E. and M. Olivera-Aguilar 2012 “Investigating Measurement Invariance Using Confirmatory Factor Analysis.” In “Handbook of Structural Equation Modeling” ch.23: 380-392.
- 宮野勝 1998 「不公平は「江戸の華」」豊田秀樹編『共分散構造分析【事例編】』北大路書房：第 1 章 1-8 頁。
- Miyano, Masaru 2015 “Measuring Political Interest Using Anchoring Vignettes: Empirical Evidence from Japan.” 『中央大学社会科学研究所年報』19：33-45.
- 宮野勝 2016 「「政治的関心」の測定—ヴィネット基準点法（Anchoring Vignettes Method）を用いて—」宮野勝編『有権者・選挙・政治の基礎的研究』中央大学出版部。
- 豊田秀樹編著 1998 『共分散構造分析：入門編—構造方程式モデリング—』朝倉書店。
- Welzel, Christian and Ronald F. Inglehart 2016 “Misconceptions of Measurement Equivalence: Time for a Paradigm Shift.” *Comparative Political Studies* July 2016 vol.49-8: 1068-1094.

