

# 国際比較／文化比較調査における 測定と比較可能性の確認のための統計的技法\*

——多集団確認的因子分析と確認的最小空間分析——

真 鍋 一 史\*\*  
前 田 忠 彦\*\*\*  
清 水 香 基\*\*\*\*

## I. はじめに

本稿は、真鍋、前田、清水（2021）の続編である。そこで、前稿では何をしたか、そして本稿では何をしようとしているか、について説明するところから始める。

まず、前稿で何をしたかという、それはつぎのようにまとめることができる。国際比較／文化比較調査の実践とその方法論的な研究への関心の高まりにともなって、「測定の等価性／不変性」というテーマをめぐって多くの理論的・実証的・方法論的研究がなされるようになってきた。そこで、前稿では、このような研究領域におけるさまざまな研究の内容とその系譜を、伝統的な「文献研究」の手法にもとづいて概観するとともに、国際比較／文化比較というコンテキストにおいて、これまで最もよく用いられてきたとされる「多集団確認的因子分析（multigroup/multiple group confirmatory factor analysis：MGCFA）」に焦点を合わせ、そのアイデア・技法・手続きについて解説した。

では、本稿では何をするのかというと、それ

は、国際比較／文化比較調査における「測定の比較可能性（measurement comparability）」——別の表現をとるならば、「測定の等価性／不変性」——を実証的に確認（confirm）するための統計的諸技法に関する検討、具体的にいうならば、「多集団確認的因子分析」と「確認的最小空間分析（confirmatory smallest space analysis）」との方法論的な比較検討、ということである。

そこでさらに、では、なぜそのような方法論的検討を行なうのか、そして、なぜMGCFAとSSAに焦点を合わせるのか、という「問い」が出てくる。

まず、後者については、国際比較／文化比較調査における「測定の比較可能性」の実証的な確認のための統計的技法としては、さまざまなものが開発されてきている。例えば、Davidv et al.（2014）では、つぎのようなものがあげられている。

- ・探索的因子分析（exploratory factor analysis：EFA）
- ・多集団確認的因子分析（multigroup confirmatory factor analysis：MGCFA）
- ・多次元尺度法（multidimensional scaling：MDS）
- ・項目反応理論（item response theory：IRT）

\*キーワード：測定の比較可能性、国際比較／文化比較調査、統計的技法、多集団確認的因子分析、確認的最小空間分析

\*\*関西学院大学名誉教授、青山学院大学名誉教授、統計数理研究所データ科学研究系客員教授

\*\*\*統計数理研究所データ科学研究系准教授

\*\*\*\*北海道大学大学院文学研究科助教

・潜在クラス分析 (latent class analysis: LCA)

では、このようなさまざまな技法のなかから、なぜ本稿において、因子分析法の系列のなかから MGCFA を、そして、多次元尺度法の系列のなかから確証的 SSA を、それぞれ選び出すことになったかという、じつは、このような選択をめぐっては、この研究領域における、二人の「先達」を中心とする研究事例があったからにはほかならない。それは、つぎの二つである。

〈1〉 Shalom Schwartz et al. による一連の「人びとの価値観モデル」の構成 (construction) に関する研究——「環状連続体モデル」から「ヒエラルキカルな三層構造モデル」へ——

ここで、前者のモデル構成においては SSA が、そして後者のモデル構成においては CFA (確証的因子分析) / MGCFA が、それぞれ用いられた。

〈2〉 Shaul Oreg et al. による一連の「人びとの変化に対する資質的な抵抗尺度」の開発研究

この研究は、心理尺度の開発研究において、CFA/MGCFA と確証的 SSA が併用された先駆的な研究事例であるとされている。

これら二つの研究事例においては、このように、MGCFA と SSA が併用されて、独自の研究が進められた。本稿の国際比較／文化比較調査における「測定の比較可能性」の探究は、これら「先達」の足跡を踏まえて計画されたのである。以上が、本稿において、MGCFA と SSA に焦点を合わせるようになった理由である。

つぎに、前者の、なぜこれら二つの統計的技法について方法論的検討を行なうのか、という「問い」に対して答えなければならない。

そのような「答え」の第1は、どのような統計的技法にも、それぞれ「利点」と「限界」があり、それらの問題を越えてデータ分析を進める——それは、いかえれば、「国際比較／文化比較研究における方法論的な質の向上」(Steenkamp and Baumgartner, 1998, p.88) をめざすことといえる——ためには、複数の技法を併用することが必

要であり、そのような併用の具体的な方略を検討することが課題となるというものである。これは、いうまでもなく、きわめて重要なポイントといわなければならない。そして、そのためには、これまでに開発されてきたさまざまな統計的技法の方法論的検討が不可欠のテーマとなってくる。

じつは、このような方法論的検討についても、別の「先達」による試みがある。まさに「先達はあらまほしきものなり」というべきであろう。それは、Braun and Johnson (2010) による、An Illustrative Review of Techniques for Detecting Inequivalences である。この文献においては、国際比較／文化比較調査における測定の比較可能性の確認のために用いられる諸技法——「基礎的な技法 (basic techniques)」から「先端的な技法 (advanced techniques)」にいたる諸技法——のさまざまな特徴 (features) を、いくつかの基準 (criteria) によって整理する試みがなされている。

まず、検討のために取りあげられた諸技法は、つぎのとおりである。

- ・ 諸項目ごとの「度数分布 (distribution)」「平均値 (mean)」「相関係数 (correlation coefficient)」などの基礎的な技法
- ・ 「信頼性係数 (Cronbach's alpha)」
- ・ 「探索的因子分析 (exploratory factor analysis: EFA)」
- ・ 「多次元尺度法 (MDS)」
- ・ 「多集団確証的因子分析 (MGCFA)」
- ・ 「多水準モデリング (multilevel modeling)」
- ・ 「項目反応理論モデル (IRT models)」

そして、これらの諸技法について、以下のような点からアセスメントを行なう。

- ① その技法が、測定の比較可能性の迅速な概観 (quick overview) を可能にするか。
- ② 多くの国ぐに／文化を、同時に取り扱う (handle) ことを可能にするか。
- ③ 測定の比較可能性に問題のある国／文化を識別する (identify) ことを可能にするか。
- ④ その技法が、国ぐに／文化についての総合的な「要約測度 (summary measures)」を提供することが可能か。
- ⑤ その技法が、項目のレベルにおいて、あるい

は、テストのレベルにおいて、測定の比較可能性を取り扱うか。

以上のような検討の結果を一覧表の形で取りまとめたのが、つぎの表1である（ここでは、本稿での問題関心に合わせて、MDSとMGCFAの比較に絞って、もとの表を作り直している。なお、MDSの第2カラムは、large numberをどう捉えるかで、議論が分かると考えられる）。

表1 MDSとMGCFAの有利な点と不利な点

	MDS	MGCFA
Quick overview	Yes	No
Handling of large number of countries	Difficult	Difficult
Identification of individual cases	Yes	No
Summary measures across countries	No	Yes
Item or test level	Item	Both

確かに、Braun and Johnson (2010) による以上のような統計的技法の整理は、これまでに開発されてきた諸技法の「全体像の把握」という点からして、きわめて重要な試みといわなければならない。しかし、われわれのめざすところは、このような諸技法の「全体像の把握」というところにとどまるものではない。われわれがめざしているのは、前稿で述べたように、そのつぎの段階の「新しい飛躍の種の発見」である。では、そのような「発見」は、いかにして可能となるであろうか。じつは、われわれは、上述の〈1〉と〈2〉の二つの研究事例の「再検討」をとおして、そのような契機がもたらされることになるのではなかろうか、と考えている。

以上の二つの研究事例は、この領域における「研究の発展」が、substantiveな「理論的考察 (theoretical consideration)」と「統計的技法 (statistical technique)」との出逢いをとおして着実にもたらされるものであることを示唆している。そして、そのような「研究の発展」が、前者の場合は「価値観理論 (モデル) の構成」、そして、後者の場合は「心理尺度の開発」という形でなされたという点が注目されるのである。

こうして、以上において、「本稿で何をするか」を明らかにしてきた。それは、この研究領域における「新しい飛躍の種の発見」であり、そのため

に、上述の先行研究事例の詳細な再検討を行なうということである。そのような再検討は、後者の研究事例から始める。つまり、本稿では後者の Shaul Oreg. et al. による研究事例の方法論的な再検討を行なう。前稿と同様、本稿は、このようなテーマで共同研究を続けてきた3人の共同執筆の形で構成される。I、II B、III (1) (2) (3)、1、2、3 B、C、IV1、2、3、5、6 (1) (4) (5) (6) は真鍋、II A-2、III 3 A、IV4、6 (2) (3) は前田、II A-1、IV6 (2) (3) は清水によって原稿が準備され、共同討議の繰り返しをとおして、現在の形へとまとめあげられてきたのである。

## II. 多集団確証的因子分析と確証的最小空間分析

### A. 多集団確証的因子分析

#### A-1. MGCFAの利用・測定不変の必要条件・その結果

「多集団確証的因子分析 (MGCFA)」とは、確証的因子分析 (CFA) を開発したことで知られる Jöreskog (1969) が、1971年の論文 (Jöreskog 1971) で提案した「複数の標本に同一の観測変数を適用して得られたデータにおいて、これらの複数の標本から同一の因子構造が得られるかどうかを検証する」(清水和秋, 1994) ために開発された技法——当初 (Jöreskog 1971) は「複数標本における同時因子分析 (simultaneous factor analysis)」と称されていた——である。以下ではまず、国際比較研究の文脈において、MGCFAが取り上げられることになってきた背景ともいえるべき研究上の関心について述べ、この技法から示される様々な分析の結果が持つ「意味」について解説を試みる。

1. MGCFAが国際比較／文化比較調査における「測定の等価性／不変性」を問題とするものであることは、すでに前稿で述べた通りである。国際比較研究の文脈において、上記の問題は大きく次のような二つに分けて捉えることができる。一つは、因子分析の技法によって、複数の観測変数のセットから単一ないしは複数の潜在変数 (つまり因子) を取り出すという場合に、ある国・文化に

において特定の因子構造が確認されたとして、それと同様の因子構造を他の国・文化においても想定 (assume) できるだろうか、また、それはどのように確認できるだろうか、という問題である。もう一つは、もし仮に、同一の観測変数のセットから共通の因子構造を持つ概念が取り出せたとして、そうした概念における得点 (score) の大小について、国・文化間で意味のある比較 (meaningful comparison) をすることができるのかどうかという問題である。

2. しかしながら、MGCFA の技法を用いることで、どのような調査データを扱う場合でも上に述べたように関心を満たすことができるかという点、そうではない。MGCFA は CFA のいわば延長線上において開発が進められてきた技法であり、したがって通常の因子分析モデルと同様、使用するデータの性質として、次のような条件が求められる (Cohen 2008)。

- (1) 使用される観測変数が連続変数であること
- (2) 諸変数の分布が多変量正規分布 (multivariate normal distribution) に従うこと
- (3) 観測変数と潜在変数 (因子) の関係が線形であること

このうち例えば (1) のような仮定が満たされない、2 値あるいは順序尺度水準の観測変数が用いられている場合は、多集団因子分析の文脈であるか、それとも単一集団のデータへの適用であるかを問わず、順序尺度の場合にはポリコリック相関を用いた分析を行うとか、2 値の場合には項目反応モデルのような適切なモデルを適用する必要がある (豊田 2014 など参照)。国際比較の文脈での順序尺度の観測変数の応用事例は、例えば Davidov (2018) に見ることができる。

(2) の多変量正規分布の仮定については、観測変数が連続変数であるようなケースについて、パラメータ推定法に最尤法を用いる際に必要となるものだが、この仮定のデータへのあてはまりを検討するための方法や、この仮定からデータが逸脱している (あてはまっていない) 際の対処法については、例えば単一集団の文脈では狩野・三浦 (2002) を参照することが有益であろう。国際比

較研究の文脈でのこの仮定からの逸脱に関しては、より複雑な検討が必要となる可能性がある。

(3) の仮定の適切性について取り扱った Bauer (2005) は、観測変数が通常の因子分析モデルとは異なり潜在変数の 2 次関数であるというモデル (あるいは仮定) の下で得られるデータに、通常のように線形の関係を仮定した (適切でない) 因子分析モデルを適用した場合に、集団間での等値性が誤って否定されてしまうケースの起こり易さについて、シミュレーション研究により評価している。具体的には、もし観測変数と潜在変数の関係が集団間で同じであったとしても、両者の関係が非線形的であり、因子平均が集団間で異なっている場合には、線形モデルを当てはめて得られる切片と因子負荷量に違いが生じてしまい、不変性が認められないという結果になることがあることを示している。

3. データが以上の条件を満たしている場合、MGCFA によって集団間 (あるいは時点間) における「測定の等価性/不変性」の検討を行うことができる。MGCFA では「測定の等価性/不変性」を「configural invariance」「metric invariance」「scalar invariance」の 3 つのレベルに区別して捉える。言うまでもなく、それぞれ、統計的にどのような条件を満たした時にそれが確認され、その結果から分析者がどのような意味を読み取ることができるのか (どのように解釈すべきか) ということが異なってくる。

#### (1) Configural Invariance

異なる複数の集団において、ある測定の道具 (measurement instrument) を構成する諸項目 (観測変数) の、「顕著な因子負荷量 (salient factor loadings)」と「顕著ではない因子負荷量 (non-salient factor loadings)」の配置 (configuration) が同じである場合、configural invariance が認められる (Steenkamp and Baumgartner 1998)。上記の「顕著ではない因子負荷量」は、かならずしもゼロに制約しなければならぬというものではないが、慣行的にゼロとして制約されることが多いようである。要約的な説明の仕方がとられる際に

は、「因子構造が集団間で同一である」(Davidov et al. 2014)、「すべての集団が因子負荷量について同一のパターンを有している」(Cieciuch et al. 2014) などといった表現が用いられる。

Configural invariance は、構成概念(抽出された因子)の基本的な意味と構造を、国や文化を跨いで調査するといった場合に、ある構成概念が国や文化を超えて同じように概念化できるかどうかを確認する上での第一の条件となる。Configural invariance が認められることで、はじめて構成概念の因子構造が国や文化を超えて共通であると言えるようになる。ただし、それが認められたとしても、構成概念が観測変数に対して持つ効果(あるいは意味)の大きさが、集団間で異なっているという場合が考えられる。したがって、より高いレベルの不変性が確認されるまでは、定量的な比較を行うことは控えるべきである(Steenkamp and Baumgartner 1998)。Configural invariance の有無についての検定は、概念構造の国際比較/文化比較という文脈において、それ自体が重要な分析的関心を有するものであることは言うまでもないが、それと同時に、より高いレベルの測定の等価性/不変性の検定へと分析の歩を進めていく上でのベースラインとしても位置付けられることになる(Davidov et al. 2014)。

## (2) Metric Invariance

Metric invariance は、上記の configural invariance において求められる「顕著ではない因子負荷量」の配置が集団間で等しいという制約に加え、「顕著な因子負荷量」についてパラメータの等値制約を置くことで検定される。なぜならば、因子負荷量というものは、構成概念の得点の差異が、観測変数の得点の差異にどのように関連しているかについての情報を有するものだからである(Steenkamp and Baumgartner 1998)。したがって、もし集団間で因子負荷量が等しいということが確認されれば、それは構成概念の尺度間隔(intervals of scale)が等しい——つまり、構成概念における1単位の増加がそれぞれの観測変数に対して持つ効果が、集団間で等しい——と言えるようになるのである(Davidov et al. 2014, Cieciuch et al. 2019)。

Metric invariance が認められたとすれば、その項目の得点の差(difference score)が持つ意味について、国や文化間で意味のある比較をすることができるようになる。もう少し具体的に言えば、metric invariance が認められた項目と、他の外部変数との非標準化回帰係数や共分散について、比較できるようになるということである(Davidov et al. 2014, Cieciuch et al. 2014; 2019)。ただ、研究者の関心によっては、国や文化ごとの、観測変数や構成概念の得点の平均値比較を行いたいということもあるだろう。しかし、Metric invariance が確認されたというだけでは、まだこうした比較を行うことはできない。というのも、観測変数の切片(仮に構成概念の得点がゼロだったときの観測変数の値)が国ごとに異なっている可能性があるからである。したがって、この時点ではまだ、もし集団間で平均値の差が認められたとしても、それが切片の差によって生じたものなのか、それが構成概念の平均値の差によって生じたものなのかの区別をつけることができないのである。

## (3) Scalar Invariance

Scalar invariance は、集団間で、因子負荷量だけでなく、観測変数の切片も等しいという制約を置くことで検討される。Scalar invariance が認められたとすれば、異なる集団に属する回答者であったとしても、構成概念の得点が等しければ、観測変数の値もまた等しいということが期待される(Davidov et al. 2014)。したがって、集団間での観測変数の平均値の差は、集団間の構成概念の平均値の差によって生じているものであると理解することができるようになるのである(Steenkamp and Baumgartner 1998)。

以上のような「測定の等価性/不変性」の条件が満たされることによって、ようやく次にあげるような「意味のある比較」を行うことができるようになる(Steenkamp and Baumgartner 1998, Davidov et al. 2014, Meinting et al. 2020)。

- ・ 集団間での構成概念の素点の比較
- ・ 集団間での構成概念の平均値の比較(例えば、国ごとのランキングを作成するなど)
- ・ 集団間で平均値が比較可能であることを前提としたマルチレベル分析

ここまで解説を行ってきた *configural invariance*, *metric invariance*, *scalar invariance* のすべてが認められなければ、集団によって測定の尺度がそもそも異なっているのだから、構成概念や観測変数の得点の比較をしたとしても「意味がない (meaningless)」ということになる。したがって、もし集団間で何らかの構成概念の平均値比較を行うおうとするならば、少なくとも1つの構成概念につき、上記3つのレベルの「測定の等価性/不変性」の条件を満たす、2つ以上の観測変数を用意することが求められる (Steenkamp and Baumgartner 1998)。

しかしながら、Meintinger et al. (2020) によれば、*scalar invariance* が認められることは稀であるという。その理由の一つとして、伝統的な MGCFA による検定の方法が「厳密 (exact)」を求めすぎており、「厳格 (strict)」になされすぎることが指摘されている。こうした文脈において、近年、「ベイジアン近似的等価性検討アプローチ (Bayesian approximate approach)」や「Alignment 最適化によるアプローチ (alignment optimization approach)」が注目を集めるようになってきたことは、すでに前稿で述べたとおりである。

## A-2. MGCFA による分析の具体的な手順

1. MGCFA を利用した測定不変性 (MI) の検討の具体的な手順についても簡単に述べておこう。検討手順を提案した複数の文献があるが、複数の文献で多くの場合その手順はフローチャートのような形で提案される。基本的に、より制約が緩い [1] *configural invariance* モデルの適合度を検討する段階 (あるいは、さらにその前段階) から、[2] 因子負荷の値にのみ集団間の等値制約を入れる *metric invariance* モデルを経て、一番制約が厳しい [3] 切片項も集団間で等しい *scalar invariance* モデルへと制約を強めていく方向で分析を進める、という説明が一般的である。想定されるモデル間の比較のために考えられる複雑な分析の流れについては、例えば Steenkamp and Baumgartner (1998, p.83) の図1が包括的であり、参照して欲しい。モデル内のパラメータの等値制約を緩めるべきか否かの判断には、修正指標

(*modification index*) という統計量が用いられることも多い。モデル間の比較のためには、SEM で一般的に用いられるモデルの全体的な適合度指標を参照すればよいが、国際比較調査の文脈で推奨されているのは、CFA や RMSEA といった指標である。例えば、Cieciuch et al. (2019, p.161) は、これらの指標の値の変化量に基づく判断基準を提示している。一般にサンプル・サイズが比較的大きい状態で分析を行うことが多い国際比較調査では、カイ2乗統計量を利用した尤度比検定で特定のモデルの適合に関わる帰無仮説が棄却されてしまうケースが多く、またネストしたモデル、特に制約をより厳しくしたモデルと元のモデルの間のカイ2乗統計量の差の検定でも、制約の厳しいモデルが (有意に適合度が下がったとの判断で) 支持されないということが生じやすい。こうした事情もあり、サンプル・サイズの影響を比較的受けにくい適合度指標が好んで利用されているものと考えられる。

2. [1][2][3] のようなモデル比較は、比較的定型的に進められる分析ではあるが、それぞれの結果によって判断が分かれていくような、検討プロセスの総体について記述しているものは必ずしも多くない。その中で Cieciuch et al. (2019) の概観は有益であろう。若干の注記を施す目的で、筆者 (前田) が同論文中の図を翻訳・翻案したものが、図1である。

図最上部の、「完全な測定不変性の検証」の部分について、ここでは [3] の *scalar invariance* がテストされることを示しているが、前段階として [1] [2] の検討があることは言うまでもない。したがってその部分に詳細なプロセスが含まれることになるが、前述のように Steenkamp and Baumgartner (1998) により補足することができる。この部分を通過する (*scalar invariance* が満たされる) と、右下の「実質科学的な (*substantive scientific*) 検討・分析を続ける」ことが可能になるが、これは集団の間での因子平均値の比較が可能になることを意味していよう。

他方、これが満たされない場合の直下の四角内では、部分的な MI について述べているが、ここにも、「部分的に (特定の項目についてのみ) 厳

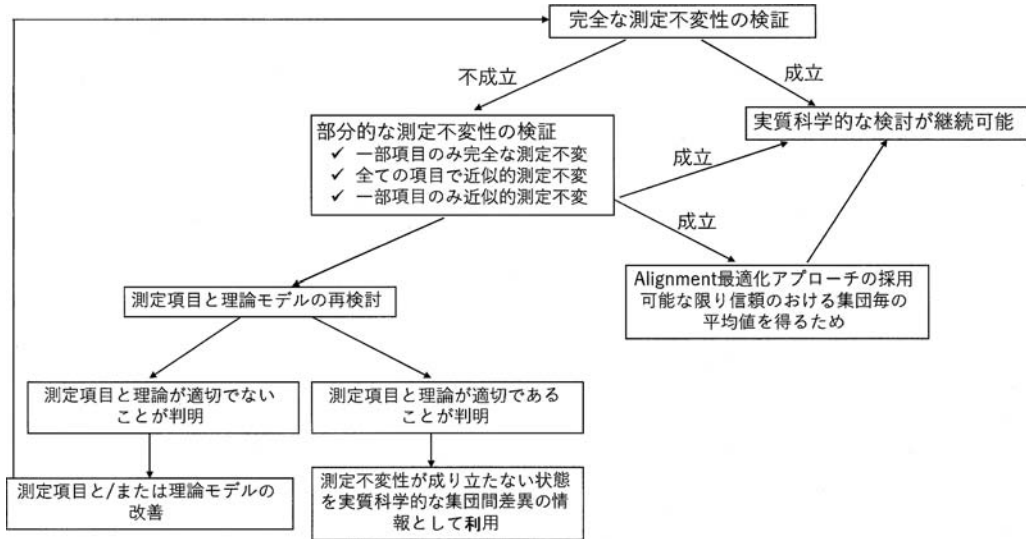


図1 測定不変性を検討するためのプロセスの概観 (Ciecuch, Davidov, Schmidt, and Algesheimer, 2019 を翻訳・翻案)

密な MI が成立している」か、「全ての項目で近似的な MI が成り立っている」か、「部分的に、近似的な MI が成立している」か、といったいくつかの段階があり得ることを意味している。最後のパターンが最も制約が緩い状態になっている。この条件がある程度よい状態で満たされていれば、再び集団間で因子平均値を比較するような実質的な分析が続けられることになる。

「部分的な測定不変性」も満たされていない場合は、測定道具（具体的は尺度に含まれる質問項目）と適用しているモデルの妥当性を検討せよ、場合によっては、項目を改良して（新しい調査を行ったデータを得て）最初の検討に戻るべし、とされている。

3. 最後に、MGCFA による検討に関するいくつかの留意点を上げよう。一つ目の留意点は、集団間の等値制約を確証的因子分析のどのパラメータにまで適用するかという点に関連している。純粋に MGCFA という統計的手法を検討する立場からは、さらに制約が強いモデルも検討することは可能かつ自然だという点である。scalar invariance モデルに加えて、さらに [4] 集団間で各変数の測定モデルにおける誤差分散が等しいという制約を入れるモデル、あるいは、[5] 複数の因子（構成概念）がある場合には、因子間共分散の行列が

集団間で等しいという制約を入れるモデル、などが考えられる。実は [5] の仮定は、[4] のような仮定を置いた後にさらに追加されるという順番で課されるとは限らず、[2] の因子負荷行列への制約の次に考えられるべき制約であるとも考えられる（例えば、狩野・三浦（2002）では [4] と [5] のどちらを先に検討してもよいと説明している）。[5] の仮定は、metric invariance の後に、複数の構成概念（因子）間の共変動（因子が共有する成分とでも言い換えられよう）が、集団間で等しいという制約であり、この意味で複数の構成概念の性質がその相互関係も含めて集団間で共通であるという状態を表現したモデルである。成立する因子分析の性質の類似という意味では重要な検討視点であると思われる。

MI の文脈で [5] の仮定が必ずしも論じられないのは、[5] は複数因子を持つモデルを検討する際に生じるモデルである一方、国際比較・文化間比較の文脈では、測定不変性は個々の構成概念ごとに検討される、すなわち一因子のモデルで検討されるという文脈が多く、因子間“共分散”のパラメータがそもそも生じない、という事情によるであろう（なお、ⅢA でレビューする Oreg et al. (2011) では、複数因子の MGCFA での MI の検討が行われており、この指摘に該当しない）。なお、測定項目が定まった一因子モデルの場合に

は [1] の *configural invariance* モデルの意味合いもやや分かりにくく、現実には、「非ゼロの母数の位置が等しい」ではなく、「各項目が当該因子に大きな負荷量を持つ（その意味で非ゼロである）」という主旨の検討が行われることになる。

他方、[4] のモデルが国際比較・文化間比較調査の MI の文脈で必ずしも検討されない明確な理由は不明である。この文脈での MI は [3] の *scalar invariance* の時点ですでに成立が危ぶまれる状態であることが多く、さらに制約を強めたモデルを検討することは実用面で不可能と見なされているといった事情であろう。一因子モデルで検討される場合に [5] の検討はなく、[1] [2] [3] [4] の順に制約を強めたモデルを仮定することがおそらくは自然であるが、[4] の制約の要否の検討は、しばしば不要とされ、[3] の成立を以って、因子平均の集団間比較が可能であると見なされているのは前述のとおりである。Cieciuch et al. (2019) や Vandenberg and Lance (2000) はこの第 4 の制約についても言及しているが、前者は通常は最初の 3 つの制約の検討までで十分である旨を述べている。この注意を踏まえた上で **II A-2** の以下の記述でも、検討の範囲を基本的に [1] [2] [3] の範囲までに限定しよう。また一因子モデルでの検討とする。したがって [2] での制約は「因子負荷行列」ではなく「因子負荷ベクトル」への制約との表現を採用する。

二番目の留意点は、すでに述べたとおり、伝統的な MGCFA による検定の方法が「厳密 (exact)」を求めすぎており、「厳格 (strict)」になされすぎるという指摘に関係している。分析手続きに即して言えば、等値制約は、理想的な状態としては因子負荷ベクトル、切片であれば切片のベクトルの全ての要素が集団間で等値されたモデルの成立だけを検討すること、一般には *full metric invariance* とか *full scalar invariance* などのように呼ばれる完全な不変のモデルを考えることだけでは十分でない（ことが多い）という点である。一部のパラメータベクトルの一部の要素のみ全集団間で等値可能だが、別の一部のパラメータベクトルについては等値可能でない、といった部分的等価性 (*partial invariance*) のみ成立するモデルの

許容可能性を検討しなければならないケースが多い。この *partial invariance* の問題へのアプローチについて、バイズの近似的等価性アプローチと、Alignment 最適化のアプローチという二つの有力な方法を、真鍋・前田・清水 (2021) において紹介したので参照されたい。

さらに、二番目の留意点を敷衍すると、部分的等価性の許容についても、集団の数が多くなれば、特定の変数（調査項目）に関するパラメータ（因子負荷または切片）の中のどれが、どの集団とどの集団では等値可能であるか、というパターンの組み合わせが増えてしまい、詳細に検討することは現実的に難しくなる点にも留意が必要であろう。例えば、2 集団であれば、一つのパラメータについてであっても、 $\lambda_{(1)} = \lambda_{(2)}$  か、 $\lambda_{(1)} \neq \lambda_{(2)}$ （ここでの (1) (2) のような添字はグループを表すものとする；以下同様）のどちらかの可能性しかないが、4 集団あれば、[1]  $\lambda_{(1)} \sim \lambda_{(4)}$  が全て等しい (1 通り)、[2]  $\lambda_{(1)}$  から  $\lambda_{(4)}$  のうち一つだけ仲間はずれ (4 通り)、[3] 4 集団のうちの 2 つのみ等しく、残る 2 集団とも、相互に等しくない、つまり 3 つのグループに分かれる (6 通り)、[4] 4 集団のうち 2 つずつ等しい 2 群に分かれる (6 通り)、[5] 全ての集団で相互に異なる (1 通り) のような、多くの可能性がある。集団の数が増えると、この可能性の数は膨大なものになり、さらにこれがパラメータベクトルの個々の要素に対して考慮可能なことから、理論的には比較可能なパターンは 1 つのパラメータについて可能な数のべき乗の形で増大し、悉皆的に比較検討することは実質的に不可能である。現実的な *partial invariance* の検討では、こうした膨大な可能性の検討は省かれ、全ての国（集団）の間での等値の仮定が成立するか否かを、パラメータごとに検討する、といった折衷の方法が採られることが多いようである。なお、すでに前稿で紹介した Alignment Optimization のアプローチ (Asparouhov and Muthen, 2014) は、こうした問題を実用的なレベルで解決する点でも刮目すべき提案と言える。

## B. 最小空間分析

ここでは「ガットマン・スケール」によって名



を馳せた、Louis Guttman によって開発された「最小空間分析 (SSA)」と呼ばれる統計的技法についての解説を試みる。真鍋は、この技法については、すでにさまざまな形での解説を行なっている (真鍋、1993, 2001, 2002, 2003, 2016, 2021 ほか)。そこで、本稿では、その記述は、必要最小限の事項に絞ることとする。

1. SSA は、通常の統計的技法と違って、それを単独のデータ分析の技法として捉えることが困難である。それが、単にデータ分析のための道具にとどまるものではないからである。いいかえれば、SSA という技法は、Guttman の「ファセット・アプローチ (facet approach)」と呼ばれる方法論的な枠組みのなかに位置づけられて、はじめて意味のある (meaningful) ものとなるからである。この点をめぐっては、この研究領域における日本の第一人者の一人と目される狩野裕 (2002 b) が、Guttman の方法論的な立場を、確証的因子分析を導入した Karl Jöreskog のそれと対照させながら、つぎのように述べていることが注目される。

「Jöreskog は、探索的な因子分析モデルの基本は崩さず、因子に関する仮説を推定に活かす検証的なパラメトリックモデルを構築し、現在の構造方程式モデルへと発展させた。それは洗練された方法論であったが、ファセットのような質問紙の作成からデータ解析までを含む一般的な枠組みではなかった。ファセットはモデルを用いた推測ではなく、あくまでも記述的な方法論の枠組みで検証的かつ頑健な方法をめざしたものであるといえよう。」(pp i ~ ii)。

では、狩野のいうところの「質問紙の作成からデータ解析までを含む一般的な枠組み」としての「ファセット・アプローチ」がどのようなものであるかという、それについては、Samuel Shye (1978) による整理が、これまでの Guttman の方法論研究において、最も「要領を得たもの」といえる。その証拠に、Guttman 自身も、その多年の研究の成果について語る際には、この Shye の提

案した「枠組み」に沿う形をとっている (Guttman, 1980)。それは、「ファセット・アプローチ」を「ファセット・デザイン (facet design)」「ファセット・アナリシス (facet analysis)」「ファセット・セオリー (facet theory)」の三つ領域からなる「知の体系」として捉えるというものである。このような「ファセット・アプローチ」の全体像の詳細な解説については、真鍋による上記の関連諸文献を参照されたい。ただ、「ファセット・セオリー」については、後述する SSA による分析結果である「SSA マップ/図 (diagram)」の「読み取り/解釈 (interpretation)」の仕方との関連で、やや詳細に述べておかなければならない。

2. 繰り返しになるが、ファセット・アプローチは、ファセット・デザイン、ファセット・アナリシス、ファセット・セオリーの三つの領域から構成される。そして、SSA という統計的技法は、ファセット・アナリシスの重要な部分をなすものとして位置づけられるのである。

#### (1) ファセット・デザイン

①観察 (つまり、質問紙調査) のための概念枠組みの準備、②質問文と回答の形式——scalable question items (尺度化可能な質問項目) と rating method (評定法)——の選択、③調査の仮説的図式を文章の形で表現する独自の技法であるマッピング・センテンス (Mapping Sentence) の構成。

#### (2) ファセット・アナリシス

仮説検証型のデータ分析の技法、例えば、「尺度分析 (Scalogram Analysis: Scale Analysis)」「部分スケログラム分析 (Partial Order Scalogram Analysis: POSA)」「最小空間分析 (Smallest Space Analysis: SSA)」「中央値回帰分析 (Median Regression Analysis)」などの開発。

#### (3) ファセット・セオリー

質問紙調査に対する回答として捉えられる人間行動の諸法則とその理論的根拠の定式化: 「第1の法則」「第2の法則」「多調回帰の法則」などの構築。

i) 第1の法則

第1の法則とは、「態度 (attitude)」や「関与 (involvement)」などの人間行動については、それぞれについての諸項目間の関係は単調関係を示し、相関係数はプラス (あるいは、せいぜいゼロ) となり、マイナスにはならないというものである。

例えば、政治学の領域でなされてきた人びとの政治関与に関する調査研究では、「ある仕方では政治に関与する人は、ほかの仕方でも政治に関与する傾向がある」という知見 (finding) が見出され、そこから「政治関与の累積性」という経験的一般化 (empirical generalization) が導かれてきた (Milbrath, 1965=1976) が、これも政治学の領域の固有の法則というよりも、Guttman の第1の法則の一つの事例にすぎないといわなければならない。

さらに、コミュニケーション行動の研究領域で検証されてきた「あるメディアでコミュニケーションをする人は、ほかのメディアでもコミュニケーションをする傾向がある」という命題も、この法則の一つの事例にすぎないと考えられる (真鍋, 1998)。

こうして、社会科学の研究においても、これまで多くの重複研究 (redundancy) がなされてきたことがわかる。第1の法則の定式化によって、このような問題に対する一つの解決策が提示されたともいえるのである。

ii) 第2の法則

第1の法則が、質問諸項目間の関係 (Pearson の「積率相関係数」や Guttman の「弱単調性係数」) がすべてプラスになるというその関係の「(プラス-マイナスの) 符号 (sign)」に関する法則であるのに対して、第2の法則は、その関係の「(大-小の) 大きさ (size)」に関する法則である。この法則が「領域の法則 (Regional Law)」と呼ばれるのは、SSA の描き出す幾何学的形状 (configuration) によって、それら諸項目間の関係の構造が視覚的に空間の領域 (region) として捉えられるからである。Guttman は、多くの大規模な質問紙調査のデータを用いて、さまざまな Regional Laws を構築してきたが、それらはすべてつぎの点から派生してきたものである。質問諸項

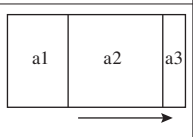
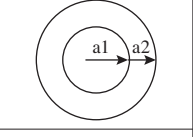
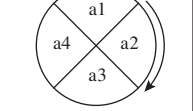
ファセットの役割	→ 空間の分割	
Axial	矩形をいくつかの小さな矩形にスライスするように分割する。	
Modular	共通の原点のまわりにいくつかの同心円を描いて空間を分割する。	
Polar	共通の原点からの区分線が円をいくつかのくさび形 (V字型) に分割する。	

図2 ファセットの役割と regions との対応関係

目の内容 (domain) についてのファセットの諸要素 (elements) は、それと同数の regions に分割される SSA の空間に対応する。ファセット (の諸要素) が空間の分割において果たす役割には三種の種類がある。ファセットがランク・オーダー (rank order: 賛-否、好-嫌、高-低、大-小などの一次元的な順序) をもたないものである場合は polar、ファセットがランク・オーダーをもつものである場合は modular か axial というのがそれである。前者に対応する理論は Circumplex、後者に対応する理論は Simplex と呼ばれる。こうして、このファセットの三種類の役割が組み合わせられて、交差する分割線が cylinder (円筒形)、cone (円錐形)、sphere (球形)、cube (立方体) のような幾何学的な形状を描くことになる。それぞれの形状に対応する理論は Cylandrex、Conex、Spherex、Multiplex と呼ばれる。また modular と polar が組み合わせられた形状に対応する理論は Radex と呼ばれる (図2 ファセットの役割と regions との対応関係、を参照されたい)。

iii) 多調回帰の法則

これは異なる種類 (varieties) の人間行動の相互間の関係についての法則である。具体的にいうならば、intensity (強度)、closure (開閉)、involvement (関与) は、attitude (態度) に対してそれぞれ多調関係となり、順に U (あるいは V) 字型、N 字型、M 字型の回帰 (regression) を示すというものである (ただ、これらの諸法則は、

本稿での議論と直接にかかわるところがないので、ここでは省略する)。

3. 以上を踏まえて、つぎに、このような「ファセット・アプローチ」の一つの領域である「ファセット・アナリシス」の一技法として位置づけられる SSA についての解説に移る。SSA は、多次元尺度法 (multidimensional scaling) の系列に属し、相関マトリックスに示された  $n$  個の変数/項目間の関係を、 $m$  次元 ( $m < n$ ) の空間における  $n$  個の点の距離の大小によって示す方法である。相関が高くなるほど距離は小さくなり、逆に相関が低くなるほど距離は大きくなる。通常は諸変数間の関係を視覚的に描写するために二次元 (平面) あるいは三次元 (立体) の空間布置が用いられる (SSA についての、テクニカルな解説は、Amar and Toledano, 2001 を参照されたい)。

以上は、SSA のテクニカルな側面からする、ごく簡潔な解説である。そこでつぎに、このような記述を構成する諸部分のいくつかの点について、もう少し説明しておきたい。

(1) SSA に用いられるデータは、少なくとも「順序尺度変数/項目 (ordinal variable/item)」でなければならないとされている (Braun and Johnson, 2010, p.385)。

(2) 変数/項目間の関係の測度 (measure) としては、「ピアソンの相関係数ではなく、単調変換による非計量型の相関係数 (単調性係数)」(狩野, 2002 b, p.i) が用いられる。因みに、SSA へのコンピュータ・ソフトウェア・プログラム HUDAP では、「弱単調性係数 (weak monotonicity coefficient)」が用いられる (詳細については、林・鮑戸編, 1976 を参照されたい)。

(3) 以上の「相関係数」——「弱単調性係数」——を用いて、「相関マトリックス」を作成する。これが、SSA 実行のための「入力/オリジナル係数 (input/original coefficient)」となる。これにもとづいて、アウトプットとしての「SSA マップ/図」が作成される。それは、二次元あるいは三次元の「ユークリッド空間 (Euclidian space)」に諸変数——具体的にいうならば、質問諸項目の番号——が印字された「空間布置図 (spacial plot)」である。

(4) では、「SSA マップ」は、もとの「相関マトリックス」を、どのくらい適切に描き出しているであろうか。このような、「相関マトリックス」と「SSA マップ」との「適合度 (degree of fit)」は、「疎外係数 (alienation coefficient)」によって示される。それは、「ストレス指数 (stress index: Kruskal のストレスともいわれる) の一種であり、0 から 1 までの値を取り、0 が「完全な適合度 (perfect fit)」を表わす。Guttman (1968) の経験則 (rule of thumb) では、「評価できる適合度 (evaluating fit)」は 0.15 か、あるいは、それより小さい値とされたが、その後、0.20 までは (Borg and Lingoes, 1987)、あるいは、0.25 まででさえ (Ben-Shalom and Horenczyk, 2003)、「満足できる適合度 (satisfactory fit)」といわれるようになった。

4. では、このような SSA によるデータ分析の結果、つまり「SSA マップ」と呼ばれる「空間布置図」については、どのような「読み取り」がなされうであろうか。これが SSA によるデータ分析の最大のポイントとなる。繰り返しになるが、「空間布置図」は SSA という技法によるデータ分析の「結果」であり、「読み取り」はそのような結果の「解釈」である。実証科学において、「結果」と「解釈」ははっきりと区別されるべきものである。SSA マップにおける、このような両者の違いを、真鍋は、生物学者の福岡伸一のアイデア (福岡, 2010) を借用することによって、つぎのように比喩的に説明している。例えば、夏の夜空に輝く星々をそのままカメラに収めたとするならば、その星々の写真はそのような被写体が撮影された「結果」である。そして、その写真の画面上にいくつかの星座を区分していくとするならば、それは、まさしくそのような結果の「解釈」というものである。天空に輝く星々に、星座という「意味づけ」——つまり「解釈」——を施したものであるからにほかならない。

SSA という統計的技法の中心には、「近接仮説 (contiguity hypothesis)」という考え方がある。そして、質問紙調査というものは、その質問紙 (調査票) で用いられる「言葉の意味」をめぐる実証的な測定 of 技法であり、したがって、そのような

質問紙調査のデータ分析は、まさに調査者と被調査者の両方の側における「意味空間／意味連関」の同一性の探究ということになる。そして、Guttman の基本的な考え方からするならば、調査で用いられる質問諸項目の意味内容が近い場合には、それら諸項目の SSA マップにおける位置（空間的距離）も近いものとなる。そのような「近さ」を手掛かりとして、諸項目の領域区分がなされる。こうして、SSA マップの「空間分割図 (spatial partition)」が完成する。それは SSA マップの「解釈」である。

5. SSA マップの「読み取り」を、以上のように理解しておくとするならば、では、Guttman とその共同研究者は、このような「読み取り」をどのように進めていったのであろうか。ここでは、SSA マップから何らかの「法則性」を発見する「想像力・洞察力・発想力」が要求される。そして、そのような発見の能力に支えられて構築されてきたものが「ファセット・セオリー」にほかならない。しかし、そのような知的営為はそれで終わるわけではない。いったん構築された「ファセット・セオリー」にもとづいて、つぎに、特定の「調査データ」についての「空間布置図」からの「空間分割図」の作成が可能となるのである。これは、実証科学の領域における「知的営為の循環過程」ともいべきものである。こうして、「(SSA マップにおける) 諸領域 (regions) の間の分割線は、理論から導かれるものでなければならず、統計的な手続き (statistical procedure) の結果であってはならない」(Braun and Johnson, 2010, p.385) という指摘の意味が、具体的に納得されるものとなる。そして、そのような確認をとおして、Guttman の SSA という統計的技法については、それが「ファセット・アプローチ」のなかに位置づけられて、そして、より直接的には「ファセット・セオリー」と関連づけられて、初めて意味のあるものとなるということが、具体的に理解されることとなるのである。まさに、Guttman の方法論は、三位一体の「知の体系」といべきものであろう。

6. 以上において、SSA という統計的技法を、そ

れをその一部に含む「ファセット・アプローチ」と呼ばれる「知の体系」と関連づけながら解説してきた。繰り返しになるが、SSA は「ファセット・デザイン」と「ファセット・セオリー」とを連結する「技法」として見事に完成されたのである。しかし、SSA の技法としての発展は、そこにとどまるものではなかった。その後、Samuel Shye によって、それまでの SSA は「探索的 SSA (exploratory SSA)」と性格づけられるとともに、新たに「確証的 SSA (confirmatory SSA)」の体系的な手続きが開発され、プログラム化されるに到った——プログラム・ソフトは、Faceted Smallest Space Analysis (FSSA), FSSAWIN と呼ばれている (Shye, 1991, 1997) ——。

7. では、「確証的 SSA」は、従来の「探索的 SSA」と、どの点が、どのように、違っているのであろうか。それは、以下の3点にまとめられる (Shye, Elizur, Hoffman, 1994)。

(1) 「確証的 SSA」では、「探索的 SSA」によって示唆された SSA マップの空間領域分割 (space regional partition) が、より客観的 (objective) な仕方で確認される。

(2) 「探索的 SSA」では、そのような「領域区分」が、分析者が①印象論的な検討 (impressionistic examination) を行なうこと、②分割線 (partition line) を手書きで描くこと、によってなされてきたが、その操作がコンピュータ化された (computerized partition) ということである。

(3) その手続き (procedure) は、すでに述べた「ファセット・セオリー」に従って、①axial、②modular (radial)、③polar (angular) の順で、それぞれの空間分割が、事前に「ファセット・デザイン」の形で明細化された (pre-specified) 諸仮説と、どの程度まで一致しているか——空間分割の適合度 (goodness-of-fit) ——が、「分離指数 (separation index)」と呼ばれる数値的な測度 (numerical measure) によって判断される、というものである。①②③の順での、SSA マップのコンピュータによる空間分割の事例が図3に示されている。

「分離指数」は、それぞれの領域 (region) から逸脱した (deviant) 諸変数／項目の数 (num-

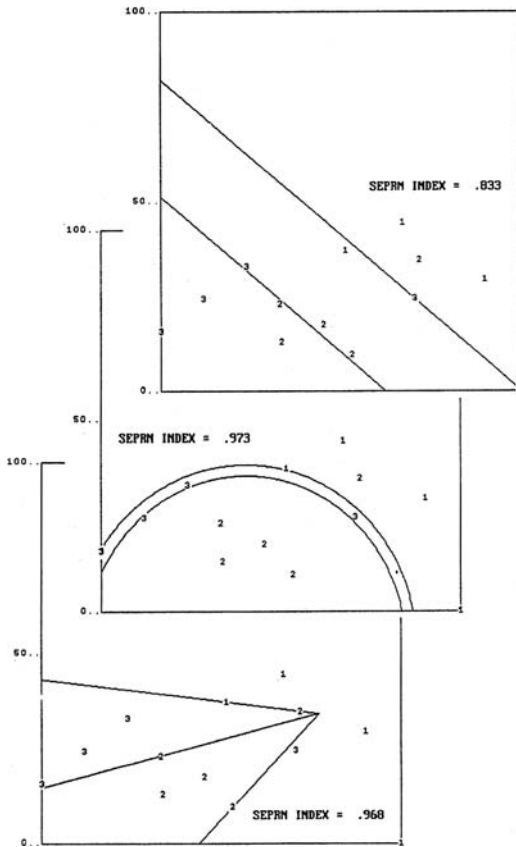


図3 SSA マップのコンピュータによる空間分割とそれぞれのケースの分離指数 (separation index) の値 (Borg and Shye, 1995, p.142)

ber)ではなく、それらの「距離の総和 (sum of distances)」にもとづいている。その点についてのテクニカルな解説は、Borg and Shye (1995, pp.143-146)を参照されたい。この指数は、0から1までの値を取り、1はすべての変数/項目が、それぞれ仮説に示された領域内に位置していることを示している。Oreg et al. (2011)は、0.90か、それ以上の値が「満足できる適合度 (satisfactory fit)」を示すものであるとしている。

以上は、SSA という統計的技法についての、方法論的な諸文献 (methodological literatures) にもとづくまとめである。では、このような技法は、いわゆる「応用研究 (applied research)」といわれる領域において、実際に、どのように利用されるのであろうか。そのような研究事例の一つが、つぎのⅢで取りあげる Shaul Oreg et al. による

「人びとの変化に対する資質的な抵抗尺度」の開発研究である。

### Ⅲ. 研究事例

——Shaul Oreg et al. (2011) の「人びとの変化に対する資質的な抵抗尺度」の開発研究——

本章は、ここで、なぜ、研究事例の一つとして、この文献を取り上げるのかについての説明から始める。これについては、以下の点があげられる。

(1) 本稿は、I. はじめに、のところで述べたように、MGCFA と確証的 SSA との方法論的な比較検討に焦点を合わせている。そして、この文献は、Oreg et al. の指摘するように、「構成概念の次元性 (construct dimensionality) をテストするために、確証的因子分析 (CFA) と多次元尺度法 (MDS) の両方を同時に利用した研究は、これまでわずかしがなく、とくにある特定の尺度の測定の等価性 (measurement equivalence) をテストするために MGCFA と確証的 SSA の両方を採用したのものとしては、本研究が初めてのものである」(p.253) として位置づけられる。そして、そうであるならば、われわれの文献研究において、このパイオニア的な研究事例に注目することは、容易に理解されるであろう。

(2) Shye, Elizur, Hoffman (1994, p.120) は、「SSA という技法は、その結果が領域の近接パターン (regional contiguity pattern) という点から注視され、解釈される場合に、『理論構築』と『尺度化』にとって、最大の力を発揮するものとなる」(p.120) と述べている。

そして、そうであるならば、前者の「理論構築」——Shye らの視座からするならば、それは「ファセット・セオリー」の構築ということになるが、本稿の筆者らからするならば、「価値観理論」などの「substantive な理論」の構築も射程に入ってくる——ということについては、しばらく措くとして、後者の「尺度化」ということについては、Oreg et al. の研究事例は、まさにそのような「尺度開発」についての議論の展開の可能性

を、CFA/MGCFA と SSA との方法論的な比較をとおして、探究したものであり、本稿におけるこのような研究事例の再検討は、Shye らの指摘を具体的に確認する試みということになる。

では、なぜ、このような確認を試みるかという点、それは、それによって、この研究領域においても、methodological な「知の蓄積 (cumulative knowledge)」(Blalock, 1989, p.15) がもたらされるであろうことを確信するからにほかならない。

(3) 前稿においては、国際比較／文化比較調査における「測定の等価性／不変性」に関する研究の系譜とその内容を概観したが、ここでは、この研究領域における諸文献を、

- ・「国際比較調査の実践と方法に関する研究」
- ・「統計的方法に関する研究」
- ・「統計的方法の国際比較／文化比較研究への応用に関する研究」

に分けるとともに、今後の筆者らによるこの研究領域に対する貢献は、③の「応用研究」における創造性 (creativity) の探究という点にあることを示唆した。そして、Oreg et al. による「人びとの変化に対する資質的な抵抗尺度」の開発研究は、まさにそのような「応用研究」の先駆的な研究事例の一つといわなければならない。これこそが、ここで、この文献を取りあげる理由である。

以上を踏まえて、本章では、Oreg et al. によるこの研究事例の「問題関心」「調査方法」「分析結果」「CFA/MGCFA と 確証的 SSA の 比較検討」について、やや詳細に紹介していくことにする。

## 1. 研究の問題関心

個々人は、変化の諸状況 (change situations) に対する典型的な反応 (typical response) が異なる。このような変化に対する典型的な反応 (typical reaction) の違いは、人格特性 (personality trait) の一つ、すなわち「変化に対する資質的な抵抗 (dispositional resistance to change)」として概念化 (conceptualized) された (Oreg, 2003)。そして、そのような変化に対する抵抗の特性を捉える「測定の尺度 (measurement scale: RTC scale)」が、一連の諸研究をとおして確立されて

きた。それら諸研究においては、RTC 尺度の「構造妥当性／構成概念妥当性／併存的妥当性／予測的妥当性 (structural/construct/concurrent/and predictive validities)」が立証 (demonstrated) されてきた。人は、変化に対して資質的に抵抗する傾向があればあるほど、ある特定の具体的な変化に対してより否定的な態度を表わす傾向があり、そして自分から進んでそのような変化を受け入れる傾向が少ない (例えば、Oreg, 2006 ; Oreg, Nevo, Metzger, Leder, and Castro, 2009)。

以上のような、変化に対する抵抗の特性は、つぎの4つの次元 (dimensions) からなるものとされる。

(1) 決まりきったことへの要求 (routine seeking) : 人は、どの程度まで安定した (stable)、決まった環境 (routine environments) を享受し、追求するか。

(2) 感情的な反応 (emotional reaction) : 人は、外部から強いられる (imposed) 変化に対して、どの程度、ストレスや居心地の悪さ (uncomfortable) を感じるか。

(3) 短期的な焦点整合 (short-term focus) : 同じく、感情的な源泉 (affective sources) からでてくるものであり、人が変化についての、一方の「短期的な不快感 (inconvenience)」と、他方の「長期的な利得 (benefit) の可能性 (potential)」に、どの程度、心を奪われるか。

(4) 認知的な融通性のなさ (cognitive rigidity) : もう一つの別の考え方や見方をしてみることにに対して、どの程度、頑固であり、気が進まないか。

これら4つの異なる次元は、異なるコンテキストにおいて、顕現化 (salient) してくるものであるが、「合成 RTC スコア (composite RTC score)」を用いるならば、自発的な (voluntary)／強制的な (imposed) 諸条件のさまざまなコンテキストにおいて、個々人の「変化への反応」が予測できるということが示されてきた (Oreg, 2003, 2006)。

このような一連の研究は、substantive な点からしても、methodological な点からしても、きわめ

て興味深いものといわなければならない。ところが、そこには一つの「問題」が残されていた。それは、このような一連の研究が、アメリカ合衆国とイスラエルにおいて収集されたデータにのみもとづくものであるということであった。まさに、ここから、「変化に対する資質的な抵抗」という「構成概念」と、それを捉える「測定の尺度」の、アメリカ合衆国とイスラエルの文化的セッティングを越える「一般化可能性 (generalizability)」の探究という問題関心がでてくることになった。

こうして Oreg et al. (2008) は、世界の 17 か国から集められたデータにもとづいて、「この研究領域において最も強力 (most powerful) で、用途の広い (versatile) アプローチ」(Steenkamp and Baumgartner, 1998, p.78) とされる MGCFA を用いて、このような「一般化可能性」を実証的に確認した (Oreg et al., 2008)。

しかし、このような結果は、測定の等価性／不

変性のテストのために開発されてきたほかの技法を用いても、同じように確認できるであろうか。このような方法論的な問題関心から、この文献では、CFA/MGCFA と確証的 SSA の比較検討がなされることになったのである。

## 2. 質問紙調査の方法とその結果

### (1) 調査の方法

#### ①調査の対象国

調査の対象国は、オーストラリア、中国、クロアチア、チェコ、ドイツ、ギリシャ、イスラエル、日本、リトアニア、メキシコ、オランダ、ノルウェー、スロバキア、スペイン、トルコ、イギリス、アメリカ合衆国の 17 か国である。

#### ②調査の対象者

心理尺度の開発研究においては、大学の学部学生が調査対象者に選ばれることが多いが、ここでも、以上の 17 か国の 4201 名の学部学生に対して

表 2 RTC 諸項目と 17 か国における平均因子負荷量

Item	Factor	Mean Standardized Loading	Loading SD
1. I generally consider changes to be a negative thing.	Routine seeking	0.54	0.14
2. I'll take a routine day over a day full of unexpected events any time.		0.64	0.10
3. I like to do the same old things rather than try new and different ones.		0.70	0.08
4. Whenever my life forms a stable routine, I look for ways to change it. <sup>a</sup>	Emotional reaction	0.44	0.11
5. I'd rather be bored than surprised.		0.50	0.09
6. If I were to be informed that there's going to be a significant change regarding the way things are done at school, I would probably feel stressed. <sup>b</sup>		0.64	0.08
7. When I am informed of a change of planes, I tense up a bit.		0.72	0.08
8. When things don't go according to plans, it stresses me out.		0.64	0.08
8. If one of my professors changed the grading criteria, it would probably make me feel uncomfortable when if I thought I'd do just as well without having to do any extra work. <sup>b</sup>	Short-term focus	0.54	0.10
10. Changing plans seems like a real hassle to me.		0.62	0.11
11. Often, I feel a bit uncomfortable even about changes that may potentially improve my life.		0.72	0.09
12. When someone pressures me to change something, I tend to resist it even if I think the change may ultimately benefit me.		0.49	0.10
13. I sometimes find myself avoiding changes that I know will be good for me.	Cognitive rigidity	0.50	0.09
14. I often change my mind. <sup>a</sup>		0.48	0.17
15. I don't change my mind easily.		0.63	0.11
16. Once I've come to a conclusion, I'm not likely to change my mind.		0.68	0.08
17. My views are very consistent over time.		0.64	0.13

Source : Oreg, S., Bayazit, M., Arciniega, L., Armenakis, A. A., Barkauskiene, R., ... van Dam, K., *Journal of Applied Psychology*, 93 (4), 935-944, 2008. ©2008 by the American Psychological Association. Reproduced with permission.

<sup>a</sup> These items are reverse-coded.

<sup>b</sup> When used in a job setting, these items are rephrased to fit the organizational context.

質問紙調査が実施された。調査対象者の等質化を高めるために、「教育歴」と「年齢」による「マッチング・サンプル (matching sample)」が構成された。

サンプル・サイズは、各国平均が241名——最も多いギリシャが386名、最も少ないスロバキアが171名——で、中国(194名)とスロバキア(171名)を除いて、それら以外の国々には、いずれも200名以上となった。このようなサンプル・サイズは、CFAにおいて、安定した結果(stable results)を導くことのできるものといわれている(Kline, 1998)。

③調査は、大学における授業の評価(credit)あるいは条件(requirement)という形で、学習の一環として実施された。

(2) 質問紙の作成

①質問紙——RTC 尺度を構成する17の質問項目(ステートメント)と回答者の属性項目——は、翻訳-逆翻訳(translation-back-translation)のプロセスをとおして、それぞれの参加国の言語(native language)に翻訳された。

②個々人の変化に対する志向/反応(orienta-

tion/reaction)に関する17のステートメントと、それぞれの対応する4つの次元については、表2を参照されたい。なお、それぞれのステートメントに対する回答の選択肢(response options)としては、1「強くそう思わない」から6「強くそう思う」までのレンジ(range)が採用された。

(3) MGCF/SSA に先立つ予備分析

①信頼性係数

17の調査対象国における、4次元で構成されるRTC 尺度の信頼性係数——クロンバックの $\alpha$ 係数(Cronbach's alpha)——は、表3に示した。信頼性係数は、すべて0.70、あるいはそれ以上の「満足できるレベル(satisfactory level)」に達していることがわかる。

②確証的因子分析

調査対象国における「測定の等価性(measurement equivalence)」を立証する(establishing)第1のステップとして、それぞれの国のサンプルに対して別々にCFAが実行(run)された。ここでは、Oreg et al. (2008)で確認された因子間に相関関係が見られる「4因子解(a four-factor solution)のモデル」(図4)が仮定されたが、結果

表3 17か国における信頼性係数、記述統計量、適合度指数

Country	N	RTC $\alpha$ Chronbach	RCT Mean	RCT SD	$\chi^2_{(107)}$	RMSEA <sup>a</sup>	CFI <sup>b</sup>	GFI <sup>c</sup>
Australia	251	0.82	3.09	0.57	172.56	0.050	0.93	0.93
China	194	0.85	3.14	0.62	170.07	0.055	0.94	0.91
Croatia	246	0.84	3.01	0.61	158.88	0.045	0.97	0.93
Czech Republic	224	0.84	3.13	0.56	184.24	0.057	0.92	0.91
Germany	206	0.77	3.12	0.48	131.36	0.033	0.97	0.93
Greece	386	0.72	3.03	0.50	227.29	0.054	0.93	0.94
Israel	241	0.85	3.15	0.59	193.42	0.058	0.93	0.92
Japan	337	0.75	3.22	0.52	198.46	0.051	0.91	0.93
Lithuania	212	0.77	2.86	0.51	171.39	0.053	0.92	0.91
Mexico	265	0.79	2.79	0.58	216.74	0.062	0.92	0.90
Netherlands	205	0.85	3.17	0.52	177.59	0.058	0.94	0.91
Norway	266	0.84	2.91	0.56	218.21	0.063	0.92	0.91
Slovakia	171	0.79	3.27	0.51	184.28	0.065	0.90	0.89
Spain	288	0.81	3.01	0.58	165.97	0.044	0.95	0.94
Turkey	241	0.77	3.03	0.54	188.86	0.056	0.90	0.91
United Kingdom	204	0.78	3.02	0.51	190.22	0.062	0.90	0.90
United States	264	0.83	3.05	0.54	160.90	0.044	0.95	0.94
Mean	247.12	0.80	3.06	0.55	183.08	0.050	0.93	0.92

Source : Oreg, S., Bayazit, M., Arciniega, L., Armenakis, A. A., Barkauskiene, R., ... van Dam, K., *Journal of Applied Psychology*, 93(4), 935-944, 2008. ©2008 by the American Psychological Association. Reproduced with permission.

<sup>a</sup> Root-Mean-Square Error of Approximation.

<sup>b</sup> Comparative Fit Index.

<sup>c</sup> Goodness-of-Fit Index.



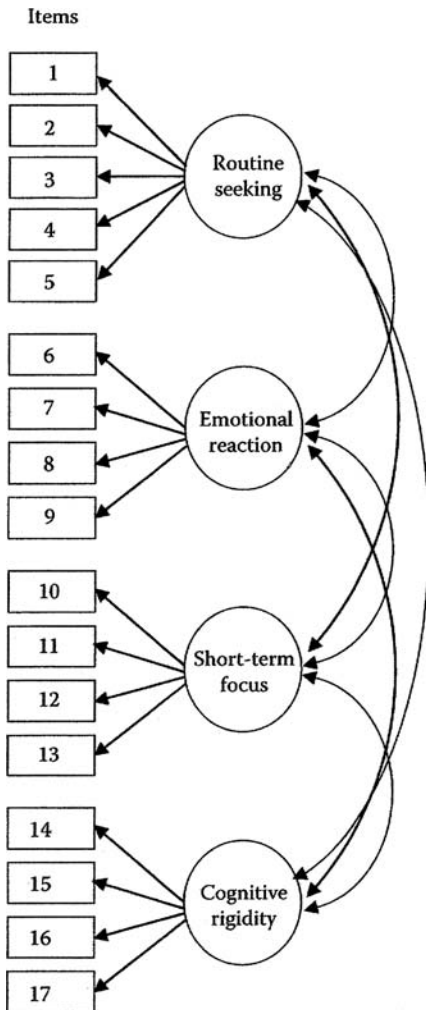


図4 RTCの4因子モデル (Oreg et al., 2011, p.263)

(表3)は、4因子のRTC尺度は、対象国のすべてにおいて、少なくとも「満足できる適合度 (satisfactory fit)」を示した。ただ、例外はスロバキアで、スロバキアではGFIは0.90よりもわずかに低い結果(0.89)となった。

③RTCの部分尺度間の相関関係

Oreg (2003)の諸知見と同様、RTCの部分諸尺度(subscales)間には、有意な相関関係が見られる(表4)。最も高い相関関係は、「感情的な反応(emotional reaction)」と「短期的な焦点整合(short-term focus)」との間のものであり、最も低い相関関係は、「感情的な反応」と「認知的な融

表4 17か国におけるRTC部分尺度間の平均相関係数

Subscale	1	2	3
1. Routine seeking			
2. Emotional reaction	0.49		
3. Short-term focus	0.61	0.77	
4. Cognitive rigidity	0.23	0.16	0.21

<sup>a</sup> Correlations (estimated) among RTC subscales were derived within the CFAs. (Oreg et al., 2011, p.264)

通性のなさ(cognitive rigidity)」に関するものである。

3. MGCFAと確認的SSAによる分析結果

A. MGCFAの結果

Oreg et al. (2011)では、MGCFAによる測定不変性(Measurement Invariance, 以下MI)の分析方針を述べているが、それは要約すれば次のようにまとめられる。(1)前提として、分析の目的をOreg et al.は「個人レベル(individual-level)」のMIの検討と定めていることから、検討すべき不変性の範囲をII A-1、およびII A-2に述べた「configural invariance」と「metric invariance」に限定する(p.258)。そして、(2)モデルの適合度指標として主にCFIとRMSEAを用い、やや補助的にGFIを参照する。制約がより強いモデルでの適合が、有意に悪くなっていないことを以って、制約が強いモデルの成立可否を判定するための手順を先行研究の推奨に基づいて定める。(3)多数の、(文化的な)異質性の高い国を含む条件では、完全な不変性(full invariance)が成立するとみることが多いので、部分的な不変性(partial invariance)を許容する(具体的には一部の項目の因子負荷が国間で異なってもよいとする)。これらはII A-2で述べた図1での最上部「完全な測定不変性の検証」とその左下「部分的な測定不変性の検証」に相当するものである。

以上の方針のうち、(3)については、17か国というグループ数の多さから自然な設定といえる。部分的な不変性の検討の必要性和その手段については、真鍋・前田・清水(2021)も参照されたい。(2)については分析方法上の先行研究の推奨に基づくものであり、Oreg et al.はネストした

モデル間の適合度比較のために所謂 *chi-square difference test* (より制約の強いモデルでのモデルの適合の悪さを示すカイ2乗値の増分) の検討結果は(論文中で結果は示されているものの)採用していない。この手続きが比較的瑣末な *invariance* からの逸脱でも、より制約的なモデルは有意に適合度を落としたものと判定されがちであることをその理由としている。他方、CFI については(より制約の強いモデルで)適合度が過度に減少していないと判断できる差分を特定して検討する方針を示している。GFI は最も一般的であり古くから用いられている CFA (またはより一般的な構造方程式モデル SEM) での適合度指標である。分析方針の(1)についての議論は、後述する。

以上の分析方針と MGCFA に先立つ国ごとの CFA による予備分析での4因子モデルの適合のよさの確認を経て、著者らは MGCFA による *configural invariance* と *metric invariance* に関する検討結果を述べている。ここでの検討は4因子に対して17項目が配置された17行4列の因子負荷行列に対する検討である点に留意する。

まず、*configural invariance* のモデル、すなわち因子負荷行列の自由母数の位置が17か国の間で同じである(値は異なってもよい)とするモデルの適合度は十分に高く、また因子負荷も想定した因子に確かに有意な負荷を持つことが示されたとされる。

さらに、*metric invariance* を仮定したモデルでも満足できる適合度が得られたこと、すなわち、*configural invariance* からの大きな悪化がないことを述べている。しかしながら CFI 指標ではやや大きい低下が観察されたので、*metric invariance* モデルすなわち全ての因子負荷の値を17か国間で同じと仮定するのはやや制約的過ぎる可能性が示唆され、このためいくつかの項目でその仮定を緩めることが検討されている。つまり Oreg et al. は、図1における部分的な測定不変性の検討の中の *partial metric invariance* の検討手順を踏んでいる。Oreg et al. は、RTC 尺度内のどの項目で文化間の意味合いの差が大きいのかという点について依拠できる先行研究がないことと、パーソナリティ尺度などでは否定的ワーディングを用いた(い

わゆる)逆転項目で、文化間で反応傾向の差をもたらしやすいという経験則を踏まえて、RTC 尺度内での逆転項目である項目4、14で因子負荷の等値制約を緩めるという方針を定めた。実際にこの制約を緩めることで *metric invariance* からの適合度の改善を得ることができたことを示している。項目4、14については次の確証的 SSA でも考察対象とされるので、その内容についてはそちらで示すこととしよう。

以上が Oreg et al. の MGCFA による分析結果の概要である。ここでは一部の項目(逆転項目である2項目)を除いて *partial metric invariance* が成り立ち、「変化に対する資質的な抵抗」という構成概念は、4つの下位側面を持つ共通のものとして、用意された17項目により17か国間での比較可能な(妥当な)形で測定されていることを述べようとしている。

## B. 確証的 SSA の結果

1. Oreg et al. は、17か国の調査データについて、Shye (1997) によって開発されたソフトウェア、FSSA: Faceted Smallest Space Analysis を用いて確証的 SSA を実行し、二次元の SSA マップを描いた。結果は表5と図5に示したとおりである。

2. 確証的 SSA の結果の解説は、17か国における SSA マップの「疎外係数」——「相関マトリックス」と「SSA マップ(空間布置図: spatial plot)」との適合度——の検討からスタートする。それらは、表5の第2カラムに示されている。では、これらの数値はどのように評価されるであろうか。すでに述べたように、これまで、① Guttman (1968)、② Borg and Lingoes (1987)、③ Ben-Shalom and Horenczyk (2003)、の三つの基準が提案されている。ここで、③の基準を採用するならば、「疎外係数」は、すべての国で0.25の境界点(cut-off point)をクリアしていることになる。しかし、①と②の基準を採用するならば、問題となる国ぐにがでてくる。Oreg et al. は、これらの国ぐについては、二次元以上の SSA マップの検討の必要性を示唆している。

表 5 17 国の SSA 疎外係数と分離指数 (Oreg et al., 2011, p.269)

Country	Alienation Coefficient	Four-Factor Solution		Three-Factor Solution	
		# of Deviant Items	Separation Index	# of Deviant Items	Separation Index
Australia	0.21	3	0.90	1	0.95
China	0.24	3	0.95	0	1
Croatia	0.15	0	1	0	1
Czech Republic	0.19	0	1	0	1
Germany	0.20	0	1	0	1
Greece	0.19	3	0.94	0	1
Israel	0.21	0	1	0	1
Japan	0.21	3	0.97	0	1
Lithuania	0.23	4	0.79	2	0.95
Mexico	0.20	3	0.92	0	1
Netherlands	0.16	0	1	0	1
Norway	0.14	0	1	0	1
Slovakia	0.19	4	0.95	0	1
Spain	0.19	0	1	0	1
Turkey	0.23	0	1	0	1
United Kingdom	0.21	4	0.88	0	1
United States	0.18	0	1	0	1
Mean	0.19	1.59	0.96	0.18	0.99

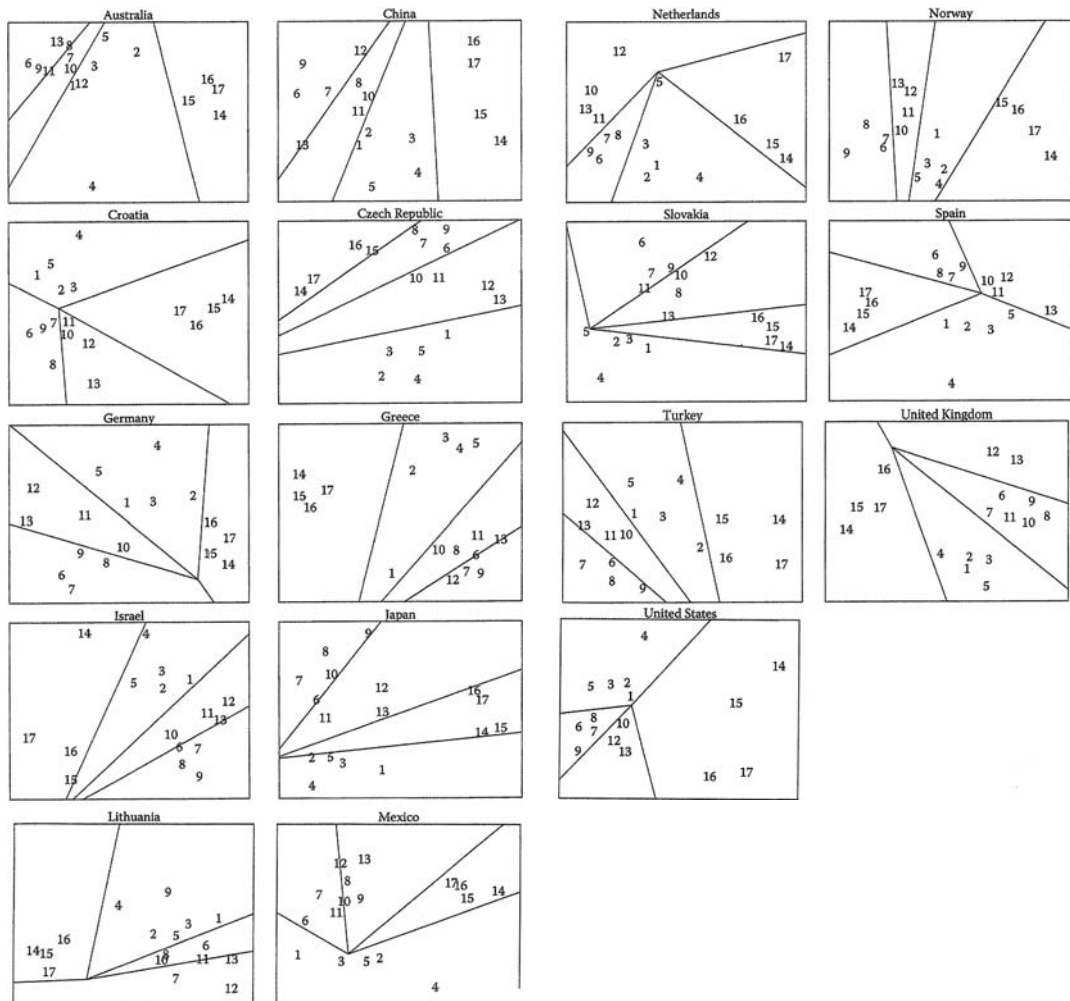


図 5 17 国の SSA マップ—Angular/Polar 空間分割— (Oreg et al., 2011, p.266-268)

3. つぎに、「仮説にあげられた尺度の構造」と「SSA マップの空間分割図 (spatial partition)」との適合度の検討に移る。これは、いいかえれば、それぞれの国の「空間布置図」については、どのような「空間分割図」が最も適切であるか、という「問い」である。このような「問い」に対する「答え」が、FSSA の実行によってもたらされる。Oreg et al. は、FSSA の“angular—Guttman 用語では polar—partition”を採用する。

では、Oreg et al. は、なぜ、“axial partition”、あるいは、“radial—Guttman の用語では modular—partition”ではなく、“angular partition”を選択したのであろうか。ここでは、以下の二つの理由があげられている。

①どの「空間分割」の方法を選ぶかについて、とくにこれまでの諸研究 (Oreg, 2003 ; Oreg et al. 2008) から示唆される「概念的な基礎 (conceptual basis)」といったものはない。

②それが、これまで発表されてきた文献において、最もよく利用されてきた方法であり、最もよく知られた方法である。

4. このような、“angular/polar partition”の方法による 17 か国の SSA マップの「空間分割図」の検討——その適合度の検討——は、どのようなものであろうか。このような検討については、「MGCF A」と「確証的 SSA」には、大きな違いがある。繰り返しになるが、それはつぎの点である。

MGCF A：適合度の確認は、「国ぐに間の全体的なモデル適合度指数 (indexes of overall fit of the model across samples)」によって判断される。

確証的 SSA：適合度の確認は、①各国の SSA マップを別々 (separately) に検討することによって、②それらの国ぐにの SSA マップにおける諸項目の分割のパターンに同じような一貫した (consistent) 傾向が見られるかを検討することによって、なされる。

確かに、このような確認については、①逸脱諸項目の数 (deviant items)、②分離指数 (separation indexes) という二つの客観的な証拠 (objective indication)、をあげることができる。しかし、

それらが、SSA マップにおける諸項目の布置の全体的なパターンを視覚的に検討すること (visually examining) に取って代わるものではない。

5. 以上を踏まえて、表 5 と図 5 から、つぎのような結果を読み取ることができる。

(1) まず、four-factor solution/four-region partitioning の場合は、表 5 から 9 か国——クロアチア、チェコ、ドイツ、イスラエル、オランダ、ノルウェー、スペイン、トルコ、アメリカ合衆国——において「完全な支持 (full support)」が得られ、残りの 8 か国——オーストラリア、中国、ギリシャ、日本、リトアニア、メキシコ、スロバキア、イギリス——において「部分的な支持 (partial support)」——とくに、イギリス (0.88) とリトアニア (0.79) で「分離指数」の値が小さい——が得られたことがわかる。

そして、図 5 から、前者の 9 か国においては、四つの領域区分が明らかであるのに対して、後者の 8 か国においては、“emotional reaction”の諸項目 (6~9) と “short-term focus”の諸項目 (10~13) が、SSA マップ上で相互に混ざり合って位置していることがわかる。

(2) つぎに、three-factor solution/three-region partitioning の場合は、表 4 からするならば、モデル適合度は 15 か国において「完全な支持」——分離指数は 0——が得られ、残る 2 か国——オーストラリアとリトアニア——において「部分的な支持」——分離指数は 0.95 まで大きくなった——が得られた。

そして、図 5 からするならば、オーストラリアとリトアニアの 2 か国を除く、15 か国の SSA マップにおいては、“routine seeking (項目 1~5)”と “cognitive rigidity (項目 14~17)”と “emotional reaction + short-term focus (項目 6~13)”の三つの領域区分が明らかである。

では、ここで、例外とした二か国の SSA マップがどのような形状 (configuration) のものであるかということ、オーストラリアの場合は、仮説では “short-term focus”の項目として位置づけた項目 12 “When someone pressure me to change something, I tend to resist it even if I think the change

may ultimately benefit me”が、“routine seeking”の領域のなかに入り込んでいる。そして、リトアニアの場合は、仮説では“emotional reaction”の項目として位置づけた項目9 “If one of my professors changed the grading criteria, it would probably make me feel uncomfortable even if I thought I’d do just as well without having to do any extra work”が、同じく“routine seeking”の領域のなかに入りこんでいる。

6. 最後に、「変化に対する抵抗」に対して、「それとは逆の／反対の／対立する方向でワーディングがなされた諸項目 (negatively worded items)」の SSA マップ上での位置についての検討がなされている。それらは、つぎの二項目である。

4 “Whenever my life forms a stable routine, I look for ways to change it.”

14 “I often change my mind.”

17か国の SSA マップのそれぞれにおける、これら二項目の位置を見ていくならば、つぎのようなことがわかる。それは、これら二項目は、はじめの仮説において考えられたそれぞれの領域内に位置しているものの、①それぞれの領域内のほかの諸項目の群の内側に混ざり合うのではなく、それらの群の外側のところに、②angular/polar のそれぞれの分割線に近いところに、③SSA マップの外縁に近いところに、位置する傾向があるということがわかる。

こうして、これら二つの項目は、SSA マップ上のほかの諸項目とは相対的な差異性 (relative distinctiveness) を示すものであることが、実証的に確認されたのである。

### C. MGCFA と確証的 SSA の比較検討

1. 以上において、Oreg et al. は、「変化に対する資質的な抵抗」という構成概念が、異なる国ぐににおいて、等価な意味をもつかどうかを、Oreg, et al. (2008) と同じデータを用いて、CFA/MGCFA と確証的 SSA という二つの統計技法によって、実証的にテストした。その結果、これら

の方法によってもたらされた諸知見には、「共通部分 (an overlap)」とともに「異なる諸側面 (distinct aspects)」があることがわかった。そして、このような結果は、「本研究で用いられた測度 (measure: 具体的には 17 項目からなる測定の尺度)」と「その根底にある構成概念 (its underlying construct)」へのより深い洞察 (insights) を提供するものとなる。

2. 若干の例外はあるものの、概して、以下のような結果が導かれた。

- ・ 確証的因子分析 (CFA) によって、「変化に対する資質的な抵抗」の国際比較の妥当性 (cross-national validity) が確認された。
- ・ MGCFA によって、17か国のすべてにおいて、partial metric invariance の証拠 (evidence) が獲得された。「変化に対する抵抗」とは逆の／反対の／対立する方向でワーディングがなされた二項目 (negatively worded items) を除いて、それら以外の諸項目は 17か国において「不変性のある因子負荷量 (invariant loadings)」を示すものとなった。
- ・ 確証的 SSA によって、尺度の構成の再現性 (replication) に関する以上の結果が支持 (support) された。「4 領域解 (four-region solution)」は、17か国のうちの9か国において「完全な適合度 (perfect fit)」を、そして残りの8か国において「よい適合度 (good fit)」を達成した。事前の仮説どおり、構成概念の4次元——routine seeking、emotional reaction、short-term focus、cognitive rigidity——に対応する諸項目は、SSA マップの四つの領域にクラスター化 (clustered) される結果となった。

3. CFA の結果 (表4) に示された「emotional reaction の次元」と「short-term focus の次元」との間の高い相関 (0.77) に対応して、確証的 SSA の結果 (図5) においても、これら二つの次元の諸項目のクラスター (item clusters) は、常に隣接 (adjacent) している。そして、上述の8か国において「完全な適合度」が得られなかったのは、それぞれの SSA マップにおいて、これら二つの領域間で相互に諸項目の混ざり合いが見られ

ることによるものである。この結果は、「変化に対する抵抗」については、「4次元／因子／領域解」にくらべて、「emotional reaction の次元」と「short-term focus の次元」を統合した「3次元／因子／領域解」の方が、より「節約的 (parsimonious)」で、「より頑健 (robust)」な尺度構成であることを示唆している。

では、このことから、「4次元／因子／領域解」よりも「3次元／因子／領域解」を選択すべきであるという結論が導かれるであろうか。Oreg et al. は、そのような考え方は採用しない。それは、つぎのような理由からである。

①CFA と確証的 SSA を組み合わせた分析の結果は、「emotional reaction の次元」と「short-term focus の次元」との区別には、相変わらず意味がある (still meaningful) ということを示している。

②「特定の変化に対する個々人の応答 (response)」を予測するために RTC 尺度を用いた多くの研究において、これら二つの部分尺度 (subscales) の区別にもとづいて、多くの諸知見が蓄積されてきている。

以上を踏まえて Oreg et al. は、「4次元／因子／領域解」が適切なものとするのである。

4. 確証的 SSA の結果は、以上において述べてきたような構成概念の「全体的な構造 (overall structure)」の検討にとどまるものではない。SSA という統計的技法の特徴は、それを越えて、特定の項目 (specific items) の、ほかの諸項目との関連における、それらの SSA マップ上での位置 (position) についての、直観の観察を可能にするというところにある。このような SSA の性質 (nature) を利用して、以下において、いくつかの項目の検討結果が示される。

(1) 「変化に対する抵抗」とは逆の／反対の／対立する方向でワーディングされた諸項目 (項目 4 と項目 14) については、まず MGCF A によって、その「変則性 (irregularity)」が確認されたが、つぎに SSA マップにおけるそれらの諸項目の位置の検討をとおして、その「変則性」が再確認されることになった。

(2) 逸脱諸項目——仮説における「空間分割 (領域区分)」と実際の SSA マップにおける諸項目の「空間布置」の結果に差異が見られる諸項目を「逸脱諸項目」と呼ぶ——の検討が、以下のようになされる。

SSA によるデータ分析の主要な結果の一つが、「4 因子／領域解は、調査対象国の 9 か国においては、完全な支持 (full support) が得られたが、残りの 8 か国においては、部分的な支持 (partial support) が得られるにとどまった」というものであった。そこで、これら 8 か国の SSA マップについての詳細な検討が課題となる。その結果、つぎのようなことが明らかとなる。それは、これら 8 か国の SSA マップの半数において、項目 8 と項目 12 が「逸脱項目」となっているということであった。

まず、項目 8 “When things don't go according to plans, it stresses me out” は、事前の仮説においては “emotional reaction” の部分尺度の一部をなすものとして位置づけられた。ところが、SSA による分析の結果、それが隣接する (adjacent) “short-term focus” の領域に入り込んでいることがわかった。しかし、この結果は理解／納得することのできるものといえる。それは、計画を変更するということは、短期的には迷惑なものであっても、長期的には利益をもたらす可能性があると考えられるからである。

つぎに、項目 12 “When someone pressures me to change something, I tend to resist it even if I think the change may ultimately benefit me” は、事前の仮説においては “short-term focus” の部分尺度の一部をなすものとして位置づけられた。ところが、SSA による分析の結果、それが隣接する “emotional reaction” の領域に入り込んでいることがわかった。しかし、この結果も、理解／納得することのできるものであろう。それは、利益をもたらす可能性のある変更には抵抗することが、いかに感情的な諸要素によって駆り立てられたものであるかは想像に難くないからである。このような結果の解釈 (interpretation) ——「事後的な説明 (post hoc explanation)」——は、「心理的リアクタンスの理論 (theory of psychological reactance)」 (Brehm, 1966) という理論的説明 (theoretical ex-

planation) につながるものであることを示唆している。

以上が、Oreg et al. による、SSA マップにもとづく「逸脱諸項目」の検討の結果である。ただ、このような結果については、つぎのような点に注意が喚起されている。

①以上の項目 8 と項目 12 については、17 か国のなかの 13 か国において、仮説どおりの領域区分が示されており、以上の逸脱事例は 4 か国において観察されたにすぎない。

②項目 8 と項目 12 を除いて、それ以外の逸脱諸項目については、一貫性というものは見られない。

③逸脱事例が見られる国ぐにに、一つの「共通項 (a common denomination)」ともいうべきものを見出すことはできない。

5. 4 因子／領域解についての、SSA マップによる再検討の結果、以下の点が明らかとなった。

4 因子／領域解は、多くの国ぐににおいて、全体的な支持 (overall support) が得られたが、そのような支持はリトアニアとイギリスにおいてはやや不十分 (poorer) であった。

まず、リトアニアでは、「分離指数」は 0.79 とかなり低い。そして、CFA の結果、4 因子／領域解にくらべて、3 因子／領域解の適合度がより高く (superior)、より適切 (appropriate) であることがわかっている。しかし、リトアニアにおいて、なぜ、そのような結果となったかについては、「理論的説明」といえるようなものはいまがない。

つぎに、イギリスでは、「分離指数」は 0.88 で、Oreg et al. の採用した 0.90 という境界点 (cutoff point) をわずかに下回る。このような SSA マップの検討結果は、CFA におけるイギリスの「モデル適合度 (model fit)」の低さ——RMSEA が 0.062、CFI と GFI が 0.90——に対応するものとなっている。しかし、以上のような結果について、いまだ「理論的説明」といえるものがない現状にあっては、そのことについて何かはっきりした結論を導くに先立って、イギリスについてのさらなるデータの収集を続けることが必要

であるという。

6. 上述のイギリスについてのデータ分析では、CFA と SSA で共通する結果が示されたが、これら二つの方法で、相異なる結果が導かれることもある。例えば、以下のような例をあげることができる。

スロバキアは、CFA では、RMSEA が 0.065、CFI が 0.90、GFI が 0.89 という最も低い適合度を示したが、確証的 SSA では中程度の適合度であった。では、なぜ、このような結果となったのであろうか。一つの可能性として、それには、CFA の結果に影響を及ぼすであろうデータの「正規性 (normality)」あるいは「線形性 (linearity)」という問題が関係しているかもしれないが、それらは「正規性」や「線形性」を前提としない確証的 SSA にとっては関係のないことであるかもしれない。

それとは逆に、リトアニアとオーストラリアは、確証的 SSA ではやや低い適合度が示されたが、CFA では比較的高い適合度指数が得られた。確証的 SSA におけるリトアニアとオーストラリアのやや低い適合度は、リトアニアの場合の項目 9、オーストラリアの場合の項目 4 のような、ある特定の項目の極端な「逸脱」と関係しているかもしれない。このような「逸脱」は確証的 SSA においては「歪んだ結果 (skewed results)」をもたらすものであるが、CFA においては「強い痕跡 (strong trace)」を残すものではないようである。

以上のような CFA と確証的 SSA という二つの統計的方法による諸知見 (findings) の違いは、それら二つの方法の相互に補い合う価値 (complementary value) に光を当てるものとなる。

#### IV. Oreg, et al. (2011) の研究事例の課題と展望

国際比較／文化比較調査における測定の比較可能性の確認のための統計的技法としては、さまざまなものが開発されてきている。本稿では、これら諸技法のなかから、因子分析法の系列に属する CFA と MGCFA、そして多次元尺度法の系列に

属する確証的 SSA を選び出し、それら技法のアイデア・手続き・特徴の概要を記すとともに、さらにこれらの三つの統計的技法を用いて、RTC 尺度と呼ばれる心理的尺度の開発研究を進めてきた Oreg. et al. (2011) の研究事例を取りあげ、その問題関心・方法・結果について紹介してきた。

そこで、最後に、この領域における今後の研究の発展の視座から、Oreg. et al. (2011) の研究事例の課題と展望を取りまとめておきたい。

では、そのような「課題と展望」をどのような構成とするかということ、ここでは Oreg. et al. (2011) の論述の流れに沿って、そこでいくつかの問題を取りあげるとともに、それぞれについての方法的な議論の展開を試みるという仕方を取りたい。

## 1. Oreg et al. の共同研究の問題関心

Oreg et al. の、このテーマに関する研究の展開は、「実証科学における研究過程論」——例えば、Wallace (1971=2018) ——という視座からして、きわめてオーソドックスなものといえることができるであろう。

・まず、「変化に対する資質的な抵抗」が、人びとの人格特性の一つとして概念化された (Oreg, 2003)。

・それにもとづいて、そのような構成概念を実証的に測定するための RTC 尺度が開発されるとともに、その構造妥当性・構成概念妥当性・予測的妥当性などが立証されてきた (Oreg, 2006; Oreg et al., 2009)。

・以上の一連の研究は、アメリカ合衆国とイスラエルにおける質問紙調査のデータ分析を踏まえたものであった。そこで、そのような文化的セティングを越えるデータの収集をめざして17か国調査が実施されるとともに、その結果の MGCFA を用いたデータ分析がなされた (Oreg, 2008)。

・このようなデータ分析の結果は、MGCFA 以外の統計的技法を用いても、同じように確認できるであろうか。このような方法的な問題関心から、MGCFA と確証的 SSA の比較検討がなされることになった (Oreg. et al., 2011)。

さて、われわれ——真鍋、前田、清水——の問題関心は、以上のような「実証科学における研究過程論」の枠組みを越えて、さらに社会科学のこの領域における「研究の歴史」を探るという視座にある。このような視座からするならば、Oreg. et al. (2011) における「CFA・MGCFA と確証的 SSA の併用」というアイデアは、1970年代からの実証的な社会科学の方法的な系譜における“methodological pluralism”——全体として研究成果における方法の多様性に関して用いられる——や“multimethod approach”——一つの研究においていくつかの異なる方法が用いられること——という考え方 (Payne and Payne, 2004=2008) と、どのようにつながってきているであろうかという問題関心がでてくる。いうまでもなく、ある研究者の方法的な関心というものは、その研究者の主體的な関心にもとづく独自のものととともに、その時代と社会の関心と通底する「共同主観的」(廣松, 2017) なものでもあるからにはかならない。

さらにいえば、筆者(真鍋)のより具体的な仮説として、つぎのようなことがあげられる。それは、このような「いくつかのデータ分析の技法の併用」という方法的関心の背後には、「社会調査データの二次分析 (secondary analysis)」の発展という研究動向があったという仮説である。では、二次分析とは何かというと、この点の詳細については、真鍋 (2012) を参照されたいが、一次分析 (primary analysis) が、社会調査を実施して、データを収集し、その分析をとおして知見を導くことを指すのに対して、そのようにしてすでに収集されたデータを用いて、新たな仮説や方法にもとづいて、新たなデータ分析を行ない、新たな知見や解釈を導き出すことを二次分析と呼ぶ、というのが一般的な考え方である。こうして、調査データの統計的分析ということに関して、さまざまな技法の利用の試みが広がっていったというのが筆者の仮説である。このような問題関心の探究は、今後のきわめて興味深いテーマの一つとなるであろう。

## 2. 調査方法

Oreg. et al. (2011) は、Oreg. et al. (2008) の



17 개국調査データを踏まえたものである。そこで、Oreg et al. は、この 17 개국調査の方法をめぐる二つの課題について記している。

①17 개국調査は、大学生を対象になされた質問紙調査であって、そのデータはランダム・サンプリングにもとづくものではない。したがって、サンプルは、それぞれの国の国民文化 (national culture) を代表するものではない。

②調査対象国の 17 개국は、世界の国民文化の全てを代表するものではない。

いうまでもなく、以上の二点は、国際比較／文化比較調査のあり方という視座からして、今後に残された重要な課題であることは間違いない。しかし、「Schwartz の概念枠組みにもとづく価値観の国際比較——ドイツと日本における『大学生調査』のデータ分析——」と題する国際共同研究 (真鍋, 2018 b, 真鍋, Jagodzinski, Davidov, Dülmer and Hommerich, 2020 a, 2020 b) を継続してきた筆者 (真鍋) からするならば、Oreg et al. の「大学生調査」の方法については、より喫緊の課題を指摘しておくことに意味があると考ええる。それは、以下の三点である。

### (1) 大学生調査の問題

上述の筆者 (真鍋) らによる国際共同研究においては、ドイツのケルン大学、日本の北海道大学と青山学院大学の三つの大学の学生を対象に質問紙調査が実施された。その結果、質問項目によっては、ドイツの大学と日本の大学との違いを越えて、北海道大学と青山学院大学との違いが大きいということがわかった。このことは、それぞれの国で、どの大学を選ぶかによって結果が大きく変わるということを示唆している。この点は、Cieciuch et al. (2019) による「国際比較／文化比較における測定の不変性に関する方法論的な文献のレビュー」において、この研究領域における重要な注意事項としてあげられた「特定の調査対象国の national population samples に内在する異質性 (heterogeneity)」という問題の一つの具体的な事例であるといえよう。確かに、このような問題は、国際比較／文化比較調査データの高度な統計的技法による分析に先立って、十分に検討しておくなければならない事項といわなければならない

い。

### (2) 質問紙作成の問題

質問紙調査の統計的分析における「データの水準」という問題に関しては、これまでさまざまな方法論的な議論がなされてきている。例えば、一方の多次元尺度法という技法においては、データは「順序尺度 (ordinal scale)」でなければならないとされてきた (Borg and Groenen, 2005)。そして、他方の MGCFA は、「連続分布 (continuously distributed)」／「正規分布 (normally distributed)」のデータのための技法であるとされてきた (Davidov et al., 2011)。もっとも、後者については、その後「シミュレーション研究」にもとづいて、たとえデータが「順序尺度」のものであっても、MGCFA はうまくいく (works well) ということが明らかにされた (Welkenhuysen-Gybels and Billiet, 2002 ; Welkenhuysen-Gybels, 2004 ; De Beuckelaer, 2011)。そして、このような研究成果を踏まえて、Davidov et al. (2011) は、「ヨーロッパ社会調査 (ESS)」のデータを用いて、リッカート・スケール型の順序尺度の質問諸項目についても、MGCFA の利用が可能であることを確認した。

以上のような研究動向を踏まえていえば、Oreg らは、RTC 尺度の開発のための質問文の作成にあたって、どのようにして、scalable な質問諸項目を準備したか、そして、どのようにして、それらが統計的技法によるデータ分析に耐えうるものであるかの検討を行なったか、という疑問がでてくる。

しかし、Oreg et al. (2011) においては、このような方法論的な議論は、十分になされているとはいえない。

### (3) 回答の選択肢の形式の問題

RTC 尺度の諸項目としては、表 2 に示した 17 の質問文 (ステートメント) が作成された。そして、それぞれの質問文に対する回答の選択肢 (response option) においては、1 「強くそう思わない (strongly disagree)」から 6 「強くそう思う (strongly agree)」までの「アグリーメント・スケール」の形式が採用されている。

このような回答の選択肢の形式に関しても、その国際比較／文化比較調査のコンテキストにおける「等価性」の研究がなされてきている。そのような研究事例の一つが、ドイツ・マンハイムの「調査・方法・分析センター (ZUMA・当時)」と米国シカゴ大学の「全国世論調査センター (NORC)」の共同研究として実施された MINTS (Research into Methodology of Intercultural Surveys) プロジェクトであり、筆者 (真鍋) はこの共同研究の背景・方法・結果をとりまとめたレビュー論文を書いている (真鍋, 2003)。

しかし、Oreg et al. (2011) においては、このような問題については、方法論的な関心は全く示されていない。

#### (4) 質問紙の翻訳の問題

質問紙 (調査票) の翻訳-逆翻訳のプロセスにおいては、翻訳者と逆翻訳者が、それが「最も適切な翻訳 (most appropriate translation)」であるかどうかについて、合意 (agreement) に到達するまで、“Original Version” と “Back-translated Version” との間の違い (differences) について議論を続けた、と記されている (p.256)。しかし、「最も適切な翻訳」ということが、どのような内容であるのか、そして、そのような判断をめぐって、どのような議論がなされたのか、さらに、両者の合意はどのように達成されたのか、などについての記述は全くない。

この点は、すでに述べたように、「Schwartz の概念枠組みにもとづく価値観の国際比較」の国際共同研究を経験してきた筆者 (真鍋) にとっては、きわめて残念なことである。Oreg et al. の共同研究においては、日本も調査対象国の一つとなっている。そして、筆者らの国際共同研究においては、英語から日本語への翻訳と、その逆翻訳においては、さまざまな問題がでてくることが確認されている (真鍋, 2018 b)。いうまでもなく、これまで、国際比較／文化比較調査における質問紙 (調査票) の翻訳の問題については、すでに多くの先行研究がなされている (Harkness et al. eds., 2003)。ところが、それらの文献においては、筆者らがその問題の所在を確認してきた、とくに英語と日本語の二言語間の翻訳の問題に焦点を合わ

せた研究は未だ十分になされていない。斯学の現状に鑑みて、Oreg et al. の共同研究において、このような質問紙の翻訳をめぐる問題が検討されていないということは、やはり惜しまれることといわなければならない。

### 3. MGCFA に先立つ予備分析

Oreg et al. (2011) においては、MGCFA と確証的 SSA という二つの技法をめぐる方法論的な比較検討に先立って、①17か国における四次元で構成される RTC 尺度の信頼性係数、②17か国における確証的因子分析 (CFA)、③17か国における RTC の部分尺度間の平均相関係数、の検討がなされている。いうまでもなく、これらの検討は MGCFA と確証的 SSA に進むための準備作業である。しかし、それで十分かという点、必ずしもそうとはいえない。例えば、Braun and Johnson (2010) は、国際比較／文化比較調査における「比較可能性」の検討のための「基本的な技法」の一つとして「回答の度数分布 (distributions across response categories)」の検討をあげるとともに、このような「基本的な技法」が、調査データの比較における重要な問題点 (pitfalls) の所在、「高度な技法」によるデータ分析の結果が「人工物 (artifact)」である可能性、を示唆する重要な技法であるとしている。繰り返しになるが、筆者 (真鍋) は、このような技法を最大限に活用することをとおして、「価値観の国際比較」に関する国際共同プロジェクトにおいて、「高度な技法」を用いたデータ分析において見過ごされる可能性のあるさまざまな諸知見を提示してきたのである (真鍋ほか, 2020 a, b)。

### 4. MGCFA の結果

Oreg et al. の議論のうち MGCFA に関わる方法や結果について、議論すべき若干の点について、筆者 (前田) の見解を述べておこう。

(1) 第一に、この分析は複数因子で構成される尺度での検討である点の一つの特徴であろう。RTC には4側面があるとしつつも、例えば表3で  $\alpha$  係数は17項目からなる尺度全体について求めている (このためかなり高い値が確保されてい

る) 点からも分かるように一つの得点としての RTC 得点を求めるという使い方もされる尺度である。これは心理尺度などでよく見る扱い(複数の下位尺度で構成される上位尺度を単一の得点としても用いる)のようにも思えるが、いくつかの点に注意を喚起しておきたい。まず、「 $\tau$  等価」測定を仮定する  $\alpha$  係数の前提と矛盾する面が残る点である。 $\tau$  等価の仮定とは古典的テスト理論の文脈で、 $x_1 = \tau + \varepsilon_1$ ,  $x_2 = \tau + \varepsilon_2$  のように、真値部分  $\tau$  を共通成分として持つことを意味する。他方、4 因子があると仮定すると、 $\tau$  には  $\tau_1 \sim \tau_4$  のような 4 成分があることが仮定されていることになる(もちろんこれら 4 つの成分には相関があつてよい)。

この点は、例えば  $\tau$  等価測定といった仮定を必ずしも要しない信頼性係数の推定を用いることで表面的には解決されるが、別の論点も含むように筆者には思えるので、少し観点を変えて説明する。国際比較/文化比較の文脈で、特定の構成概念を扱って MI を検討する場合には、単一の構成概念を扱い、国間で同じ測定項目のみを扱う文脈では configural invariance の意味あいがいまははっきりしないケースも多くなるように思われるが、この研究の場合は全体として「変化に対する資質的な抵抗」という一つの大きな構成概念を扱いつつ、それらが 4 つの因子を持つ 17 項目という尺度を扱っているため、configural invariance の因子負荷行列  $\Lambda$  の自由母数の配置が等しいということの意味ははっきりしている。他方、このような複数の因子を扱い、因子間相関(行列としては  $\Phi$  で表記されるものとする)の値について表 4 で言及しているが、その言及は「CFA モデルで推定した因子間相関の 17 か国での平均」の形であつて、「その相関の 17 か国での異同」については論じていない。 $\Phi$  の異同を論じ、そのことを通じて、下位構成概念(下位尺度)間相互の関係が国間で異なるか否か、という考察を MI の議論に含めなくてよいのかという点について、MI の方法論を提供し、また利用している研究者の間でコンセンサスが得られているのかが不明であり、筆者はこの点の議論が必要と考えている。

(2) この論文で扱ったデータについての CFA

に基づく検討、特に先に言及した因子間相関についてのより詳細な検討は、先行研究、Oreg et al. (2008) を参照する必要がある。本稿に転載した表 4 を見ると、尺度間の相関の値は、マチマチであつて、下位尺度 2-3 間はかなり高い一方、4 は他の 3 つの下位尺度との相関が全般に低く、この下位尺度はやや異質な印象を受ける。ただし尺度 2 と 3 をひとまとめにした 3 因子モデルとの比較を行つても適合度が悪くなるとの言及はある。テスト理論的な用語では下位尺度 2 (emotional reaction) と 3 (short-term focus) の相関が高いことによる 2 と 3 の「弁別的妥当性」(異なる概念間では相関が低くなるべきである)に関わる不十分さが、いくつかの国で問題を起こしていると要約できるかも知れない(p.236-237 冒頭のパラグラフでこの言及がある:ただしこの後、先行研究に言及する形で、この 2 つの下位尺度には弁別的な妥当性がある旨を主張している)。

(3) もう 1 点、Oreg et al. の MGCF A での議論での前提についてコメントしておく。III 3 A で紹介したように、Oreg et al. は分析の前提として、MGCF A による不変性の検討の目的について、国レベル(country-level)での構成概念の(比較可能性の)確定のために、個人レベル individual-level での構成概念の MI の確認が必要ということ、そして注意を configural と metric invariance に限定することを述べているのだが、個人レベル、国レベルという用語を導入している動機が、筆者には十分理解できない。尺度に対応する因子の平均点を国間で比較するというところで、国レベルや集計レベル(aggregate level)での構成概念という枠組みを導入したのだろうが、この枠組み(国レベル構成概念の設定)による比較可能性を、個人レベルでの MI の成立の可否の求めることの意味は自明ではないというのが筆者(前田)自身の見解である。ただし、この論文で configural と metric のレベルでの invariance に議論を限定したことは、統計的な性格のかなり異なる確証的 SSA との比較を目的とすることからの要請でもあると考えられる。というのも SSA 自体は項目間の相関の情報のみを用いることから、MGCF A におけるグループ間での切片の等値性の

比較に相当する（したがって、上記二つの invariance よりもさらに制約的な scalar invariance の検討を行う）手続きを含まないからである。以上は、重要な論点であるので、本稿の最後（IV6 (6)）に、別の視点から、今後の課題の一つとして問題提起を行う。

(4) 全般的に結果部分を参照すると、結局確証的 SSA で異質性が高いという結論に至る項目は、概して表 2 で平均の因子負荷が小さい項目と言ってよいかもしれない（項目 4, 12, 14）。こうした細部だけではなく、分析からもたらされる全体像についての違いについて、Oreg et al. の立場をまとめておくことに、より大きな意義があるだろう。

## 5. 確証的 SSA の結果

(1) Oreg et al. は、17 개국調査のデータについて、Shye (1997) の FSSA と呼ばれる確証的 SSA のソフトウェアを用いて、二次元の SSA マップを描いた。いうまでもなく、確証的 SSA の目的は、RTC の“four-region partitioning”あるいは“three-region partitioning”という空間分割の確認である。そして、そうであるならば、Oreg et al. が、そのデータ分析を、なぜ、「二次元の SSA マップ」の作成とその検討にとどめ、「三次元の SSA マップ」の作成とその検討を試みなかったのか、という疑問がでてくる。

筆者（真鍋）の方法論的イマジネーションからするならば、例えば、二次元の SSA マップでは、「emotional reaction の諸項目」と「short-term focus の諸項目」が混ざり合う国ぐににおいても、三次元の SSA マップでは、それらがもう一つの次元、つまり「高さ」の次元からして弁別（識別）可能になるということもありうるのではなかろうか、と考えられるのである。このような可能性の探求は、きわめて興味深い今後の方法論的な課題の一つとなるであろう。

(2) Oreg et al. は、以上のような二次元の SSA の「空間布置図」について、その「空間分割」の方法としては、“angular—Guttman の用語では polar—”を採用した。そして、その理由とし

て、①Oreg et al. の先行研究には、どの方法を選ぶべきかを示唆する「概念的な基礎 (conceptual basis)」といったものがあるわけではない、②“angular/polar partition”が「最もよく利用されてきた方法であり、最もよく知られた方法である」、という二点をあげている。

しかし、Guttman のファセット・セオリーと SSA についての方法論的研究を続けてきた筆者（真鍋）の立場からするならば、このような説明は決して納得できるものではない。

まず、①については、Oreg et al. は、RTC 尺度の開発研究における substantive な「概念的な基礎」には言及しているものの、SSA という統計的技法と切っても切れない関係にある「ファセット・セオリー」には全く言及していない。しかし、すでに、本稿 II B のところで解説したように、SSA は「ファセット・セオリー」と関係づけられて、初めて意味のある (meaningful) データ分析の技法となる。具体的にいうならば、Guttman の「ファセット・セオリー」には、ファセットが空間分割において果たす役割に関する法則——「第二の法則」と呼ばれる——が含まれる。それは、ファセットがランク・オーダーをもつものである場合は、axial あるいは radial/modular、そして、ファセットがランク・オーダーをもたないものである場合は、angular/polar と呼ばれる空間分割の役割を果たす、という法則である。そして、Oreg らは、「変化に対する資質的な抵抗」という心理的特性を、“routine seeking” “emotional reaction” “short-term focus” “cognitive rigidity” の四つの次元で構成した。では、これら四つの次元に何らかのランク・オーダーを想定することができるかということ、社会科学のこれまでの知的インベントリーからして、そのような想定は不可能であろう。そして、そうだとするならば、上述の「第二の法則」から、このような RTC の部分尺度の諸項目は、SSA マップにおいて、angular/polar の形状の領域区分を示すであろうことが予測されることになるのである。

このような「予測」をめぐる実証的な確認の試みは、ファセット・セオリーの発展の視座からして、きわめて重要な研究テーマとなるであろう。

つぎに、②の「angular/polar partition が最もよ

く利用されてきた方法である」ということについては、本当にそうなのだろうかという疑問がでてくる。それは、すでにあげた筆者による文献研究からするかぎり、すぐに納得できるものではない。さまざまな substantive なテーマについての研究において、“axial partition” “radical/modular partition” “angular/polar partition” の三つの方法が、それぞれのケースに応じて適切に用いられてきたというのが実情なのではなからうか。

この点についても、よりシステムティックな文献研究にもとづく探索が、今後に残された重要な課題となることは間違いない。

(3) Oreg et al. の確証的 SSA の結果の検討は、CFA の結果との比較という視座から、“four-factor solution/four-region partitioning” あるいは “three-factor solution/three-region partitioning” の、いずれがより適切であるかを確認するというところに絞ってなされている。確かに、RTC 尺度の開発研究という Oreg et al. の研究プロジェクトの性格からして、そのようなデータ分析の目標と方向は、十分に納得できるものであるといわなければならない。しかし、それと同時に、この研究が CFA/MGCFA と確証的 SSA という二つの異なる統計的技法の方法論的な比較の試みであることを考えるならば、それは、これら技法を用いて、RTC 尺度を構成する 17 の質問諸項目を、その内容、形式、ワーディングなどの点から再検討する絶好の機会にもなりうるものであった。そして、そのような再検討は、ファセット・デザインのアイデアにもとづいて、実験計画法的に進めることで可能になるはずのものであった。しかし、このようなアイデアについては、ここでその全容を詳細に説明するだけの紙面の余裕はない。この点については、筆者（真鍋）が星川啓慈・大正大学教授を代表とする科研費基盤研究（A）の「8か国における宗教意識調査」のデータ分析において試みた「因子分析と最小空間分析に関する方法論的検討」（2018 a）を参照されたい。

(4) RTC 尺度を構成する 17 項目のうち、二項目は、「変化に対する抵抗」に対して、それとは逆の／反対の／対立する方向でワーディングがな

された質問文、具体的にいうならば、「変化に対する志向／追求／受容」を表現する質問文となっている。そして、その結果、これら二項目は、SSA マップにおいては、ほかの諸項目と明らかに「差異性 (distinctiveness)」あるいは「変則性 (irregularity)」を示すものとなっている。

Oreg et al. は、このような諸知見は、「変化に対する資質的な抵抗」尺度の構成において、「変化に対する抵抗」とは逆の／反対の／対立する方向でワーディングがなされた項目を用いることの適切性 (appropriateness)」に関する方法論的な関心を惹起する、と述べている。しかし、Oreg et al. は、このような問題の所在を指摘するにとどまり、それをさらに、社会調査の実践的な方法論上の問題へと発展させ、その解決のための具体的なアイデアを提示する、というところにまでは到らなかった。

## 6. CFA/MGCFA と確証的 SSA の比較をめぐる議論

(1) Oreg et al. は、「変化に対する資質的な抵抗」という構成概念と、それを捉える RTC 尺度が、異なる国々において、等価な意味をもつかどうかを、Oreg, et al. (2008) と同じデータを用いて、CFA・MGCFA と確証的 SSA という二つの統計技法によって、実証的にテストした。その結果、これらの技法によってもたらされた諸知見には「共通部分」とともに「異なる諸側面」があることがわかった。このような結果は、「RTC 尺度」とその根底にある「構成概念」へのより深い洞察を提供するものであるという。

以上の記述は、Oreg, et al. (2011) の結論ともいべきものといえよう。ここで、筆者らの問題関心は、つぎの点にある。それは、CFA・MGCFA と確証的 SSA という二つの統計的技法の併用のもたらす諸知見が「より深い洞察 (greater insights) を提供する」という記述の部分である。では、なぜ、筆者らは、このような部分に関心をもつのかというと、それは、筆者らが、前稿（真鍋、前田、清水、2021）において、この研究領域における文献を、①統計的方法に関する文献、②国際比較／文化比較調査の実践と方法に関する文献、③統計的方法の国際比較／文化比較

調査のデータ分析への応用に関する文献、の三種類に分類したと関わってくる。この基準からするならば、Oreg, et al. (2011) は、③の「応用研究」として分類される文献といえる。じつは、筆者らが、この研究領域において、貢献しようと見定めているところは、この③に分類される研究にはかならない。そして、この③に分類される研究の顕著な特徴の一つが上述の Oreg らのこの記述の部分、つまり「より深い洞察」という表現／用語の部分にある。それは、「より深い洞察」という表現／用語が、統計学の領域における technical language ではなく、社会科学の領域における substantive language である——Jacques Herman (日本語訳、1993) の「社会学の言語 (Les Langage de la sociologie)」という考え方を踏まえている——からである。筆者らは、このような「応用研究」の文献における、いわゆる nontechnical language——統計学の専門用語以外の用語——に関心をもっている。それは、それらが「応用研究」の領域における「新しい飛躍の種の発見」を示唆するものとなるのではなからうか、と考えているからにはかならない。

では、Oreg et al. は、この「より深い洞察」という用語で、何を意味しようとしたのであろうか。その手がかりは、これら二つの統計的技法によって明らかにされてきた「共通部分」と「異なる諸側面」に関する Oreg et al. の記述を精査することによって得られるのではなからうか。筆者らの事前の期待からするならば、そのような「より深い洞察」は、一方において、人びとの「変化に対する抵抗」という心理特性あるいはパーソナリティ特性についての、本研究においてなされた実証科学的な探究を、substantive にさらに深めていく契機となる内容とともに、他方において、以上の二種類の統計的技法の特徴についての、その応用の視座からする理解を、methodological により深めていく契機となる内容を、同時に、含むもののはずであるということであった。

ところが、Oreg et al. (2011) についての、以上の二点に焦点を合わせた文献内容の精査から、つぎのことが明らかになってくる。まず、①については、いくつかの知見についての仮説的な考察とともに、先行研究における substantive な理論

——例えば、「心理的なりアクタンスの理論」——との関連づけが試みられるものの、それ以上の議論の展開はなされていない。

つぎに、②については、CFA/MGCFA と確証的 SSA によって異なる結果が導かれるケースについて、その原因をめぐって、CFA/MGCFA の場合はデータの「正規性」あるいは「線形性」という問題がかかっている可能性を示唆するとともに、確証的 SSA の場合はある項目の SSA マップにおける「逸脱」がデータ全体に「歪んだ結果」をもたらすという問題がかかっている可能性を示唆している。しかし、このような可能性の示唆は、どこまでも示唆にとどまるものであって、そのことをさらに方法論的に探索する試みはなされていない。

以上から、Oreg, et al. (2011) においては、「より深い洞察」が、具体的にどのような内容であるかについては、十分に議論されないままとなっている。この点は、筆者らの主張する「応用研究」における「創造性 (creativity)」という視座からして、きわめて残念な結果であるといわざるをえない。

(2) Oreg et al. では、繰り返し構成概念の測定尺度の妥当性 (validity) について言及されており、Oreg et al. の MGCFA や SSA による分析の結果は、全体として「変化に対する資質的な抵抗」の「国際比較上の妥当性 (cross-national validity)」を示唆するものであると結論づけている。しかし、これらの技法での検討によって確認されるのは、尺度の構造妥当性 (scale's structural validity) に限られている点に注意すべきだろう。もし、異なる集団において、(質問文の原文が同じという意味で) 同一の心理的な事柄について尋ね、その回答の構造 (複数の質問項目間の回答の相互関連性) に共通性が認められたとしても、集団間でそうした心理的な事柄の帰結としての行動が全く異なるとすれば、それは同一の現象を測っていると言えるだろうか。

(3) 尺度を開発するという場合には、その出発点において、使用する質問項目の意味するところが、対象となる特定の概念に関わるものであるこ

とが自明であるか、少なくともそう解釈すること自体の合理性 (reasonability) が要求される。しかし、そうした自明性や合理性は、研究者が属する特定の文化に立脚したものである。このことは表面的妥当性 (face validity) の問題が国際比較／文化比較調査においていかに扱われるべきかという点を指摘したものであるとも言える。その意味では、測定尺度の「国際比較上の妥当性」を問題とするにあたっては、MGCFA や SSA によって異なる国・文化において同様の測定モデルが利用可能であるという結果が得られたとしても、その結果自体の妥当性について考慮しなければならないのではないのか。通常は、表面的妥当性や、あるいは尺度項目の内容的妥当性 (content validity) についての検討は、国際比較を考える前の単一の国／文化での開発を想定した場合にすでに行われて、基本的にクリアしているもの、との前提があると考えられる。それがそのまま複数の国での妥当性も自動的に保証するというのではないのではないのか、という点を筆者らは指摘しておきたい。それでは国際比較の文脈における表面的妥当性、内容的妥当性を確認する作業はどのようなものであるべきなのだろうか。繰り返しになるが、MGCFA による測定不変性の議論は、構造的な妥当性の共通性について述べているだけであり、表面的妥当性や内容的妥当性が、同じ構成概念の名称と、翻訳された測定項目によって、それぞれの国で並行して確認されていることを以てクリアされたとするのか、それとも、さらに何か条件が必要であるのか、引き続き検討されるべきであろう。

(4) 以上に述べてきた二種類の技法もたらす結果として、Oreg et al. は、つぎのような諸知見を報告している。

- ・CFA によって、「変化に対する抵抗」の国際比較の妥当性が確認された。

- ・MGCFA によって、17か国における partial metric invariance の証拠が獲得された。

- ・確証的 SSA によって、RTC の四つの次元——routine seeking, emotional reaction, short-term focus, cognitive rigidity——に対応する諸項目は、SSA マップの四つの領域にクラスター化 (clus-

tered) された。

さて、ここで問題は、確証的 SSA による知見を「諸項目は四つの領域にクラスター化された」と表現している点である。なぜ、この点を問題と考えるかという、SSA という技法の基本的な性格からして、「領域はクラスターではない (Regions are not clusters)」(Shye, Elizur with Hoffman, 1994) とされているからにほかならない。

この点をもう少し具体的に解説しておくならば、それは、つぎようになる。例えば、SSA マップにおいて、空間分割された、ある領域内にプロットされる点の数が少ない場合、それは、そのような領域に対応する質問項目の数が少ないからであって、事前に多くの質問項目を準備しておくならば、SSA マップのその領域はプロットされる点でいっぱいになるはずである。こうして、SSA マップにおいては、そのような点 (質問項目) の「クラスター (群れ／まとまり)」ではなく、そのような点 (質問項目) が布置される「領域」にこそ意味があるのである。この点は、じつは、事前のファセット・デザインにおいて準備される「質問項目と Structuple との対応表」——例えば、真鍋 (2018 a, p.27) ——と SSA マップ——同じく p.32——と対応させることで明らかとなる。

以上からして、Oreg et al. は、なぜ、SSA マップの空間分割において、「領域」と「クラスター」を明確に区別しなかったのであろうか。それは、Oreg et al. が、SSA という技法を、ほかの統計的技法の場合と同じように、単独のデータ分析の技法として取り扱い、それをファセット・アプローチ——ファセット・デザイン、ファセット・アナリシス、ファセット・セオリー——の三位一体の知の体系のなかに位置づけようとしなかったからではなかろうか。この点は、Oreg et al. の確証的 SSA の利用の最大の問題点といわなければならない。

(5) Oreg et al. は、「CFA・MGCFA と確証的 SSA との比較をめぐる議論」を締めくくるに当たって、「これら二つの方法による諸知見の違いは、それら二つの方法の相互に補い合う価値 (complementary value) に光を当てる」(p.239) と

述べている。

しかし、筆者らからするならば、このような表現も、相変わらず、必ずしも明確なものとはいえない。それは、「それら二つの方法の相互に補い合う価値」の内容が具体的に書かれていないからにはかならない。確かに、Oreg et al. (2011) においては、これら二つの方法によって導かれた諸知見には「共通する部分」とともに「異なる部分」があることが明らかにされるとともに、そのような結果には、それぞれの技法の特徴がかかわっているであろうことが示唆された。しかし、ここでも、そのような示唆は、どこまでも示唆にとどまっておき、そのことを、さらに technical に解明する試みがなされるわけでもなく、また、それらの諸知見をめぐって「応用研究」としての独自の substantive な視座が提示されるわけでもなかった。繰り返しになるが、この点も、やはり惜しまれる結果といわなければならない。

(6) 最後に、Oreg et al. は、本研究で残された、つぎのようなもう一つの方法論的な問題をあげている。

本研究は、「変化への資質的な抵抗」という構成概念は、複数の国／文化において、「個人的レベル (individual level)」で、その意味 (meaning) が共有 (share) されていることを示唆した。こうして、RTC 尺度はさまざまな国／文化において、個々人の「変化への資質的な抵抗」を比較するために利用可能となった。では、「変化への資質的な抵抗」の「集合的レベル (aggregate level)」においても、同じように、さまざまな国／文化の比較は可能であるかという、それは全く別の問題である。「個人的レベル」における「測定の特価性」は、より高次の諸単位間の比較を行なうための「必要な基準 (necessary creation)」ではあっても、「十分な基準 (sufficient)」ではない (van de Vijver and Poortinga, 2002) という。

確かに、Oreg et al. の指摘するように、「個人的レベル」から「集合的レベル」への実証的研究の展開のためには、一方の「理論的な枠組みの準備」と、他方の「データ分析の手続きの開発」という二つの側面における知的創造が要請されるこ

とになる。このような問題に関しては、Hofstede (1980 = 1984)、Hofstede, Hofstede and Minkov (2010 = 2013)、Schwartz and Bilsky (1987)、Schwartz (1992, 1994)、Schwartz et al. (2012)、Triandis (1995 = 2002)、Smith and Bond (1998 = 2003) をはじめとして、かなりの研究の蓄積がある。こうして、ここでも、このテーマに関する文献レビューにもとづく「研究の全体像の把握」と「新しい飛躍の種の発見」が、今後の重要な課題となるのである。

#### 文献

- Asparouhov, Tihomir and Muthén, Bengt, O. (2014). Multi-Group Factor Analysis Alignment. *Structural Equation Modeling*, 21(4), 495-508.
- Amarr, Reuven and Toledano, Shlomo (2001). *Hudap Manual with Mathematics and Windows Interface* (2<sup>nd</sup> ed.). The Hebrew University of Jerusalem.
- Bauer, Daniel J. (2005). The Role of Nonlinear Factor-to-Indicator Relationships in Tests of Measurement Equivalence. *Psychological Methods*, 10(3), 305-316.
- Ben-Shalom, Uzi and Horenczyk, Gabriel (2003). Acculturation Orientations: A Facet Theory Perspective on the Bidimensional Model. *Journal of Cross-Cultural Psychology*, 34(2).
- Blalock, Jr. Hubert M. (1989). Toward Cumulative Knowledge: Theoretical and Methodological Issues. In Eulau H. ed., *Crossroads of Social Science*, Agathon Press.
- Borg, Ingewer and Lingoes, James C. (1987). *Multidimensional Similarity Structure Analysis*. Springer.
- Borg, Ingewer and Shye, Samuel (1995). *Facet Theory: Form and Content*. Sage Publications.
- Borg, Ingewer and Groenen, Patrick J. F. (2005). *Modern Multidimensional Scaling* (2<sup>nd</sup> ed.). Springer.
- Braun, Michael and Johnson, Timothy P. (2010). An Illustrative Review of Techniques for Detecting Inequivalences. In Harkness, J. et al. eds. *Survey Methods in Multinational, Multiregional, and Multicultural Context*. John Wiley and Sons.
- Brehm, Jack, W. (1966). *A Theory of Psychological Reactance*. Academic Press.
- Cieciuch, Jan, Davidov, Eldad, Schmidt, Peter, Algesheimer, René, and Schwartz, Shalom H. (2014). Comparing Results of an Exact vs. an Approximate (Bayesian) Measurement Invariance Test: A Cross-



- Country Illustartion with a Scale to measure 19 Human Values. *Frontiers in Psychology*, 5, 1-10.
- Cieciuch, Jan, Davidov, Eldad, Schmidt, Peter, Algesheimer and René (2019). How to Obtain Comparable Measures for Cross-National Comparisons. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 71, 157-186.
- Cohen, Arie (2008). The Underlying Structure of the Beck Depression Inventory II: A Multidimensional Scaling Approach. *Journal of Research in Personality*, 42, 779-786.
- Davidov, Eldad, Dalter, Georg, Schmidt, Peter, and Schwartz, Shalom H. (2011). Testing the Invariance of Values in the Benelux Countries with the European Social Survey: Accounting for Ordinality. In Davidov et al. eds. *Cross-Cultural Analysis*. Routledge.
- Davidov, Eldad, Meuleman, Bart, Cieciuch, Jan, Schmidt, Peter, and Billiet, Jaak (2014). Measurement Equivalence in Cross-National Research. *The Annual Review of Sociology* 40, 55-75.
- De Beuckelaer, Alain and Swinnen, Gilbert (2011). Biased Latent Mean Comparisons Due to Measurement Noninvariance: A Simulation Study. In Davidov et al. eds. *Cross-Cultural Analysis*. Routledge.
- 福岡伸一 (2010). 『ルリボシカミキリの青』 文藝春秋.
- Guttman, Louis (1968). A General Nonmetric Technique for Finding for the Smallest Coordinate Space for a Configuration of Points. *Psychometrika*, 33, 469-506.
- (1980). Recent Structural Laws of Human Behavior. (=真鍋一史抄訳「人間行動に関する最近の構造的諸法則」)『慶應義塾大学新聞研究所年報』14, 1-16.
- Harkness, Janet A. et al. eds. (2003). *Cross-Cultural Survey Methods*. Wiley.
- 林知己夫、鮑戸弘共編 (1976). 『多次元尺度解析法』サイエンス社.
- Herman, Jacques 著, 原田哲・樋口義弘訳 (1993). 『社会学の言語』白水社.
- 廣松渉 (2017). 『世界の共同主観的存在構造』岩波文庫.
- Hofstede, Geert (1980). *Culture's Consequences*, Sage. (=1984, 萬成博、安藤文四郎ほか訳『経営文化の国際比較』産業能率大学出版部.)
- Hofstede, Geert, Hofstede, Gert Jan, and Minkov, Michael (2010). *Cultures and Organizations: Software of the Mind*, (3<sup>rd</sup> ed.) McGraw-Hill. (=2013, 岩井 八郎、岩井紀子訳『多文化世界』有斐閣.)
- Jöreskog, Karl G. (1969). A General Approach to Confirmatory Maximum Likelihood Factor Analysis. *Psychometrika* 34, 183-202.
- (1971). Simultaneous Factor Analysis in Several Populations. *Psychometrika* 36, 409-426.
- 狩野裕・三浦麻子 (2002). 『グラフィカル多変量解析 - AMOS、EQS、CALIS による目で見る共分散構造分析 -』現代数学社.
- 狩野裕 (2002 a). 「構造方程式モデリングは、因子分析、分散分析、パス解析のすべてにとって代わるのか？」『行動計量学』29.2, 138-159.
- (2002 b). 「まえがき」木村通治、真鍋一史、安永幸子、横田賀英子『ファセット理論と解析事例』ナカニシヤ出版.
- 木村通治、真鍋一史、安永幸子、横田賀英子 (2002). 『ファセット理論と解析事例』ナカニシヤ出版.
- Kline, Rex. B. (1998). *Principles and Practice of Structural Equation Modeling*. Guilford Press.
- 真鍋一史 (1993). 『社会・世論調査のデータ解析』慶應義塾大学出版会.
- (1998). 『国際イメージと広告』日経広告研究所.
- Manabe, Kazufumi (2001). *Facet Theory and Studies of Japanese Society*, Bier'sche Verlangsanstalt, Bonn, Germany.
- 真鍋一史 (2003). 『国際比較調査の方法と解析』慶應義塾大学出版会.
- (2012). 「社会科学はデータ・アーカイブに何を求めているか」『社会と調査』(社会調査協会)有斐閣、8, 16-23.
- (2016). 「ファセット・アプローチと価値観研究——Lois Guttman とその共同研究者の知的世界の探索——」『関西学院大学社会学部紀要』123, 9-32.
- (2018 a). 「宗教意識の国際比較——『因子分析』と『最小空間分析』に関する方法論的検討——」『青山地球社会共生論集』3, 1-52.
- (2018 b). 「〈研究ノート〉Schwartz の『価値観研究』の方法論的な検討」『関西学院大学社会学部紀要』129, 75-94.
- 真鍋一史, Jagodzinski, Wolfgang, Davidov, Eldad, Dülmer, Hermann, and Hommerich, Carola (2020 a). 「S. Schwartz の概念枠組みにもとづく価値観の国際比較——ドイツと日本における『大学生調査』のデータ分析——」『関西学院大学社会学部紀要』133, 87-107.
- 真鍋一史, Jagodzinski, Wolfgang, Davidov, Eldad, Dül-

- mer, Hermann, and Hommerich, Carola (2020 b). 「S. Schwartz の概念枠組みにもとづく価値観の国際比較 (Ⅱ) —— ドイツと日本における『大学生調査』のデータ分析——」『関西学院大学社会学部紀要』135: 1-20.
- 真鍋一史 (2021). 「ファセット・アプローチとウェルビーイングの研究——Louis Guttman とその共同研究者の足跡——」『関西学院大学社会学部紀要』136, 1-28.
- 真鍋一史・前田忠彦・清水香基 (2021). 「国際比較／文化比較調査における測定 の等価性／不変性の研究：多集団確証の因子分析 (MGCFA) を中心として」『関西学院大学社会学部紀要』137, 1-28.
- Meitinger, Katharina, Davidov, Eldad, Schmidt, Peter, and Braun, Michael (2020). Measurement Invariance : Testing for It and Explaining Why It is Absent. *Survey Research Methods* 14(4) 345-349.
- Milbrath, Lester W. (1965). *Political Participation : How and Why Do People Get Involved in Politics?* Rand McNally & Company (=1976, 内山秀夫訳『政治参加の心理と行動』早稲田大学出版部.)
- Oreg, Shaul (2003). Resistance to Change : Developing an Individual Differences Measure. *Journal of Applied Psychology*, 88, 680-693.
- (2006). Personality, Context, and Resistance to Organizational Change. *European Journal of Work and Organizational Psychology*, 15, 73-101.
- Oreg, Shaul, Bayazit, Mahmut, Vokola, Maria..... van Dam, Karen (2008). Dispositional Resistance to Change : Measurement Equivalence and the Link to Personal Values Across 17 Nations. *Journal of Applied Psychology*, 93(4), 935-944.
- Oreg, Shaul, Nevo, Ofra, Metzger, Hila, Leder Naftali, and Castro, Dotan R. (2009). Dispositional Resistance to Change and Occupational Interests and Choices. *Journal of Career Assessment*, 17(3), 312-323.
- Oreg, Shaul et al. (2011). Measurement Equivalence of the Dispositional Resistance to Change Scale. In Davidov et al. ed., *Cross-Cultural Analysis : Methods and Applications*. Routledge.
- Payne, Geoff and Payne, Judy (2004). *Key Concepts in Social Research*. Sage. (=2008, 高坂健次ほか訳『キーコンセプト ソーシャルリサーチ』新曜社.)
- Schwartz, Shalom H. and Bilsky, Wolfgang (1987). Toward a Universal Psychological Structure of Human Values, *Journal of Personality and Social Psychology* 53(3), 550-562.
- Schwartz, Shalom H. (1992). Universal in the Content and Structure of Values : Theory and Empirical Tests in 20 Countries. In Zanna M. ed., *Advance in Experimental Social Psychology*, 25, Academic Press.
- Schwartz, Shalom H. (1994). Are There Universal Aspects in the Content and Structure of Values? *Journal of Social Issues*, 50, 19-45.
- Schwartz, Shalom H. et al. (2012). Refining the Theory of Basic Individual Values. *Journal of Personality and Social Psychology*, 103(4), 663-686.
- 清水和秋 (1994). 「Jöreskog と Sorbom によるコンピュータ・プログラムと構造方程式モデル」『関西学院大学社会学部紀要』25(3), 1-41.
- Shye, Samuel ed. (1978). *Theory Construction and Data Analysis in the Behavioral Sciences*. Jossey-Bass Publishers.
- Shye, Samuel (1991). Faceted Smallest Space Analysis (FSSA) (version 3.01 ; computer program). Louis Guttman Israel Institute of Applied Social Research.
- Shye, Samuel, Elizur, Dov with Hoffman, Michael (1994). *Introduction to Facet Theory*. Sage Publications.
- Shye, Samuel (1997). FSSAWIN (version 1.0) Louis Guttman Israel Institute of Applied Social Research.
- Smith, Peter B. and Bond, Michael H. (1998). *Social Psychology across Cultures*. (2<sup>nd</sup> ed.) Prentice Hall Europe. (=2003, 笹尾敏明、磯崎三喜年訳『グローバル化時代の社会心理学』北大路書房.)
- Steenkamp, Jan-Benedict E. M. and Baumgartner, Hans (1998). Assessing Measurement Invariance in Cross-National Consumer Research. *Journal of Consumer Research*, 25, 78-90.
- 豊田秀樹 (2014). 『共分散構造分析 [R 編] —— 構造方程式モデリング——』東京図書.
- Triandis, Harry C. (1995). *Individualism and Collectivism*, Westview Press. (=2002, 神山貴弥、藤原武弘ほか訳『個人主義と集団主義』北大路書房.)
- Vandenberg, Robert J. and Lance, Charles E. (2000). A Review and Synthesis of the Measurement Invariance Literature : Suggestions, Practices, and Recommendations for Organizational Research, *Organizational Research Methods*, 3(1), 4-70.
- van de Vijver, F. J. R., and Poortinga, Y. H. (2002). Structural Equivalence in Multilevel Research. *Journal of Cross-Cultural Psychology*, 33, 141-156.
- Wallace, Walter L. (1971). *The Logic of Science in Sociology*. Aldine. (=2018, 渡辺深訳『科学論理の社会学——「ワラスの輪」というモデル——』ミネル

ヴァ書房.)

Welkenhuysen-Gybels, Jerry and Billiet, Jaak (2002). A Comparison of Techniques for Detecting Cross-Cultural Inequivalence at the Item Level. *Quality and Quantity*, 36, 197-218.

Welkenhuysen-Gybels, Jerry (2004). The Performance of Some Observed and Unobserved Conditional Invariance Techniques for the Detection of Differential Item Functioning. *Quality and Quantity*, 38, 681-702.

# Statistical Techniques for Assessing Measurement Comparability in Cross-National/Cross-Cultural Research: Multiple Group Confirmatory Factor Analysis and Confirmatory Smallest Space Analysis

Kazufumi MANABE

Tadahiko MAEDA

Koki SHIMIZU

## ABSTRACT

This paper is a sequel to our previous paper entitled “Measurement Equivalence/Invariance in Cross -National/Cross-Cultural Research” (Kwansei Gakuin University School of Sociology Journal, No.137). Previously, we examined the development of studies on measurement equivalence/invariance in cross-national/cross-cultural research by focusing on the statistical technique called multiple group confirmatory factor analysis (MGCFA).

Diverse statistical techniques have been developed for assessing measurement comparability such as confirmatory factor analysis (CFA), MGCFA, multidimensional scaling (MDS), item response theory (IRT), and latent class analysis (LCA). We focused on MGCFA in the previous paper because this technique has been the most widely used in cross-national/cross-cultural research.

However, any statistical technique has its advantages and disadvantages. Here, we attempt to methodologically examine these techniques to test measurement compatibility. As the first step, we deal with MGCFA and confirmatory SSA. We have two pioneering research examples that compare these two techniques.

1. A series of studies on psychological trait called “dispositional resistance to change” was initiated by Shaul Oreg.
2. A series of studies on “basic human values” was initiated by Shalom Schwartz.

In this paper, we begin with the first example. This study is the first instance to use both MGCFA and confirmatory SSA to test the compatibility of psychological trait, dispositional resistance to change. This is the discussion of the first section of this paper.

The second section summarizes the basic ideas, main features, and practical procedures of the MGCFA and confirmatory SSA techniques.

The third section presents a literature review of the empirical research on “dispositional resistance to change” by Oreg. et al (2011). Their research question, procedures, and results of data analysis are traced and examined in detail.

The final section presents a methodological discussion of the problems and prospects of their research. It also suggests a future direction for development in this field of study.

**Key Words:** measurement compatibility, cross-national / cross-cultural research, statistical techniques, multiple group confirmatory factor analysis (MGCFA), confirmatory smallest space analysis (SSA)