

国際比較／文化比較のための MGCFA の応用研究の実践的な展開*

——Schwartz の「価値観理論」の枠組みにもとづく調査研究の事例をとおして——

真 鍋 一 史**
前 田 忠 彦***
清 水 香 基****

I. はじめに

——文献研究の目的と意義——

本稿は、(1) Cieciuch, Davidov, Schmidt, and Algesheimer (2019). How to Obtain Comparable Measures for Cross-National Comparisons. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* を中心とする関連諸文献にもとづいて、(2) MGCFA (multiple group confirmatory factor analysis: 多集団確認的因子分析) と呼ばれる統計的技法の、質問紙調査のデータ分析への「応用研究 (applied studies)」が、Exact MI Test→Partial MI Test→Approximate (Bayesian) MI Test→Alignment Optimization というように発展してきたことを、(3) Schwartz の「価値観理論」の構築と、その後の発展に関する調査研究の諸事例——学術誌に発表された文献にかぎる——に焦点を合わせて、詳細に跡付けていく文献研究の試みである。

このような本稿の内容について、とくにその「目的・意義」という点から、もう少し説明しておきたい。

まず、(1) については、なぜ、本稿において Cieciuch et al. (2019) を取り上げるのかについて答えなければならない。

統計数理研究所における「共同研究」として始められた3人による「国際比較調査／文化比較調査における測定等の等価性／不変性の研究」は、欧米の学術誌を中心に発表されてきた主要な方法論的な諸文献の収集と精査から出発した。そのような文献研究の成果は、この領域における研究動向を概観するいわゆる「レビュー論文」としてまとめてきた(真鍋、前田、清水、2021、2022 a、2022 b)。この過程で、われわれは、すでにして、Cieciuch et al. (2019) に注目し、評価していたのである。その「注目と評価」のポイントは、つぎの点にあった。それは、何よりもまず、この文献が通常の意味における「レビュー論文」ではなかったということである。Cieciuch らは、Schwartz の「価値観理論」の出現とともに、そのテーマにそれぞれ独自の仕方、あるいは共同研究の形で取り組み、多くの研究成果を蓄積してきた。Cieciuch らによって、価値観調査のデータ分析という形で進められてきた一連の実証的研究は、本稿でいうところの「応用研究」の典型的な事例であった。そして、かれらの方法論的な立場——MGCFA や SEM (Structure Equation Modeling: 構造方程式モデリング) などを主要な「統計的技法」として用いる方法論的な立場——からするさまざまなデータ分析の過程とその諸結果を、自ら観察し、整理し、記述した成果こそが、この文献

*キーワード：国際比較／文化比較、多集団確認的因子分析、応用研究、Schwartz の価値観理論、決定木図、Alignment Optimization、Mplus

**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

にほかならない。この文献が、通常の「レビュー論文」ではないとした所以がまさにこの点にある。こうして Ciecuch at al. (2019) は、きわめてユニークな「応用研究」であるとともに、きわめてユニークな「レビュー論文」であるといわなければならないのである。

つぎに、(2) については、なぜ、本稿では MGCFA (や SEM) といった「統計的技法」に焦点を絞るのかについて答えなければならない。確かに、この研究領域におけるもう1つのユニークな「応用研究」の集大成である Davidov, Schmidt, Billiet, and Meuleman (2018) においては、現在、MGCFA を初めとして、MDS (multidimensional scaling: 多次元尺度法) / SSA (smallest space analysis: 最小空間分析)、IRT (item response theory: 項目反応理論)、LCA (latent class analysis: 潜在クラス分析) などさまざまな技法が開発されてきているとして、それらの「統計的諸技法」を用いたさまざまな「応用研究」の事例が取りあげられている。それにもかかわらず、本稿では、Ciecuch at al. (2019) にもとづいて、MGCFA に焦点を絞っている。その理由としては、つぎの2点をあげることができる。

①この研究領域におけるさまざまな文献において、「国際比較／文化比較研究における測定と比較可能性を確認するための統計的技法としては、MGCFA がこれまで最もよく利用されてきている」ことが指摘されている。

②本稿における筆者らの「問題関心」、そして今回の共著論文の執筆の「動機づけ」は、社会科学の領域における「統計的技法」——とくに MGCFA と呼ばれる「統計的技法」——の利用と、そのような「統計的技法」の「応用研究」の新しい創造性／可能性の探究といったところにある。これまで、さまざまな substantive なテーマへの「統計的技法」の「応用研究」が試みられる場合、そのめざすところは、そのようなテーマをめぐる諸研究の「方法論的な質の向上」であるとされてきた (例えば、Steenkamp and Baumgartner, 1998, p.88)。確かに、実証科学において、そのような「方法論的な質の向上」がきわめて重要な課題であることは間違いない。しかし、「応用研究」

の課題は、決してそこにとどまるものではない。このような「問題関心」から、筆者らは「応用研究」の新しい創造性／可能性の探究を進めてきた。

そして、そのような方法論的な「問題関心」の線上で、Schwartz, Davidov, Ciecuch らによる「価値観理論」の構築と展開は、まさにそのような「応用研究」によらなければ達成されることになかった新しい「知の創造」の試みであったという理解に到達した。こうして「応用研究」は、このように社会科学の領域における新しい「理論」の構築と、その「作り直し」(Merton, 1957=1961) をも可能にするものであることが確認されたのである (真鍋、前田、清水、2022 b)。

そしてさらに、このような「応用研究」の新しい創造性／可能性のもう1つの方向を示したのが Ciecuch at al. (2019) にほかならない、ということもわかってきた。では、その方向がどのようなものであるかということ、それは「国際比較研究／文化比較研究における測定の不変性の実証的なテストの実践 (practice) 過程」が、「その手続きと判断の体系的な流れ図」(図1) という形で明細化 (specify) されたということである。

筆者は、ここで「実践」という用語を用いた。しかし、それは筆者の独自の表現というわけではない。例えば、対応分析の領域で、世界的に広く利用されてきたテキストの1つに Michael Greenacre (2017=2020) の *Correspondence Analysis in Practice*. Taylor and Francis Group があるが、そこでは practice (実践) という用語が使われている。

さらに、日本における独自の「数量化理論」の考案者である林知己夫は、1970年代から1980年代にかけて、編著の形で、『多次元尺度解析法』と『多次元尺度解析法の実際』の2冊の研究書を出版している。前者では、多次元尺度法の諸理論が、そして後者では、そのような統計解析の実際が扱われている。ここで注目しておきたいのは、この後者の本の序文において林 (1984) が、つぎのように述べている点である。

「データによる現象解析という立場に立ち、多次元尺度解析をどう用いたら、これまで明らかに把握できなかったことがわかってく

るか、ということを中心に据えて論じてみた。このとき、データの性格をもあわせて記載し、応用が機械的にならないような配慮を加えた。(中略) データのとり方からデータ解析に到るまでの考え方やそれに添う方法、さらにそれを土台に情報をとり出す行き方が行動計量学のソフトな一面を示していると思っている。表題を『多次元尺度解析法の実際』とした所以である。」(pi)

本稿で、筆者らの「実践」という用語は、Grenacre の “practice”、そして、林らの「実際」とほぼ同じ意味内容を含むものである。とくに林が「応用が機械的にならないように」と書いている点は、重要である。それは、社会科学の領域における「統計的技法」の導入とその応用についてのあるべき方向の示唆としても読み取ることができるからである。そして、そのような「応用が機械的にならないための方略」を練りあげ、それを研究者間で共有できる形で提示するというのが、林らの意図だったのであろう。

そして、そうだとするならば、そのような「方略」は、まさに Ciecuch の「測定不変性のテストをめぐる手続きと判断の流れ図」と軌を一にする考え方であったといえるのではなからうか。

こうして、Ciecuch at al. (2019) は、「応用研究」の新しい創造性／可能性のもう1つの方向を示した重要な文献であるといわなければならないのである。

さらに、(3) については、なぜ、本稿においては Schwartz の「価値観理論」の枠組みにもとづく調査研究事例を取り上げるのかについても答えなければならない。

このような「問い」に対しては、つぎの3つの点から答えることができるであろう。

①「価値観」は、社会科学の領域において、広く多くの研究者が「問題関心」を共有することのできるテーマである。それは、「価値観」が、①諸個人、諸集団、諸社会を記述する (describe) ための、そして、②人びとの諸態度、諸意見、諸行動を説明する (explain) ための、概念・次元・尺度であるということを考えるならば、容易に理

解されるであろう。

②Schwartz の「価値観理論」は、Inglehart のそれと並んで、現代における価値観理論の双璧とされ、世界のアカデミック・コミュニティの注目を集めてきている (Datler, Jagodzinski and Schmidt, 2013)。

③現代の社会科学にとって、最も大きな出来事の1つをあげるとするならば、それは質問紙法にもとづく多くの国ぐにを対象とする国際比較調査の出現ということであろう。具体的にいうならば、「世界価値観調査 (World Values Survey: WVS)」「国際社会調査プログラム (International Social Survey Programme: ISSP)」「ヨーロッパ社会調査: European Social Survey: ESS)」などがあげられる。そして、Schwartz の「価値観理論」を実証的に確認するために作成されてきた「質問諸項目」は、早くからこれら国際比較調査の WVS と ESS の2つにおいて取り入れられてきている。じつは、実証科学としての「質問紙調査のデータ分析」という研究領域にとっては、このことはきわめて重要である。なぜならば、WVS や ESS に取り入れられたことで、Schwartz の「価値観諸項目」についての調査データは、一般公開されることになり、これらのデータへのアクセスは、基本的に「だれでも」「いつでも」「どこでも」可能となり、こうしてこの研究領域に関心を持つすべての研究者に対して独自のデータ分析の機会が提供されることになったからである。このようなデータの公開性と公共性の確立をとおして、Schwartz の「価値観理論」の枠組みにもとづく調査研究事例の検討というテーマには大きなアドバンティッジが与えられることになった。具体的にいうならば、Ciecuch at al. (2019) の文献研究にもとづいて、かれらの「応用研究」の問題関心、データ分析の方法、その結果を「追体験」するにとどまらず、われわれ自身が同じデータを用いた独自の分析方法によって、オリジナルな「知の創造」に挑戦することも可能になったのである (国際比較調査とそのデータ公開については、真鍋、2003 を参照されたい)。

II. 測定不変性のテストの実践的な手続きと判断の「決定木図」とデータ分析事例——Schwartzの「価値観理論」の枠組みにもとづく質問紙調査のデータ分析諸事例——

1. 測定不変性のテストの実践プロセス

Cieciuch, Davidov, Schmidt and Algesheimer (2019) においては、測定不変性 (measurement invariance: MI) のテストの実践プロセスについて、「決定木 (decision tree)」の形での明細化 (specification) がなされている。ここでは、この「決定木図」は、測定不変性がデータで確認されない場合——いいかえれば、測定非不変性 (measurement non-invariance) という結果が示された場合——、その結果をどのように判断し、その後どのような手続きをとるかを概観した図であると説明されている。この図については、その説明の後に、() 付きで、Author's own work と記されているものの、そのアイデアの淵源——「決定木」のアイデアの起点、発展、応用など——については何も述べられていない。

以上の説明からするかぎり、この「決定木図」

は、測定不変性のテストの実践の諸段階を整理したものとして理解することができる。

じつは、この「決定木図」は、われわれの共著論文のシリーズにおいては、今回、初めて取りあげるといわけではない。すでに、真鍋・前田・清水 (2022 a) の「MGCFA による分析の具体的な手順」の解説において、このアイデアに注目した。その部分の筆者 (前田) は、そのために、この図を翻訳・翻案したものを独自に作成するとともに、そこで1つの到達点として提示された Alignment Optimization と呼ばれる技法を、このようなテストの実践的な視座からも高く評価されるものとした。

このように、「決定木のアイデア」は、本来は、測定不変性のテストの手順を示した仕様図の提案ともいうべきものである。ところが、この図が提案された Cieciuch et al. の上述の文献においては、それは、この研究領域におけるこれまでのさまざまな文献が焦点を合わせてきた研究内容、つまり、そこで「何が、どのように統計的技法によって、どのように分析され、何が明らかとなったか」をとりまとめて示すための枠組みとしても利用されることになった。そして、その結果として、この図は、測定不変性のテストの実践的な試

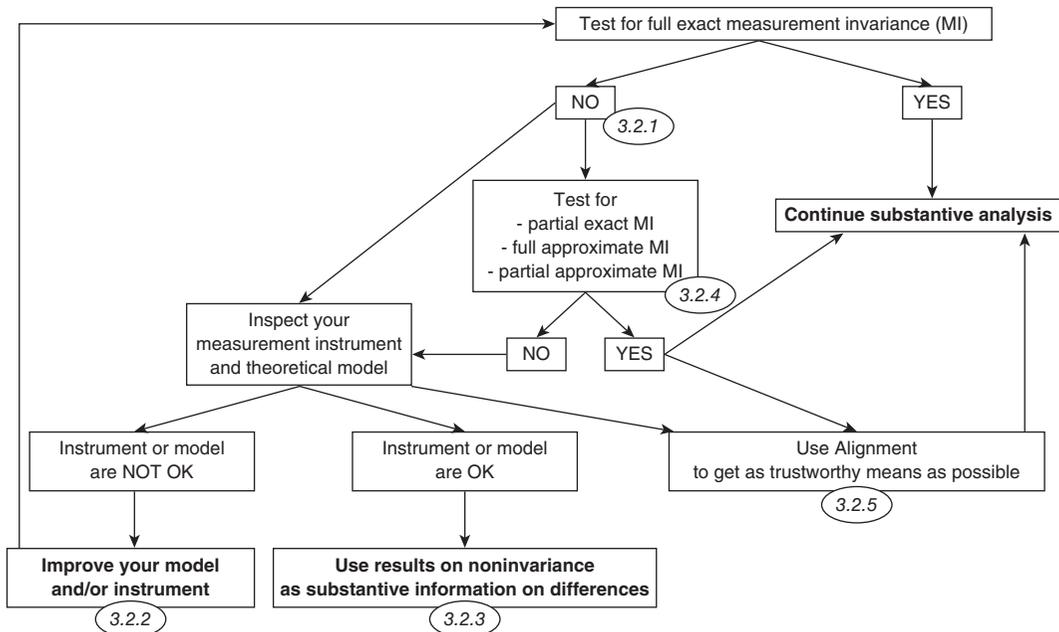


図 1 測定不変性のテストをめぐる手続きと判断の流れ図 (Author's own work)

みが、Exact MI Test→Partial MI Test→Approximate (Bayesian) MI Test→Alignment Optimization というように段階的に発展してきた歴史的な経緯を示すものとなったのである。

本稿では、以上のような Cieciuch et al. の意図にそって、この図とそれにもとづく測定不変性のテストのさまざまな試みの内容を解説するために、再度、この図をオリジナルな形で掲載しておきたい。なお、図中の楕円のなかの数字は、以下の記述のそれぞれのセクション（小節）の番号と対応するものとなっている。

2. 「決定木図」にもとづいて整理された調査データの分析事例

Cieciuch et al. (2019) において、上述の「決定木図」の枠組みを用いて整理されたのは、Schwartz の「価値観理論」の枠組みにもとづく調査データの分析諸事例（諸文献）である。そこで、この試みをめぐって、以下において、いくつかの点を確認しておきたい。

(1) Schwartz の「価値観理論」については、真鍋・前田・清水 (2022b) において、その歴史・内容・発展をめぐって、かなり詳細に解説した。したがって、ここで、それを繰り返すならば、それは、「知的情報の redundancy」の誇りを招くことになるであろう。そこで、以下においては、ここでの Cieciuch et al. (2019) のユニークな「レビュー論文」の試みの紹介に当たって、とくに必要不可欠と考えられる事項についてのみ記しておくにとどめる。

(2) この「レビュー論文」のねらいは、Schwartz らの「価値観理論」そのものの解説というところにあるのではなく、Schwartz らの「価値観研究」をとおして確立されてきた「価値観諸項目 (question items)」——具体的にいうならば、質問紙調査 (questionnaire survey) という方法での計量的な測定に用いられる「価値観諸項目 (measurement instruments)」——は、国際比較／文化比較というコンテキストにおいても利用可能なものであるかどうか、いかにすれば、測定の不変性が確認されるものであるかどうか、という

ところにある。そして、このような確認のために利用される統計的技法が MGCFA にほかならない。

(3) では、MGCFA と呼ばれる統計的技法を用いた、Schwartz の「価値観諸項目」の国際比較／文化比較の可能性の確認、いかにすれば、そのような国際比較／文化比較のコンテキストにおける測定不変性の確認は、どのような手続きと判断にもとづいて実践されるであろうか。

(4) このような測定不変性のテストに当たっては、まず Schwartz の「価値観諸項目」を用いて国際比較／文化比較の視座から実査がなされた質問紙調査のデータが必要となる。Schwartz の「価値観諸項目」のセットにはさまざまな版 (version) があり、それらを用いた国際比較／文化比較の試みには、さまざまな調査プログラム／プロジェクトがある。Cieciuch et al. (2019) では、21 の質問諸項目からなる PVQ-21 (Portrait Value Questionnaire-21) と呼ばれる短縮版 (shortened version) 質問紙を用いた「ヨーロッパ社会調査 (European Social Survey: ESS) の第1回～第7回の調査データが取りあげられた。PVQ-21 については、Schwartz (2003) と Schwartz and Cieciuch (2016) を、そして、ESS については、www.europeansocialsurvey.org を参照されたい。

(5) データ分析に際して必要となる PVQ-21 に関する諸情報が、表1にまとめて提示されている。この表の読み方については、最上段の 1.Self-enhancement—Achievement の部分を例にあげて説明するならば、以下ようになる。

①初めの1という番号は、Schwartz の10の価値観 (Schwartz の用語でいうならば、10の価値観のスタイル) の1つということであって、1という数字に特別の意味があるわけではない。

②Self-enhancement—Achievement という部分については、まず、前者の“Self-enhancement”は、Schwartz の4つの「高次の価値観 (higher-order values)」の1つである“Self-enhancement”と呼ばれる価値観ということの意味している。そして、残りの3つの「高次の価値観」は、“Self-

表1 「4つの高次の価値観」と「10の基本的な人間の価値観」と「ESSのPVQ-21の21価値観諸項目」との対応表

Item label	Items
1. Self-enhancement—Achievement	
Ipshabt	4. It's important to her to show her abilities. She wants people to admire what she does
Ipsuces	13. Being very successful is important to her. She hopes people will recognize her achievements
2. Self-enhancement—Power	
Imprich	2. It is important to her to be rich. She wants to have a lot of money and expensive things
Iprspot	17. It is important to her to get respect from others. She wants people to do what she says
3. Self-transcendence—Benevolence	
Iphlplp	12. It's very important to her to help the people around her. She wants to care for their well-being
Iplylfr	18. It is important to her to be loyal to her friends. She wants to devote herself to people close to her
4. Self-transcendence—Universalism	
Ipeqopt	3. She thinks it is important that every person in the world should be treated equally. She believes everyone should have equal opportunities in life
Ipu drst	8. It is important to her to listen to people who are different from her. Even when she disagrees with them, she still wants to understand them
Impenv	19. She strongly believes that people should care for nature. Looking after the environment is important to her
5. Conservation—Conformity	
Iprfrule	7. She believes that people should do what they're told. She thinks people should follow rules at all times, even when no-one is watching
Ipbhprp	16. It is important to her always to behave properly. She wants to avoid doing anything people would say is wrong
6. Conservation—Tradition	
Ipmo d st	9. It is important to her to be humble and modest. She tries not to draw attention to herself
Imprtrad	20. Tradition is important to her. She tries to follow the customs handed down by her religion or her family
7. Conservation—Security	
Impsafe	5. It is important to her to live in secure surroundings. She avoids anything that might endanger her safety
Iprstrgv	14. It is important to her that the government ensures her safety against all threats. She wants the state to be strong so it can defend its citizens
8. Openness—Self-direction	
Ipcrtiv	1. Thinking up new ideas and being creative is important to her. She likes to do things in her own original way
Impfree	11. It is important to her to make her own decisions about what she does. She likes to be free and not depend on others
9. Openness—Stimulation	
Impdiff	6. She likes surprises and is always looking for new things to do. She thinks it is important to do lots of different things in life
Ipadvnt	15. She looks for adventures and likes to take risks. She wants to have an exciting life
10. Openness—Hedonism	
Ipgdtim	10. Having a good time is important to her. She likes to “spoil” herself
Impfun	21. She seeks every chance she can to have fun. It is important to her to do things that give her pleasure

transcendence” “Conservation” “Openness” である。

つぎに、後者の“Achievement”については、それは、Schwartz の 10 の「基本的な人間の価値観 (basic human values)」の “Achievement” “Power” “Benevolence” “Universalism” “Conformity” “Tradition” “Security” “Self-direction” “Simulation” “Hedonism” の 1 つであることを意味している。

なお、ここでの解説で重要なポイントは、それを「原語 (英語)」で示し、「日本語訳」をしていないというところにある。それは、そのような「訳語」を確定すること自体が 1 つのプロジェクトとなるほどの大きな「エネルギーと時間」を要するものであると考えるからにはかならない。

筆者 (真鍋) は、2014 年から数年をかけて、ドイツ・ケルン大学の Wolfgang Jagodzinski 教授、Eldad Davidov 教授、Herman Dülmer 講師、上智大学の Carola Hommerich 准教授との共同研究の形で PVQ-RR 57 items (10/2013) の「日本語訳」を確定した (真鍋, 2018) が、このような質問諸項目のそれぞれが対応する「価値観の諸タイプとそれらのより高次の 4 つの価値観」を表現する「原語」の日本語訳は、それに劣らず「エネルギーと時間」を要するものとなるであろうことは明らかである。

例えば、前者の「価値観のタイプ」の 1 つとして、Schwartz は benevolence をあげている。この用語は、単に辞書的な意味を手がかりとするならば、「慈悲」という訳語が当てられるかもしれない。しかし、では、「慈悲」とは、どのような意味内容を持つ用語なのであろうか。この点について、「インド哲学・思想・宗教」研究の第一人者である中村元は、それが「仏教の伝統的な用語」の 1 つであるとして、その解説のために著書『温かなところ——東洋の思想——』(春秋社、1999 年) の 1 章 (p.4-45) を割いている。そして、その結論として、それは「温かなところ」ということになるであろうとしている。

じつは、このような論考は、「漢字で書いただけでは見えてこない言葉の本質が、やまと言葉で明らかになってくる」(中村桂子「解説」長谷川權『和の思想』岩波書店、2022 年、p.221) とい

う示唆にもつながる、「社会科学」にとってもきわめて重要な問題提起であるといわなければならないのである。

③「4 つのより高次の価値観の 1 つ: Self-enhancement」とその下位にある「10 の基本的な人間の価値観の 1 つ: Achievement」に続いて、それに対応する——それを質問文によって実証的に測定する——ESS の PVQ-21 items が示される。

では、PVQ と呼ばれる質問文の形式がどのようなものであるかを、ここでの 4 と 13 の 2 つの質問文を用いて具体的に説明しておきたい。

因みに、ここで 4 と 13 という番号が何を意味しているかということ、それは、これら 2 つの質問文が ESS の質問紙において、何番目に位置しているかを示しているものであって、それ以上の意味を持つものではない。

さて、PVQ は、質問紙法にもとづく社会調査において、「間接的方法」として分類される形式である。例えば、問 4 の場合でいえば、被調査者 (女性) に対して、「彼女にとっては自分の能力を発揮することが重要である。彼女は自分のしたことを人びとが賞賛することを望んでいる」というステートメントが示される。そして、その被調査者 (女性) は、このようなステートメントに示された女性が、自分と似ているかどうかを判断し、その程度を、1 (「まったく似ていない」) から 6 (「とてもよく似ている」) までの 6 点尺度で評定する、というのがその方法である。このように、被調査者の価値観がどのようなものであるかの判断が、それぞれのステートメントに描かれた人が、自分とどのくらい似ているかの評定 (rating) をとおしてなされることから、このような方法が「間接的方法」と呼ばれるのである。

そして、この問 4 の質問文の前に、Ipshabt という耳慣れない単語が記されているが、これは、この質問文のラベルともいえるべきものであって、この質問文を構成する 8 つの単語から任意のアルファベットを取り出して、それらを組み合わせて作成した人工合成語である。

以上で、測定不変性のテストの手順についての「決定木のアイデア」を用いた、Schwartz の

「価値観理論」の枠組みにもとづく国際比較／文化比較調査の、データ分析の諸事例（諸文献）の分類・整理・タイプ分けのための準備が整った。

3. ESS の価値観諸項目の測定不変性のテストに関する先行研究のタイプと内容

以下においては、「決定木のアイディア」にもとづく、先行研究の分類・整理・タイプ分けの結果を示すとともに、そのように分類・整理・タイプ分けされた諸研究の内容を簡単にレビューすることを試みる。すでに述べたように、以下の小節の番号は、「決定木の概観図」の楕円内に示された番号に対応する形となっている。

3.2.1 測定不変性のテスト

われわれの共著論文（真鍋・前田・清水、2021）において解説したように、測定不変性は、「多集団確証的因子分析（MGCF）」という統計的技法の開発をとおして、実証的にテストすることが可能となった。そのようなテストは、一般に、“configural” “metric” “scalar” という段階的な3つのレベルのモデルを用いて行われる。「より低いレベルのモデル」は「より低いレベルの不変性」に、そして「より高いレベルのモデル」は「より高いレベルの不変性」に対応している。このような測定不変性のテストが、ここでの先行諸研究の分類・整理・タイプ分けの出発点となる。

国際比較／文化比較の視座からする価値観調査のデータ分析という研究領域において、「測定不変性のテスト」という課題に初めて挑んだ画期的な論文が Davidov et al. (2008) であった。この論文では、Schwartz の21 の価値観項目について、ESS の第1回（2002-3）の調査データを用いて、ヨーロッパの20か国の測定不変性のテストが試みられた。同様の試みが、第2回調査データ（2006-7; 25か国）、第3回調査データ（2006-7; 25か国）についても行なわれた（Davidov, 2008, 2010）。これらのデータ分析の試みの結果は、以下の3点にまとめることができるであろう。

(1) Schwartz の「価値観理論」からするならば、「基本的な人間の価値観」は10の価値観で構成されるものとされた。しかし、「確証的因子分

析（CFA）」の結果では、これら10の価値観の7つが区別されるにとどまった。それらは、以下のものである。

- ①security
- ②self-direction
- ③stimulation
- ④hedonism
- ⑤tradition/conformity
- ⑥universalism/benevolence
- ⑦power/achievement

以上の結果について、ここで注意しておかなければならない点は、⑤⑥⑦の価値観については、それぞれの対（pairs）の相互の相関がきわめて高く（very high correlations）、そこで、これら3組の対については、それぞれ1つずつの価値観の構成概念（single value construct）として統合（unified）／合成（combined）されているということである。

(2) 以上の7つの価値観については、MGCF の結果では、metric invariance のレベルまでが立証された（established）が、scalar invariance はデータからは支持されるに到らなかった。このような結果は、ESS で測定される価値観の、国ぐに間の「平均値（mean）」による比較には、方法論的に「問題がある（problematic）」ということを示唆している。

(3) Davidov (2008, 2010) は、それにもかかわらず、これら7つの価値観については、経時的（over time）な scalar invariance が立証されることを確認した。この結果は、調査対象国ごとの価値観の時系列的な比較（longitudinal comparison）のために、それらの国ぐにの「平均値（means）」を利用することが可能であるということを示唆している。

以上から、Davidov らの画期的な「応用研究」は、人びとの「価値観」の国際比較／文化比較という試みが、測定不変性のテストという統計的技法を導入して検討してみるならば、それが必ずしも容易な試みではないことを明らかにしたのである。

では、そのような測定の不変性のテストの結果が、期待したものとならなかった場合——いかにすれば、それが「非不変性 (non-invariance)」という結果になった場合、——応用研究に携わる研究者が取るべきつぎの手続き／判断はどのようなものとなるのであろうか。図1の「決定木図」からするならば、その場合——つまり 3.2.1 の文献で、その結果が NO であった場合——の選択肢としては、つぎの2つの方途が示されている。

① “partial exact MI” “full approximate MI” “partial approximate MI” のテスト

②測定の道具 (measurement instrument : 具体的ないうならば、調査の質問項目) と理論的なモデル (theoretical model : 具体的にいうならば、Schwartz の価値観モデル) の再検討

これら2つの方途については、それぞれ、以下の点を注記しておかなければならない。

まず、①については、MI、より具体的には、MGCFA のテストにおける “full” と “partial” の区別、そして、“exact” と “approximate” の区別、についての理解が必要となる。この点については、真鍋・前田・清水 (2021) を参照されたい。

つぎに、②については、Schwartz の価値観モデル、より具体的には「円環連続体モデル／ヒエラルヒカル構造モデル (Circular Continuum Model/Hierarchical Structure Model)」についての理解が必要となる。この点については、真鍋・前田・清水 (2022 b) を参照されたい。

こうして、①から始めて、つぎに②の選択肢に進み、そして、それらの再検討の結果、「質問項目」あるいは「理論モデル」が **NOT OK** であるということがわかった場合、「質問項目」と「理論モデル」の両方またはその一方をどのように「改良 (improve)」するかがつぎの課題となる。

3.2.2 「理論モデル」と「質問項目」の両方またはその一方の改良

Schwartz (1992) において構築された「価値観モデル」は、Schwartz et al. (2012) において、さらに新しい展開を示すことになる。そこでは、もとのモデルの基本的なアイディアは堅持しながら、そのさらなる洗練化 (refining) が企てられ

たが、その「洗練化」はつぎの2つの方向で試みられた。

①Schwartz (1992) では、図2に示したように、「円環連続体」を構成する価値観が、10の種類／タイプに分けられた。ところが、Schwartz et al. (2012) では、図3に示したように、それら10の価値観の種類／タイプの多くについて、さらなる細分化——例えば、「Security」を「Personal Security」と「Social Security」に分けるといような細分化——を行なうとともに、新しく「Humility」と「Face」の2つの種類／タイプを追加することによって、19の精緻化された価値観に分けられた。

②ヒエラルヒカルな構造という点については、より高次の価値観の因子／次元を示す2番目の同心円に加えて、その外側にさらに3番目と4番目の同心円が描かれることになった。3番目が Social Focus と Personal Focus のコントラストを示す同心円であるのに対して、4番目は Growth/Anxiety-Free と Self-Protection / Anxiety-Avoidance とのコントラストを示す同心円である。こうして、「19の精緻化された価値観」「10の基本的な人間の価値観」「3つの段階からなるより高次の価値観」という「ヒエラルヒカルな三層構造モデル」が確立されることになった。

では、なぜこのような「理論モデル」と「質問

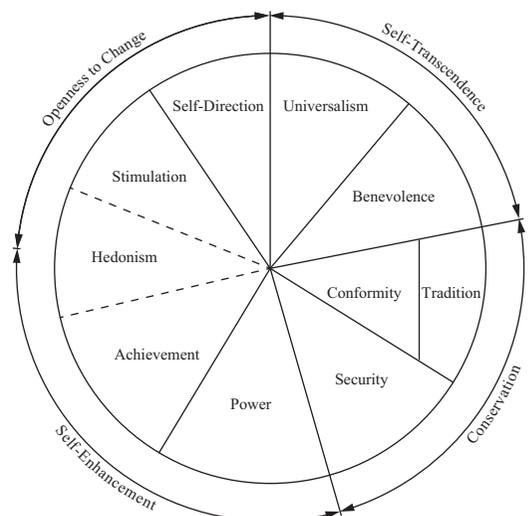


図2 Schwartz (1992) の価値観の「円環連続体モデル／ヒエラルヒカル構造モデル」
(出典：Sagiv and Schwartz, 1995)



図3 Schwartz et al. (2012) の価値観の「円環連続体モデル/ヒエラルヒカル構造モデル」
(出典：Schwartz et al., 2012)

項目」の改良 (improvement) が試みられたかという点、それは、Davidov et al. (2008)、Davidov (2008)、Davidov (2010) によって、Schwartz (1992) の「価値観モデル」と「価値観諸項目」が、国際／文化の視座からする比較可能性 (comparability) という点において問題を残しているという指摘がなされたからであり、そのような問題点を方法論的に解決することをめざして、このような改良の方策が提案されたのである。

こうして、Schwartz et al. (2012) よって、新しく再構成された——「作り直された」(Merton, 1957=1961) ——「価値観モデル」と「価値観諸項目」を用いて、それらの測定不変性を国際比較／文化比較のコンテキストにおいてテストする諸研究が行なわれてきた (Cieciuch, 2013, 2014 a, 2014 b)。

ここで、これらの諸文献の内容をごく簡潔にまとめておくならば、それはつぎのようになる。
3.2.1 のところで紹介した ESS の PVQ-21 の場合とくらべて、これらの諸文献でデータ分析に用いられた PVQ-40、PVQ-5 X (PVQ の第 5 番目の実験版の質問紙) の場合は、測定不変性が実質的によりよいものとなった。とくに後者の PVQ-5 X の場合は、すべての価値観をそれぞれ個別の

(distinct) single value として区別する (distinguish) ことが可能となるとともに、19 の精緻化された価値観のすべてについて、ほぼすべての国々において、configural invariance と metric invariance が確認された。そして 10 の基本的な人間の価値観のすべてについて、ほぼすべての国々において scalar invariance が支持される結果となった。

3.2.3. 測定の非不変性 (measurement non-invariance) を文化比較における測定の差異 (differences) の情報源として利用する方法

諸集団——さまざまな国、あるいは、さまざまな時点の調査対象のような諸集団——について、統計的に意味のある比較 (statistically meaningful comparison) を行なうためには、それに先立って、測定不変性のテストをしておくことが不可欠の前提条件 (prerequisite) となる。ところが、こうして試みられる測定不変性のテストにおいて、そのような不変性がデータで確認できないことがある。このような場合においては、諸集団間の比較には問題があり、このような比較には「偏りがある (biased)」ということになる。

社会調査における、以上のような「質問項目のもたらす偏り (item biased)」という問題を説明するために、Davidov et al. (2012) は、「マルチレベル構造方程式モデリング (multilevel structural equation modeling: MLSEM)」を応用することを提案した。そして、そのような応用研究の方法を具体的に例証するために、ESS の第 2 回調査データを利用した。

このような研究事例 (文献) についてレビューしようとするならば、それには 2 つの仕方があるであろう。1 つは、それを technical に解説するという仕方であり、もう 1 つは substantive に解説するという仕方である。前者の仕方をとるとするならば、その解説は少なくとも本稿の 1 節分の紙幅を要することになるであろう。そこで、それは、別の機会に譲ることとし、本稿では、後者の仕方をとることとし、その解説もごく簡潔に、「それによって何が明らかとなったか」をとりまとめるにとどめる。

ここで、Davidov et al. (2012) が焦点を当てた

のは、Schwartz の「基本的な人間の価値観」の1つである universalism であった。これは、「環境を保護することの重要性」を捉える項目であるが、ESS の第2回調査データにおいては、この項目はヨーロッパの国ぐににおいて、おしなべて非不変であった。では、このような非不変性は、どのように説明されるであろうか。

Davidov et al. (2012) は、ヨーロッパの国ぐにを、「人間開発指数 (Human Development Index)」のスコアの「高い国ぐに」と「低い国ぐに」に分けた場合、この universalism の価値観は、前者の国ぐにの人びとにくらべて、後者の国ぐにの人びとによってより強く支持される (endorsed) ということを見出した。では、なぜ、このような傾向が見られるのであろうか。HDI スコアの低い国ぐにの人びとは、きれいな水と空気は、まさに自分たちの生存と健康の問題であると考え。それに対して、スコアの高い国ぐに——つまり、より発展した国ぐに——の人びとは、きれいな水と空気は、もはや当たり前のことと考える。そして、このような考え方の違いから、universalism という価値観については、国ぐにに「非不変性」が観察される結果となる、というのが Davidov et al. (2012) の説明である。

3.2.4 近似的な測定不変性のテストの諸技法 (approximate measurement invariance test- ing techniques)

国際比較調査／文化比較調査における比較可能性のテストに MGCFA が利用されるようになるにともなって、そのような統計的なテストにおいて、高いレベル——例えば、scalar レベル——の測定不変性が確認されることは、きわめて稀であることがわかってきた。

では、このような問題——測定非不変性という問題——に対しては、どのように対処すればいいのであろうか。ここで、再び、さまざまな「統計的な解決策 (statistical solution)」が提案されることになる。それらが 3.2.4 に示した諸技法である。Cieciuch et al. (2019) は、Muthén and Asparouhov (2013) と van de Schoot et al. (2013) にもとづく「ベイジアン (Bayesian) の近似的 MI アプローチ (Approximate MI approach)」を取りあげる。

しかし、このアプローチの基本的な考え方とその技法のテクニカルな側面については、真鍋・前田・清水 (2021) の IV. 測定の不変性／不変性の検討法の新しい展開——Exact MI から Approximate MI へ——において、執筆者の前田がかなり詳細に解説した。そこで、本稿においては、Cieciuch et al. (2019) に即して、そのアプローチを応用することによって「何が明らかとなったか」に焦点を合わせて、以下の研究諸事例 (文献) の内容を概観するとどめる。

まず、Zercher et al. (2015) は、ESS の PVQ-21 で捉えられた人びとの価値観について、「ベイジアン近似的 MI アプローチ」でデータ分析を試みた初めての研究事例である。ここでは、Schwartz の 10 の価値観の1つ、universalism の価値観に焦点が合わされた。そして、ESS の第1回～第6回のすべての調査に参加した 15 か国を対象に「近似的な測定不変性のテスト」が行われた。それは、15 か国×6回調査=90 グループについてのテストとなった。このようなテストの結果は、つぎのようにまとめられる。

これまでの伝統的／厳密なアプローチ (traditional/exact approach) では、“scalar measurement invariance” は、これら 90 グループのなかの 37 グループにおいてのみ確立された (established) のに対して、近似的／ベイジアンのアプローチ (approximate/Bayesian approach) では、「近似的な不変性」は、90 グループのなかの 73 グループにおいて確立された (established)。つまり、universalism という価値観は、国と時点の組み合わせのすべてのケースにおいて、測定不変であるわけではないものの、それらの組み合わせのほとんどのケースにおいて比較可能であることがわかったのである。

こうして、「国際比較研究において測定不変性がデータによって支持されることは稀である」というこれまでの諸知見 (findings on lack of invariance) に対しては、実証的に異議が唱えられることになったのである。

以上の Zercher et al. (2015) の研究を、さらに一歩進めたのが Cieciuch et al. (2018) の研究事例 (文献) であった。その研究事例では、以下のような分析の仕方が採用された。

・Zercher et al. (2015) では、Schwartz の 10 の価値観の 1 つ——universalism——に焦点が合わされたのに対して、この研究では、価値観諸項目のすべて——ただし、hedonism は除く——について近似的な測定不変性のテストを試みる。

・その場合には、モデルがきわめて複雑なものとなるために、15 か国×6 回調査=90 グループについての同時テストに替わって、15 か国について、それぞれの調査回ごとに近似的な測定不変性のテストを試みる。

・価値観諸項目をそれぞれ別々に扱うのではなく、Cieciuch and Schwartz (2012) において提案された「拡大鏡 (虫めがね) 方略 (magnifying glass strategy)」——真鍋・前田・清水 (2022) を参照されたい——のアイデアにもとづいて、Schwartz の「高次の価値観 (higher-order value)」のそれぞれについて、別々に、近似的な測定不変性のテストを行なう。

・この「拡大鏡方略」をとる場合、hedonism の価値観を、隣接するいずれかの「高次の価値観」に含めるとするならば、それらの「モデル適合度 (model fit)」が有意に低下することから、その価値観はここでの分析モデルからは除外することにする。

以上のようなテストの結果、15 か国のほとんどにおいて、「高次の価値観」のそれぞれについて、近似的な測定不変性が確認されることがわかった。ただ、このような近似的な測定不変性のテストにおいて、どのくらいまで国ごとの測定のパラメーターの variability (ばらつき/散らばり) が許容されるかは、未だ十分に検討されていない。今後の重要な課題といわなければならない。

3.2.5 Alignment Optimization の応用

Cieciuch et al. (2019) が、測定不変性のテストの実践的な手続きと判断の「決定木図」の最後のステップに位置づけたのが、“Alignment Optimization” と呼ばれる技法である。この技法は、Asparouhov and Muthén (2014) で提案されたものであり、それは、Mplus 7.11 (Muthén and Muthén, 1998-2014) で実装 (implemented) されている。

では、なぜ、この技法は、そのプロセスの最後のステップに位置づけられたのであろうか。それ

は、Cieciuch et al. (2019) においてばかりでなく、Davidovei et al. (2014)、van de Schoot et al. (2015)、Meitinger et al. (2020) など、測定不変性の研究動向をめぐるレビュー諸論文においても、この技法が、最も新しく登場した技法であるばかりでなく、この研究領域の現在の段階における最も有望な技法である、とされてきたからにはかならない。

しかし、本稿では、この技法が、とくに国際比較/文化比較をめざす応用研究の便宜性を考慮に入れて開発されたものであるという点に注目したい。では、「国際比較/文化比較をめざす応用研究にとっての便宜性」とは、具体的に、どのようなことであろうか。じつは、国際比較/文化比較の視座からなされる応用研究の多くが最もよく利用してきた統計的な道具 (statistical tools) の 1 つが「平均値」であり、そのような「平均値」を用いた国際比較/文化比較に、いわば「方法論的な御墨付き」ともいうべきものを確保するために、応用研究者たちは測定不変性のテストにおいて、その最も高いレベルである scalar invariance の立証に努めてきたという背景がある。こうして、「国際比較/文化比較の視座に立つ応用研究にとっての便宜性」とは、「最も信頼に値する平均値の推定 (estimation of trustworthy means) を可能にする統計的技法の利用」ということになる。そして、まさにこのような便宜性に応えるべく開発されたのが Alignment Optimization にはかならない。

この点は、「応用研究」の創造性/可能性の探究というところに「問題関心」の射程を定める本稿において、きわめて重要な「いとぐち」となるはずのものである。しかし、そのような「いとぐち」は、本稿では、どこまでも筆者らの文献研究にもとづく確認にすぎない。ここでは、このような、社会科学の領域における「知の創造」の淵源を探る実証的な試みにおける方法論的な問題の所在の 1 つを記しておくにとどめる。

以下においては、Cieciuch et al. (2019) によりながら、Schwartz の「価値観理論」の枠組みにもとづく、人びとの「価値観」の国際比較研究/文化比較研究に、Alignment Optimization の技法がどのように利用されてきたかを概観していく。

(1) Cieciuch et al. は、表1に示された人びとの「価値観諸項目 (PVQ-21)」と、図4に示された人びとの「4つの高次の価値観 (higher-order values) の測定モデル (measurement model)」にもとづいて、ESSの第1~7回に参加した15か国を対象に、Alignment Optimizationの技法を用いて、測定不変性のテストを実施した。これは、この研究領域において知られている最大規模の応用研究ということができる。なお、Alignment Optimizationは、MplusソフトウェアパッケージVersion 7.3 (Muthén and Muthén, 1998-2014)を用いて実行された。Mplusによるテストの実行は、測定の不変性のほかのテストの場合とくらべて、とくに比較する国/文化の数が多い場合には、はるかに容易な手続きとなる。こうして、Alignment Optimizationは、substantiveな研究に取り組む応用研究者にとっては、きわめて有用な技法であるといえるのである。

(2) 表2は、105——分析対象に取りあげた15か国×ESSの第1~7回——のグループにおける、4つの「高次の価値観」のそれぞれについての、configural invarianceの「全体的なモデル適合尺度 (global fit measures)」を示したものである。この結果から、4つの「高次の価値観の測定モデル」——Conservation、Self-transcendence、Open-

ness to change、Self-enhancementの4つのモデル——は、データへの高い適合度 (good fit) を示していることがわかる。

これら4つの「高次の価値観」の因子負荷量 (factor loadings) の1995—— (Conservationについての6つの因子負荷量+Self-transcendenceについての5つの因子負荷量+Openness to changeについての4つの因子負荷量+Self-enhancementについての4つの因子負荷量=19の因子負荷量) × (15か国×ESSの第1~7回=105のグループ) = 1995——のすべてが、統計的に有意 (significant) であり、それら因子負荷量の値は、そのほぼすべてが0.4よりも大きいものであった (この点については、Brown, 2015を参照されたい)。そして、因子負荷量が0.3よりもやや低い値となったものは、わずかに10ケース——すべての因子負荷量の数1995の内の10ケース=約0.5%——にとどまった。

(3) 表3は、Mplusのアウトプットにおいて、非不変 (non-invariant) であることが示された「因子負荷量」と「切片」の数とその%である。この情報は、Alignment Optimizationの手続きをとおして計算される「平均値 (means)」が、信頼に値するものであるかどうかを判断するために重要である。Muthén and Asparouhov (2014) は、

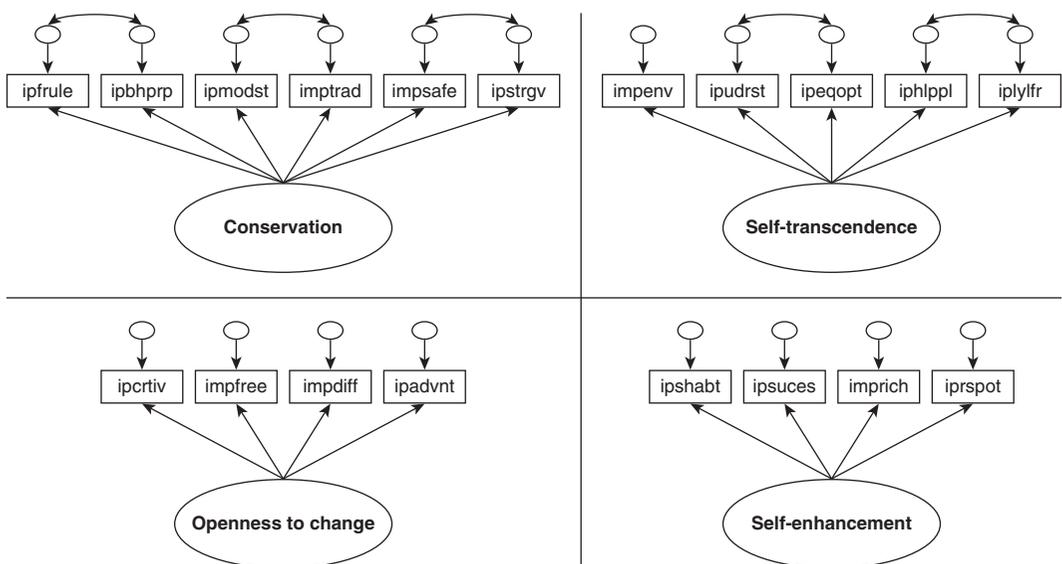


図4 4つの高次の価値観の測定モデル——拡大鏡 (虫めがね) モデル——

表2 すべての分析対象グループにおける高次の価値観についての **Configural Invariance** レベルのモデル適合度

	χ^2	df	CFI	RMSEA	SRMR
Self-enhancement	741.5	210	0.996	0.038 [0.035-0.041]	0.011
Self-transcendence	1907.0	315	0.989	0.053 [0.051-0.055]	0.015
Conservation	2758.2	630	0.988	0.043 [0.042-0.045]	0.017
Openness to change	4278.3	210	0.961	0.104 [0.101-0.107]	0.030

df degrees of freedom, CFI Comparative Fit Index, RMSEA Root Mean Square Error of Approximations; SRMR Standardized Root mean Square Residuals

表3 **Alignment Optimization** の手続きをとおして非不変であることが示された「因子負荷量」と「切片」の数とその%

	Loadings, %	Intercepts, %	Average, %
Conservation	12 (73/630)	47 (294/630)	29 (367/1260)
Self-transcendence	12 (61/525)	45 (234/525)	28 (295/1050)
Openness (without hedonism)	14 (60/420)	54 (228/420)	34 (289/840)
Self-enhancement	6 (26/420)	59 (247/420)	33 (273/840)

非不変であるパラメーターの数——つまり、「因子負荷量」と「切片」の数——が、全体の約25%という境界基準 (cutoff criterion) を越えてはならないとしている。非不変のパラメーターの数を、平均の%で見えていくなれば、それは Self-transcendence で最も低く (28%)、つぎが Conservation (29%)、続いて Self-enhancement (33%) で、Openness to change で最も高く (34%) なっている。これらの結果は、いずれも Muthén and Asparouhov (2014) によって提示された約25%という境界基準をやや越えるものとなっている。

つぎに、そのような非不変性を、「因子負荷量」の場合と「切片」の場合に分けて見ていくなれば、まず、「因子負荷量」の場合は、4つの高次の価値観のいずれについても、非不変の%は基準内のものとなっている。ところが、「切片」の場合は、4つの高次の価値観のいずれについても、非不変の%は50%ほどのものとなり、約25%という基準を大きく越えるものとなっていることがわかる。

このような知見は、これまでの Alignment Optimization 以外の測定不変性のテストの技法をとおして導かれた ESS のデータ分析の結果と、完全に対応する (corresponds) ものとなっていることがわかる。それは、ESS の価値観諸項目につ

いては、多くの場合、その測定不変性のテストの結果、「因子負荷量」は不変 (invariant) となるが、「切片」は非不変 (non-invariant) となるという知見である。

(4) 以上を踏まえて、Alignment Optimization の手続きをとおして、当該データをもとに、「最も信頼できる平均値 (most trustworthy means)」が計算されることになる。しかし、このような平均値にもとづく諸知見 (the findings for the means) は、とくに非不変の「切片」の数が50%を越える “Openness to change” (54%) と “Self-enhancement” (59%) の場合は、注意深く取り扱わなければならない。

そしてさらに、Mplus による Alignment Optimization のアウトプットにおいては、このような「平均値」にもとづいて、「調査対象国」と「ESSの第1~7回」の組み合わせのグループごとの「ランキング」が提示される。

こうして、以上のような形で示される当該データの分析結果は、特定の国や時点における人びとの価値観という substantive なテーマに関心を持つ「応用研究者」にとっては、きわめて有用な情報となるのである。

Ⅲ. おわりに——まとめと今後の課題——

本稿の目標は、「国際比較／文化比較のための MGCFA の応用研究の創造性／可能性の探究」ということであった。ここでは、応用研究という用語を、「統計数理を専門とする研究者によって開発されてきた統計的技法を社会科学の研究領域にどのように応用するかに関する研究」として位置づけてきた。そして、このような問題関心と方法論的立場から、筆者らは、この領域——つまり、MGCFA と呼ばれる「統計的技法」をめぐる研究領域——における応用研究がつぎのような創造性／可能性を示すものであることを明らかにしてきた（真鍋、前田、清水、2021、2022 a、2022 b）。

- (1) さまざまな substantive なテーマめぐる諸研究の「方法論的な質の向上（improve the methodological quality）」（Steenkamp and Baumgartner, 1998）を促進する。
- (2) 社会科学の領域における新しい「理論の構築（theory building）」を可能にする。
- (3) そのようにして構築された「理論の作り直し」（Merton, 1957=1961）を提案する。

しかし、このような応用研究の創造性／可能性は、以上のような3つの側面にかざられるわけではない。本稿では、その新しい創造性／可能性の側面に、比喩的にいうならば、「サーチライトの光」（高根、1979）を当てることにした。ここで「新しい」と書いたが、それは、そのような試みが、これまで「なされてこなかった」ということでは決してない。筆者らは、すでにして、そのような試みの具体的な例として、Greenacre の“practice”と林らの「実際」という用語で意味される試みに注目した。そして、本稿で取りあげた Ciecuch et al. (2019) の「測定不変性のテストをめぐる手続きと判断の流れ図」と「価値観の国際比較をめぐる Mplus にもとづく Alignment Optimization のイラストレーション」は、まさに Greenacre や林らの試みと軌を一にするアイデアにもとづくものであると考えた。

では、このようなアイデアに示された「知」

は、どのような「知」なのであろうか。社会科学には、人間の「知」の社会的在庫を研究対象とする「知識社会学」と呼ばれる領域がある。例えば、Berger and Luckman (1966=1977) は、このような人間の「知」のなかで、とくに「処理的な知、つまりルーティーンの遂行における実用的な能力にかかわる知」に注目する。それは、人びとの日常生活においては、このような「知」が、まさに「突出した地位を占めている」（訳書、p.72）からにはほかならない。しかし、このような「知」は、社会科学の領域における「応用研究」においても、きわめて重要なものといわなければならないのではなかろうか。じつは、そのことをあらためて確認するという意味において、ここで「新しい」と書いたのである。

こうして、本稿における筆者らの方法論的な立場は、明らかになったであろう。筆者らは、社会科学の領域の応用研究においても、そのような実践的・処理的・実用的な手続きをめぐる「知」の累積が、その重要な部分を形作ってきていると考えるのである。

以上に述べてきたような、筆者らの問題関心と方法論的な立場を踏まえて、本稿でのこれまでの叙述をまとめておこなうならば、それはつぎようになる。まず、本稿で取りあげた Ciecuch et al. (2019) の文献の内容の再確認から始める。

(1) 測定不変性のテストの実践プロセスにおいて、測定不変性がデータで確認されない場合——いいかえれば、そのテストの結果が「測定非不変（measurement non-invariance）」となった場合——その後、どのような分析の手続きをとるかについて、「決定木」の形で判断と選択を示した流れ図を提案するとともに、この「決定木の流れ図」を、この領域における先行諸研究のデータ分析の試みのタイプ分けのための枠組みとして利用することによって、そのような諸研究が Exact MI Test→Partial MI Test→Approximate (Bayesian) MI Test→Alignment Optimization というように段階的に発展してきた歴史的なプロセスを見事に描き出した。

(2) この「決定木の流れ図」において、方法論的な発展の最後の段階——つまり、現時点におけるこの研究領域の到達点——に位置づけられる Alignment Optimization と呼ばれる技法の、Mplus ソフトウェアパッケージ (Muthén and Muthén, 1982-2014) にもとづく実行プロセスの概要を解説した。

では、つぎに、本稿では、以上のような文献の内容について、「社会科学の領域における統計的技法の応用研究の創造性／可能性の探究」という視座から、どのように評価できるかについてまとめておこう。

まず、(1) についての筆者らの評価のポイントは、以下のようなものであった。測定不変性のテストの実践プロセスについての「決定木の流れ図」は、図1のキャプションの後に () 付で示された Author's own work という部分に端的に表現されているように、そのチャート化の試みそのものが、筆者らの独自の創造的なアイデアである。

では、それは、なぜ、そのように高く評価されるのであろうか。このような問いに対しては、つぎのように答えることができるであろう。いうまでもなく、「決定木」というアイデアは、ここで初めて考案されたものではない。それは、「意思決定問題での選択肢のグラフ表現」(Upton and Cook, 2008=2010, p.120) のアイデアを、「測定不変性のテストの実践プロセス」に当てはめた、「応用研究」の1つの成果ともいうべきものである。そして、そのような「応用」の意義がどこにあるかという点、それは「測定不変性のテスト」に携わる研究者が、そのようなテストをそれぞれのデータにもとづいて実践する場合に、参照すべき「青写真／ロードマップ」としてきわめて有効なものであるというところにある。

こうして、「決定木の流れ図」のアイデアは、まさに Berger and Luckman (1966=1977) のいうところの「処理的な知／ルーティーンの遂行における実用的な能力にかかわる知」として、高く評価されるべきものといわなければならない。

そしてさらに、Cieciuch et al. が、この「決定

木の流れ図」を、この領域における先行研究の方法論的な視座からのタイプ分けのための枠組みとして利用した試みも、同じく卓越したアイデアとして高く評価されるものといえよう。

最後に、以上のような「先行諸研究のタイプ」とそれぞれの「先行諸研究の内容」については、両者の対応関係——つまり、どのような「タイプ」の研究で、どのような「テーマ・内容・質問諸項目」が研究されているか——について探っていくことが、興味深い今後の課題となるであろう。

つぎに、(2) については、そこでは、Mplus にもとづく Alignment Optimization の試みが、社会科学の応用研究の領域においてきわめてプロミシングなものであることが示された。しかし、ここで、示された試みを参照するだけで、「Mplus ソフトウェアパッケージを用いた Alignment Optimization の実行の再現化」が可能になるわけではない。そのためには、より実践的な「Mplus による Alignment Optimization の実行プロセスのイラストレーション」が必要になる。林知己夫らの表現——林知己夫編著『尺度元尺度解析法の実際』サイエンス社、1984年——をもじっていうならば、それを待って、初めて「Alignment Optimization の実際」への探索がスタートすることになるのである。

(本稿は、3人による共同研究の成果を真鍋が取りまとめたものである。)

文献

- Asparouhov, T., and Muthén, B. O. (2014). Multi-Group Factor Analysis Alignment. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(4), 495-508.
- Berger, P. L., and Luckmann, T. (1966). *The Social Construction of Reality*. Doubleday and Company (= 1977, 山口節郎訳『日常世界の構成』新曜社).
- Cieciuch, J., Schwartz, S. H., and Vecchione, M. (2013). Applying the Refined Values Theory to Past Data: What Can Researchers Gain? *Journal of Cross-Cultural Psychology*, 44, 1215-1234.
- Cieciuch, J., Davidov, E., Vecchione, M., Beierlein, C., and Schwartz, S. H. (2014a). The Cross-National Invariance Properties of a New Scale to Measure 19 Basic

- Human Values. A Test Across Eight Countries. *Journal of Cross-Cultural Psychology*, 45, 764-779.
- Cieciuch, J., Davidov, E., Vecchione, M., and Schwartz, S. H. (2014b). A Hierarchical Structure of Basic Human Values in a Third-Order Confirmatory Factor Analysis. *Swiss Journal of Psychology*, 73, 177-182.
- Cieciuch, J., Davidov, E., Algesheimer, R., and Schmidt, P. (2018). Testing for Approximate Measurement Invariance of Human Values in the European Social Survey. *Sociological Methods and Research*, 47(4), 665-686.
- Cieciuch, J., Davidov, E., Schmidt, P., and Algesheimer, R. (2019). How to Obtain Comparable Measures for Cross-National Comparisons. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 71, 157-186.
- Datler, G., Jagodzinski, W., and Schmidt, P. (2013). Two Theories on the Test Bench: Internal and External Validity of the Theories of Ronald Inglehart and Shalom Schwartz. *Social Science Research*, 42, 906-925.
- Davidov, E. (2008). A Cross-Country and Cross-Time Comparison of the Human Values Measurements with the Second Round of the European Social Survey. *Survey Research Methods*, 2, 33-46.
- Davidov, E., Schmidt, P., and Schwartz, S. H. (2008). Bringing Values Back in: The Adequacy of the European Social Survey to Measure Values in 20 Countries. *Public Opinion Quarterly*, 72, 420-445.
- Davidov, E. (2010). Testing for Comparability of Human Values Across Countries and Time with the Third Round of the European Social Survey. *International Journal of Comparative Sociology*, 51, 171-191.
- Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., and Billiet, J. (2014). Measurement Equivalence in Cross-National Research. *The Annual Review of Sociology*, 40, 3.1-3.21.
- Davidov, E., Dülmer, H., Schlueter, E., Schmidt, P., and Meuleman, B. (2012). Using a Multilevel Structural Equation Modeling Approach to Explain Cross-Cultural Measurement Noninvariance. *Journal of Cross-Cultural Psychology*, 43, 558-575.
- Davidov, E., Schmidt, P., and Billiet, J., and Meuleman, B., eds. (2018). *Cross-Cultural Analysis: Methods and Applications* [2nd Edition]. Routledge.
- Greenacre, M. (2017). *Correspondence Analysis in Practice* [Third Edition]. Taylor and Francis Group (= 2020, 藤本一男訳『対応分析の理論と実践』オーム社).
- 林知己夫編著 (1984). 『多次元尺度解析法の実際』サイエンス社.
- 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.
- 真鍋一史 (2018). 「〈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・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——多集団確認的因子分析 (MG CFA) を中心として——」『関西学院大学社会学部紀要』137, 1-28.
- 真鍋一史・前田忠彦・清水香基 (2022a). 「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——」『関西学院大学社会学部紀要』138, 1-36.
- 真鍋一史・前田忠彦・清水香基 (2022b). 「Schwartzの『価値観理論』の構築とその後の展開——『円環連続体モデル／ヒエラルヒカル構造モデル』に焦点を合わせて——」『関西学院大学社会学部紀要』139, 1-41.
- Meitinger, K., Davidov, E., Schmidt, P., and Braun, M. (2020). Measurement Invariance: Testing for It and Explaining Why It is Absent. *Survey Research Methods*, 14.4, 345-349.
- Merton, R. K. (1957). *Social Theory and Social Structure*. Free Press (= 1961, 森東吾ほか訳『社会理論と社会構造』みすず書房).
- Muthén, L. K., and Muthén, B. O. (1998-2014). *Mplus User's Guide*. Muthén and Muthén.
- Muthén, B. O., and Asparouhov, T. (2012). Bayesian Structural Equation Modeling: A More Flexible Representation of Substantive Theory. *Psychological Methods*, 17, 313-335.
- Muthén, B. O., and Asparouhov, T. (2017). Recent Methods for the Study of Measurement Invariance with Many Groups. Alignment and Random Effects. *Sociological Methods and Research*.
- 中村元 (1999). 『温かなところ——東洋の思想——』春秋社.
- 中村桂子 (2022). 「解説」(長谷川權『和の思想——日本人の創造力——』岩波書店).
- Schwartz, S. H. (1992). Universal in the Content and Structure of Values: Theory and Empirical Tests in 20 Countries. In M. Zanna ed., *Advance in Experimental Social Psychology*, 25, Academic Press.
- Schwartz, S. H. (2003). A Proposal for Measuring Value Orientations Across Nations. In *Questionnaire Devel-*

- opment Package of the European Social Survey, 259-319.
- Schwartz, S. H. et al. (2012). Refining the Theory of Basic Individual Values. *Journal of Personality and Social Psychology*, 103(4), 663-686.
- Schwartz, S. H., and Cieciuch, J. (2016). Values. In Frederick T. L. et al. eds., *The ITC International Handbook of Testing and Assessment*, 106-119. Oxford University Press.
- Steenkamp, J. E. M., and Baumgartner, H. (1998). Assessing Measurement Invariance in Cross-National Consumer Research. *Journal of Consumer Research*, 25, 78-90.
- Upton, G., and Cook, I. (2008). *A Dictionary of Statistics [Second Edition]*. Oxford University Press (= 2010, 白旗慎吾ほか訳『統計学辞典』共立出版).
- van de Schoot, R., Schmidt, P., de Beucklaer, A., Lek, K., and Zondervan-Zwijenburg, M. (2015). Editorial: Measurement Invariance. *Frontiers in Psychology*, 6, 1-4.
- Zercher, F., Schmidt, P., Cieciuch, J., and Davidov, E. (2015). The Comparability of the Universalism Value Over Time and Across Countries in the European Social Survey: Exact vs. Approximate Measurement Invariance. *Frontiers in Psychology*, 6, 1-11.

Practical Development of an Applied Study
on Multiple Group Confirmatory Factor Analysis
for Cross-National/ Cross-Cultural Comparison:
Survey Research Based on Schwartz's "Value Theory"

Kazufumi MANABE
Tadahiko MAEDA
Koki SHIMIZU

ABSTRACT

In a study on measurement equivalence/invariance in cross-national/cross-cultural research initiated as a collaborative study at the Institute of Statistical Mathematics, we collected a wide range of literature on applied studies of multiple group confirmatory factor analysis (MGCFA) published mainly in European and North American journals and examined them in detail.

Through this literature review, we found that applied studies in this research field have the following creativity and potential.

1. Promote the "improvement of the methodological quality" of research on diverse topics.
2. Enable "theory building" in the field of social sciences.
3. Propose "reworking" of the theories constructed.

However, the creativity and potential of applied studies is not limited to these three aspects. On the same lines of the above literature review, we review the study by Cieciuch, Davidov, Schmidt, and Algesheimer (2019) titled "How to Obtain Comparable Measures for Cross-National Comparisons" and published in *Kölner Zeitschrift für Soziologie und Sozialpsychologie*. This publication may be considered one of the best studies on the recent research trends of MGCFA. We examine its contents in detail. This study proposes a form of a "decision tree" whereby the procedures and decisions of measurement equivalence/invariance testing are applied step by step. The stages are as follows: Exact Measurement Invariance Test → Partial Measurement Invariance Test → Approximate Measurement Invariance Test → Alignment Optimization. Furthermore, the study presents a practical illustration of how Alignment Optimization, which is positioned at the last stage, is run using the Mplus Software Package. Therefore, it is understood that the development and accumulation of such practical knowledge is crucial in the applied study of statistical techniques.

Key Words: cross-national comparison/cross-cultural comparison, multiple group confirmatory factor analysis (MGCFA), applied study, Schwartz's value theory, decision tree, Alignment Optimization, Mplus