

# 心理尺度短縮版作成における IRT の活用に関する研究

並 川 努

## 目次

第1章 はじめに .....	1
1.1. 本研究の問題と目的 .....	2
1.2. 本論文の構成 .....	5
第2章 短縮版尺度作成の現状と課題 .....	7
2.1. 短縮版作成の方法について .....	8
2.2. 日本における短縮版の作成の現状 .....	9
2.2.1. 目的 .....	9
2.2.2. 方法 .....	10
2.2.3. 結果と考察 .....	11
2.3. まとめと今後の課題 .....	19
第3章 抑うつ尺度（DSRS-C）短縮版に関する検討 .....	21
3.1. はじめに .....	22
3.2. 問題と目的 .....	22
3.3. 方法 .....	25
3.4. 結果 .....	26
3.5. 考察 .....	31
第4章 Big Five 尺度短縮版に関する検討 .....	33
4.1. はじめに .....	34
4.2. 問題と目的 .....	34
4.3. 研究1 短縮版の作成 .....	36
4.3.1. 目的 .....	36
4.3.2. 方法 .....	36
4.3.3. 結果 .....	38
4.3.4. 考察 .....	42
4.4. 研究2 短縮版の妥当性検討 .....	43
4.4.1. 目的 .....	43
4.4.2. 方法 .....	43
4.4.3. 結果 .....	43
4.4.4. 考察 .....	44
4.5. 総合考察 .....	45

第 5 章 継時的比較志向性尺度の作成と短縮版に関する検討.....	46
5.1. はじめに.....	47
5.2. 研究 1 継時的比較志向性尺度の作成.....	47
5.2.1. 予備調査.....	51
5.2.2. 調査 1.....	52
5.2.3. 調査 2.....	55
5.2.4. 総合考察.....	57
5.3. 研究 2 継時的比較志向性尺度短縮版の作成.....	59
5.3.1. 問題.....	59
5.3.2. 調査 3.....	59
5.3.3. 調査 4.....	62
5.3.4. 総合考察.....	64
5.4. 本章のまとめ.....	64
第 6 章 短縮版作成方法による差異に関する検討.....	66
6.1. はじめに.....	67
6.2. 項目の選択方法に関する検討.....	67
6.2.1. 問題と目的.....	67
6.2.2. 方法.....	68
6.2.3. 結果および考察.....	69
6.3. 得点算出方法に関する検討.....	82
6.3.1. 問題と目的.....	82
6.3.2. 方法.....	82
6.3.3. 結果と考察.....	83
6.4. まとめ.....	85
第 7 章 総合考察.....	87
7.1. 本研究のまとめと今後の課題.....	88
引用文献.....	93
謝辞.....	101
付録.....	103

## 第1章 はじめに

## 1.1. 本研究の問題と目的

評定尺度 (Scale) は、心理学の実証的な研究や心理臨床の現場などにおいて多く用いられている。「パーソナリティ」や「精神的健康」、「QOL」、「ストレス」、「態度」、「感情」など、尺度によって測定が試みられる概念は多岐に渡る。また、狭義の「心理学」だけではなく、社会学や教育学、医学、看護学など、さまざまな領域の研究や臨床においても、今日その適用範囲は広がっている。しかし、それぞれのニーズにあわせて日々さまざまな種類の尺度が開発されているため、時に尺度の乱立 (榎本, 2006, 2013) が問題として指摘されることもある。実際、開発される尺度の数は多く、たとえば、日本語の心理尺度がまとめられた『心理測定尺度集』も I (堀, 2001) から VI (堀, 2011) まで 6 巻に渡って出版され、約 300 もの尺度を収録している。また、山口 (2011) では 2009 年 7 月から 2010 年 6 月の 1 年間に「パーソナリティ研究」、「心理学研究」、「教育心理学研究」の 3 誌で発表された 115 本の論文のうち、73 本の論文が心理尺度を利用している論文であったことが報告されているが、そのうちの 25 本は、新たな心理尺度を作成することを目的にした研究であったという。心理学の他の学会誌や紀要論文、学会発表、周辺領域の研究も含めれば、その数はさらに多くなることも容易に推測できる。

しかしながら、そのように新たに作成されている尺度も、必ずしもすべてが新規の概念を測定する尺度ばかりであるとは限らない。榎本 (2013) などでも指摘されているように、概念の名称は異なるものの類似した側面を測定していたり、既存の類似尺度との違いや関連が明確に論じられないままであったりする尺度が新たに開発される場合も多い。また、そもそも開発はなされたものの、その後ほとんど活用されることがないままの尺度も少なくない。同様の指摘は、他にもさまざまな形でなされているが (たとえば、内田, 2012 など)、このような現状は、研究成果の蓄積や発展を目指していく上では、決して望ましい状態ではない。そのため、既存尺度の整理や既存尺度の活用をさらに進めていくことは、今後の尺度を用いた研究において重要な方向性の一つとなりうるだろう。

では、既存の尺度があるにもかかわらず、それらが適切に活用されないことがあるとしたら、それはどのような理由によるものと考えられるだろう。まず考えられるのは、既存の尺度に関するレビューが不足しているために存在が見落とされていたり、既存の尺度が測定している概念についての理解が十分でないために概念間の類似性に気づかなかつたりする場合である。日々新たな尺度が開発され続けている一方で、各尺度が測定している概念の特徴等に関する情報は、必ずしも十分に提供されているわけではないことも影響していると考えられる。特に、既存の尺度との弁別という観点の議論は、あまり多くないように思われる。近年、尺度の妥当性検証の必要性があらためて議論されることも多く (村山, 2012 ; 内田, 2012 ; 吉田・石井・南風原, 2012), こういった点からも、今後さらなる検討が進められることが期待される。

また、既存の尺度に何らかの問題点があり、そのままでは利用しにくい場合も、既存尺度が活用されない例として考えられる。既存尺度の信頼性や妥当性が十分に担保されていない場合だけでなく、作成からの時間経過に伴い、項目の意味合いや回答者が項目に対して抱くであろう印象が変わっていることが予想される場合などがこれにあたる。これらの場合、問題点を修正した形で尺度が改訂されるか、もしくはそういった点を踏まえた形で新規尺度が提案されることになるだろう。

しかし、既存尺度自体に問題があるわけではなくとも、尺度を利用する際の目的や、時間的制約、回答者の特徴によって、既存尺度をそのままの形で利用することが難しい場合も存在するだろう。たとえば、複数の概念間の関連を検討することを目的とした調査では、当然複数の尺度を組み合わせる必要があるが出てくる。この場合、一つひとつの尺度自体はそれほど回答に要する時間は長くなくとも、複数の尺度を組み合わせる同時に実施する場合は、全体では回答に長時間を要し、回答者に過剰な負担をかけることになる場合もある。その場合、使用する尺度の項目数はできる限り少ない方が用いやすくなる。また、多少時間がかかったとしてもできる限り正確に測定できることを重視する場合もあれば、たとえある程度精度は低くとも、短時間で効率的に測定することを重視する場合もある。心理臨床や医療、福祉、教育などの「現場」においては、特にこのように短時間で施行可能なことが重要となる場合も存在するだろう。その場合も、項目数などは少ない方が望ましいと言える。さらに、回答者の属性によっても、制約が生じる場合も考えられる。たとえば、大学生や一般成人を対象にした調査に比べ、児童や高齢者等を対象にした調査では、集中力や体力的な問題で長時間質問紙に回答することが難しい状況が生じうることも予想される。このような場合も、既存尺度がそのままの形では使用しにくいケースの一つと言えるだろう。

心理尺度は、単一の質問項目によって構成概念を測定するのではなく、複数の項目への回答から構成概念を測定するように構成されているのが一般的である。そのため、たとえば多くの研究で用いられている Rosenberg (1965) による自尊感情尺度でも、自尊感情 (Self-esteem) という一つの概念を測定するために、10 個の質問項目が用意されている。しかし、上述のように調査の実施にかけられる時間や、回答者の負担等とのバランスを考慮すると、たとえ 10 項目であっても、「多い」と感じられる場合も少なくない。このように既存尺度が長すぎる（項目数が多すぎる）と判断された場合、それをそのまま用いるのではなく一部の項目のみを用いて、すなわち短縮版を構成して、それによって測定が行われることが多いだろう。実際、「オリジナル版」に加えて尺度の「短縮版」が開発されるケースは多く、前述の『心理測定尺度集』にも「短縮版」という名称のついた尺度が複数紹介されている（たとえば、平等主義的性役割態度スケール短縮版（鈴木, 1994）など）。また、10 項目の Rosenberg の自尊感情尺度であっても、調査実施上の制約などから 5 項目や 3 項目のみが抽出されて「短縮版」として利用されるようなケースが複数報告されてい

る（並川・脇田・野口，2010）。そのため，こういった短縮版の作成というニーズは，多くの尺度においても生じるものであると考えられる。

既存尺度を「短縮化」して利用することは，既存の尺度の持つ特徴を一定程度活かしつつ，回答にかかる時間を減らすことが可能になるため，研究の効率化や調査協力者の負担軽減につながるができる。また，作成の過程で行うさまざまな分析から，既存尺度の洗練化や，そこで測定されている概念の整理にもつながると考えられることから，心理尺度を用いる上で重要な試みの一つであると言える。しかしながら，どのようにしてその尺度を「短縮化」するか，すなわちどのように「短縮版」を作成するかについては，必ずしも十分な検討がなされてはいないのが現状である。また，多くの場合，尺度の短縮版の作成自体を目的として研究が行われるのではなく，作成された短縮版を用いることによって調査等の効率化を図るために，その過程として短縮版が作成されることが多いように思われる。実際，前述の **Rosenberg** 尺度を短縮化して利用している研究は，いずれも短縮化自体が目的ではなく，あくまで調査実施上の制約などから短縮化を行っているものであった。そのため，短縮版作成のプロセスは，あまり重要視されないことも多く，それらのプロセスについては，十分に記述されないことも少なくない。**Coste, Guillemin, Pouchot, & Fermanian (1997)** などでも，短縮版作成においては厳格な方法論が不足している場合もあることが指摘されており，一般的な（オリジナル版の）尺度作成に比べ，検討が不十分であることが多いと考えられる。オリジナル版の尺度がいかに精緻に作成されていたとしても，実際に用いられる短縮版が比較的安易に限られた情報のみをもとに作成されていたとしたら，その測定自体にも疑問を持たれかねなくなってしまう。尺度の乱立を防ぎ，既存尺度を有効に活用していくためにも，どのようにして短縮版を作成すべきかについては，さまざまな観点から検討が行われていく必要があるだろう。

以上のことから，本論文では心理尺度短縮版の作成に焦点を当てた検討を行う。オリジナル版の特長を十分に活かしながら，できるだけ精度の高い測定が可能になる短縮版を作成するためには，さまざまな情報を活用し，項目選択等を行うことが望ましいと考えられる。たとえば，従来の古典的テスト理論だけでなく，**Item Response Theory**（項目反応理論・項目応答理論；以下，**IRT** とする）に基づく情報を活用することは，その重要な選択枝の一つであると考えられる。しかしながら，そのような研究がどの程度行われているのかについても，まだ十分に整理されていないのが現状である。そこで本論文では，まず短縮版尺度作成の現状を把握した上で，実際に短縮版尺度を作成する試みを複数行うことで，短縮版作成に利用可能な情報や方法の提案や，短縮版作成における課題の整理を行って行くこととする。

## 1.2. 本論文の構成

本論文の全体を通しての目的は、さまざまな場面で用いられている各種心理尺度について、その測定精度をある程度保ちつつ調査協力者の負担軽減や調査の効率化を図る短縮版の作成法を、具体的な事例をもとに検討することである。ここでは、従来の古典的テスト理論に比べ、尺度や項目に関する精緻な情報が得られると考えられる IRT を用いた検討を中心に行うこととする。

本論文の具体的な構成は以下の通りである。まず、第2章では、心理尺度短縮版がどのような方法を用いて作成されているのかについて、先行研究から現状をレビューするとともに、日本での状況について新たにデータ収集を行い、課題の整理を行う。具体的にどのような情報や分析方法を用いて短縮版が作成されているのかを示す。

また、第3章から第5章では、日本で従来多く用いられていた方法とは異なる方法として IRT を取り上げ、実際に短縮版の作成を試みる。IRT を適用した分析を行い、その情報をもとに短縮版作成を行うことを通して、その活用可能性を検討する。なお、第3章、第4章では、さまざまな調査などで利用頻度が高いと思われる既存の2尺度（DSRS-C および Big Five 尺度）を取り上げる。また、第5章では、新規の尺度（継時的比較志向性尺度）の作成を行うとともに、その尺度の短縮版を作成する形で検討を行う。いずれも短縮版のニーズが高いと考えられる領域の尺度である。

第6章では、第5章までで取り上げた3種類の尺度を用いて、短縮版の作成方法の違いが項目選択等にどのような影響を与えるのかを検討する。具体的には、因子分析を用いた方法と、第3章から第5章までで採用された IRT を用いた方法との間で、選択される項目の比較および、作成された短縮版の信頼性、妥当性の比較を行う。そして、それらから得られた知見を整理した上で、第7章では今後の短縮版作成や尺度の活用に関して課題について論じる。

なお、本論文は以下で発表された研究をもとにしている。

- ・第2章

並川 努 (投稿中). 心理尺度短縮版の作成における現状と課題

- ・第3章

並川 努・谷 伊織・脇田 貴文・熊谷 龍一・中根 愛・野口 裕之・辻井 正次 (2011). Birlson 自己記入式抑うつ評価尺度 (DSRS-C) 短縮版の作成 精神医学, **53**, 489-496.

- ・第4章

並川 努・谷 伊織・脇田 貴文・熊谷 龍一・中根 愛・野口 裕之 (2012). Big Five 尺度短縮版の開発と信頼性と妥当性の検討 心理学研究, **83**, 91-99.



## ・第5章

- 並川 努 (2010). 継時的比較志向性尺度短縮版の作成——Item Response Theory を用いた検討—— 名古屋大学大学院教育発達科学研究科紀要 (心理発達科学), **57**, 71-76.
- 並川 努 (2011). 継時的比較の個人差——継時的比較志向性尺度の作成と検討—— 心理学研究, **81**, 593-601.

## 第2章 短縮版尺度作成の現状と課題

## 2.1. 短縮版作成の方法について

尺度を用いた測定を行う際、そこに複数の項目が含まれていることは、測定の信頼性を上げることや、測定誤差を減らすこと、変数間の関係をより明確にすることなどにつながる (Franke, Rapp, & Andzulis, 2013)。しかし一方で、項目数が増えると回答に時間もかかるため、調査協力者には多くの負担をかけることにもなりうる。実際、回答にかかる時間が長くなるほど、回答率や回答の質の低下が見られるなど (Edwards, Roberts, Clarke, DiGuseppi, Pratap, Wentz, & Kwan, 2002; Galesic, & Bosnjak, 2009; Herzog & Bachman, 1981)、そこで得られるデータの自体にもネガティブな影響を与えることが指摘されている。これら項目数の多寡によって生じるさまざまな影響はトレードオフの関係にあるとも言えるため、すべての場合において最適な項目数が一義的に定まるわけではない。そのため、新規に尺度を作成する際には、当該尺度の使用目的や想定される回答者の特徴などに応じて、一定の信頼性等を確保できる項目数を維持しながらも、できるだけ回答者にとって負担の少ない尺度を構成するように工夫することが求められる。

また、そのような工夫は、新規に尺度を作成する場合だけでなく、既存の尺度を活用して調査を行う場合も同様に重要である。特に既存の尺度を利用しようとする場合は、調査実施上の制約などから、既存の尺度をそのままの形では利用しにくい場合も少なくない。たとえば、1回の調査の中で複数の尺度を併用する場合や、限られた時間の中で測定を行う必要がある場合など、既存の尺度が作成された時とは状況や目的が異なり、より負担の少ない尺度が要求されることがある。このような場合、既存の「オリジナル版」の尺度から一定数の項目を削除した「短縮版」が作成され、調査に利用されることも多い。

しかし、この短縮版作成や短縮版の利用においては、通常の（オリジナル版の）尺度作成に比べて、必ずしも検討が十分行われているとは言えない面がある。実際、短縮版を作成するためのガイドライン等も十分に整備されているわけではないのが現状である (Goetz, Coste, Lemetayer, Rat, Montel, Recchia, Debouverie, Pouchot, Spitz, & Guillemin, 2013)。また、Coste et al. (1997) などでも指摘されているように、短縮版を作成している論文の中には厳格な方法論が欠けている場合もあり、比較的安易に項目を減らして尺度を利用しているようなケースも存在する。オリジナル版がいかに精緻に作成されていたとしても、短縮版の作成において削除項目等の検討が不十分であれば、その短縮版で測定された結果については疑義が生じることにもつながりかねない。

では、短縮版を作成する際に、どのような方法が多く用いられているのだろうか。この点について Goetz et al. (2013) は、1995年から2009年に発表された英語の論文の中から、短縮版作成に関する論文をデータベース (Medline, PsycINFO) を用いて103件の論文を抽出し、分析を行っている。それによると、短縮版の項目数は、オリジナル版に比べて中央値で57% (range: -88% to -21%) であり、オリジナル版の項目数が多いほど削除さ

れる項目数も多い ( $r=.97$ ) ことが示されている。また、短縮化の方法としては、全体の40%が内容的な分析と統計的なアプローチを併用しており、49%は統計的なアプローチのみを利用していた。そして、用いられる方法で最も多いのは因子分析を用いた統計的なアプローチによる方法であり、対象となった論文の52%で採用されていた。次いで、質的な分析(46%)やクロンバックの $\alpha$ 係数を利用した方法(34%)、項目と合成得点との相関(34%)などが多く用いられていた。同様の方法で1985年から1995年の論文を対象とした Coste et al. (1997) との比較では、IRT を用いた方法が Coste et al. (1997) では0だったのに対して、新たに10件見られているという変化を除けば、短縮版作成の方法は大きく変わっていないことが示唆されている。

これら Goetz et. al. (2013) 等の研究が示している短縮版作成に関する情報は、より効率的で精度の高い短縮版の作成方法を検討する上で基礎となるものであると考えられる。しかし、先行研究で対象とされているのは、主に英語による論文のみであり、他の言語で書かれた論文における状況は明らかになっていない。そのため、日本での短縮版作成にはどのような方法が用いられることが多いのか等の短縮版作成状況についても明らかになっておらず、日本国内での現状や課題は未整理のままである。そこで、本研究では日本語の論文を対象にレビューを行い、日本における短縮版作成の現状を整理し、今後の課題について検討を行うこととする。

## 2.2. 日本における短縮版の作成の現状

### 2.2.1. 目的

短縮版の作成方法としてどのような方法が多く用いられているかについては、Goetz et al. (2013) などのレビューはあるものの、英語以外の言語で書かれた論文における状況は十分に明らかになっていない。日本においても Goetz et al. (2013) の結果と同様に因子分析の結果を基準に短縮版を作成する研究が多くみられることなどが推測されるが、実際にどういった情報をもとに、短縮版が構成されているのかについての十分なレビューは見られない。そこで、日本で行われている心理尺度短縮版の作成方法について、レビューを行い、短縮版作成の方法論について議論を行うことを目的とする。

本研究において検討する点は、主に以下の4点である。まず、短縮版作成研究がどのくらい多く行われているのかを検討するため、i) 短縮版作成に関する論文数がどのくらい抽出されるのかを示す。その後、どのような短縮版が、どのような方法で作成されているのかについて検討するため、ii) オリジナル版と短縮版の項目数、両者の比(短縮化における項目の削減率)について検討する。さらに、iii) 短縮版の作成方法(どのような分析や指標をもとに作成されているのか)について集計を行い、最後にiv) オリジナル版と短

縮版との相違点（因子構造の違いなど）についてまとめる。これらを通して、日本における短縮版作成研究の現状を整理し、今後の短縮版作成の方法について議論を行うことを目的とする。

## 2.2.2. 方法

### 2.2.2.1. 対象論文の選択条件

本研究では、目的に挙げた i) ～ iv) について検討するために、以下の条件に合致する論文を対象論文として選択し、分析を行った。対象論文としてリストアップする条件は、①1980年から2009年の30年間に発表された学術論文であり<sup>1</sup>、②既存の心理尺度をもとに新たに短縮版を作成していること、③短縮版の作成方法（どのように項目の選択を行ったのかなど）が具体的に記載されていることの3点であった。英語などの他の言語で作成された既存の短縮版を日本語に翻訳したのみの場合や、「短縮版」「簡易版」という名称がつくものの直接のオリジナル版にあたる尺度が存在せずに新規に尺度作成を行っている研究、質問紙ではなく観察法や検査法に分類される研究は除外された。また、学会発表の抄録等は、紙幅の関係で情報量が少ない場合が多いことや、後に同一の内容が論文として公開されている場合も多く、それらと内容が重複する可能性が高いと考えられることなどから、ここでは除外対象とした。なお、新規にオリジナル版の尺度作成を行っている研究でも、その論文の中でオリジナル版を作成した上で、別途その短縮版を作成している場合は対象とした。

### 2.2.2.2. 手続き

本研究では、まずデータベースを用いて候補論文のリストアップを行い、その中から上記の条件に合致する論文を絞り込む形で論文の選択を行った。具体的な手続きは以下の通りであった。

**データベースを用いた論文リストの作成** 日本語の学術論文に関するデータベースとして J-STAGE（科学技術振興機構による電子ジャーナルサイト：<https://www.jstage.jst.go.jp/browse/-char/ja/>）および CiNii（国立情報学研究所による学術情報データベース：<http://ci.nii.ac.jp/>）の二つを用いて、論文の検索を行った。まず、それぞれのデータベースで「短縮版（or 簡易版）」および「作成（or 開発）」を標題もしくは抄録に含む論文を検

---

<sup>1</sup> ここでは、Goetz et. al. (2013) との対応や、本研究が最初に行われた時期および第3章以降との整合性等を考慮し、1980年から2009年までに時期を限定して集計を行った結果を報告することとした。なお、2010年以降（～2013年まで）についても別途検討を行っており、そのデータについては、appendix2に補足として一部を記載した。また、並川（投稿中）では、1980年から2013年までのデータを報告している。

索し、調査対象候補論文とした。その結果、J-STAGE では 62 件、CiNii では 131 件の計 193 件が該当した。

しかし、この 193 件の中には、既存の短縮版を使用しているだけの論文等も多く含まれており、ここから最終的な対象論文を絞り込むと十分な数を確保できないことが予想された。そこで、標題や抄録に明示的にこれらの用語を含まなくとも短縮版作成を行っている場合も多いと予想されることから、抽出漏れを補うために J-STAGE および CiNii の全文検索で「短縮版(or 簡易版)」+「作成」+「尺度」を含む論文を抽出した。その結果抽出された 1224 件 (J-STAGE 401 件 + CiNii 823 件) を候補に追加し、延べ 1417 本を候補論文とした。なお、最終的な検索時期は 2013 年 12 月であった。

**対象論文の絞り込み・選択** 次に、1417 本について、上述の①②③の条件に合致する論文の抽出を行った。まず、①に合致しない 1980 年から 2009 年以外の論文や、学会発表の抄録等を機械的に除いた後、各論文のアブストラクトや本文を直接確認し、対象論文の絞り込みを行った。その結果、最終的に抽出された論文数は 49 本であった (具体的な文献については appendix1 参照)。

## 2.2.3. 結果と考察

### 2.2.3.1. 抽出された論文の数

今回最終的に抽出された論文数は 49 本であり、これらの論文を対象に集計を行った。(なお、一つの論文内で複数の短縮版尺度を作成している場合も見られたため、対象となった短縮版尺度は全部で 57 尺度であった。) これらの論文の中には、短縮版作成そのものを目的とした論文も多く含まれていたが、短縮版の作成が目的ではない調査研究の中で、一部の尺度を短縮化し使用している論文も含まれていた。なお、標題に「短縮版の作成」などの表現が含まれており、短縮版作成が主たる目的の一つであると判断できる論文は、49 本中 35 本であった。

また、論文が掲載されていた雑誌の種別に分けて集計を行ったところ、学会が発行している学術誌に掲載されていた論文が 27 本、大学の紀要等に掲載されていた論文が 19 本、出版社その他が発行している雑誌に掲載されていた論文が 3 本であった。今回抽出された論文では、学会誌に掲載されていたものが最も多かったものの、紀要論文においても短縮版作成の試みが多く行われていることが示された。また、掲載誌では、「パーソナリティ研究 (「性格心理学研究」を含む)」の論文が 6 本と最も多く、次いで「心理学研究」と「心身医学」の 3 本であった。なお、その他掲載されていた雑誌名をいくつか取り上げてみても、「日本老年医学会雑誌」や「体育学研究」、「体力科学」、「がん看護」、「日本看護科学会誌」、などがあり、狭義の「心理学」に限らず医学や看護学等を中心に周辺領域に幅広く短縮版作成のニーズがあることが示唆される。

次に、今回抽出された 49 論文について、発表された時期を Figure1 にまとめた。今回抽出された論文の中で最も古いのは 1989 年であり、1980 年から 1989 年ではこの 1 本のみであった。一方で、1990 年～1994 年は 4 本、1995 年～1999 年は 8 本、2000 年～2004 年は 17 本、2005 年～2009 年は 19 本と、近年になるほど論文数は増加していた。この結果は、短縮版の作成や利用の頻度が近年特に高まっていることを示す可能性も指摘できるが、今回の研究ではデータベースをもとに論文の抽出を行っていることも一定の影響を与えていると考えられる。すなわち、日本語論文のデータベースが整備されてからまだ日が浅いため、発表時期がそれよりも前の論文は、まだデータベースに登録されていなかったり、登録されていても標題等の基礎的な情報のみが検索の対象となっていたりするなどして、うまく抽出できなかった可能性がある。この点は本研究で採用した方法の限界であると言えるだろう。しかしながら、少なくとも 1990 年代から 2000 年代にかけて、短縮版の需要が一定程度あり、コンスタントに短縮版が作成され続けていることが、ここから示唆される。

### 2.2.3.2. オリジナル版と短縮版の項目数

短縮版の項目数が、オリジナル版からどの程度削減されているのかについて検討を行うため、対象 49 論文からオリジナル版尺度と短縮版尺度の項目数についてまとめ、集計を行った (Table1)。なお、対象論文の中には、オリジナル版の項目数に関する記述が省略されていたものも一部に見られたため、その場合はオリジナル版に関する論文等を参照し、確認を行った。

まず、今回対象となった尺度のオリジナル版では、下位尺度の数は様々であるものの、尺度全体での項目数の平均値は、51.5 項目 ( $SD=34.9$ ) であった。項目数の最大値は 191、最小値は 16 であり、さまざまな長さの尺度が短縮されていた。なお、中央値は 40.0 であった。同様に、作成された短縮版の項目数では、平均値は 25.1 項目 ( $SD=18.6$ )、最大値

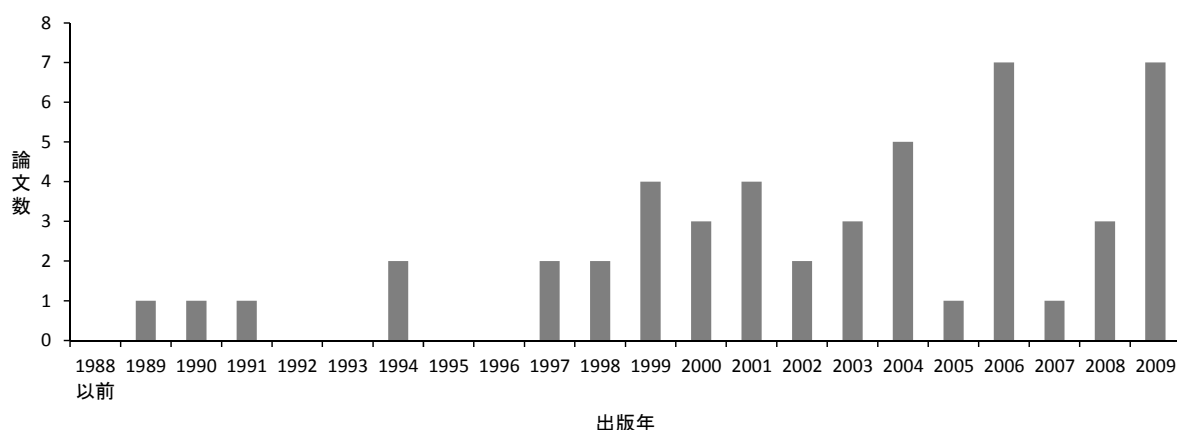


Figure1 短縮版作成論文の推移

は100, 最小値は4, 中央値は20.0であった。次に, 短縮版になることによって, オリジナル版からどの程度項目数が削減されているかについて検討した。ここでは, 尺度ごとに「短縮版の項目数÷オリジナル版の項目数×100」の値を求め, 短縮版作成による項目の削減率を計算した。その結果, 削減率はオリジナル版の平均49.7% ( $SD=15.9$ ) となっていた (range: 20.0 to 83.3)。今回の結果では, 短縮版はオリジナル版の概ね半分程度の項目数で作成されていることが示唆された。

### 2.2.3.3. 短縮版の作成方法

次に, 短縮版がどのように作成されているのかについて検討を行った。一般的に, 短縮版の「作成」には, 短縮版に用いる項目の選択から, 短縮版の信頼性・妥当性の検討までの一連のプロセスを含むと考えられる。しかし, 今回集計を行ったところ,

まずオリジナル版から一定の手続きを経て項目を選択し短縮版を(暫定的に)作成し, その後で別の検討として当該短縮版の信頼性・妥当性の確認が行われるケースが多く見られた。また, 設定した条件に合致しなかったため今回の対象論文には含めなかったが, 作成した短縮版の信頼性・妥当性の確認部分を別論文で行っているケース(たとえば上地・宮下, 2009a など)もいくつか見られた。そのため, ここでは短縮版「作成」のプロセスの中で, オリジナル版の項目から短縮版として用いる項目を選択する部分と, 短縮版の信頼性・妥当性の確認を行う部分とを便宜的に区別し, それらのプロセスのうち前者の, 項目を選択する部分に着目をし, 集計を行うこととした。すなわち, どのような分析結果や指標をもとに短縮版として用いる項目が選択されているのかについて明らかにする。なお, 一つの論文内で複数の短縮版尺度を作成している場合もあるが, 今回の対象とした49論文では, 基本的に同一論文内では同一の方法を用いてすべての短縮版の作成がなされていたため, ここでは論文単位で集計を行うこととした。集計は, Goetz et al. (2013) 等も参考にして, 作成方法を大きく質的(内容的)な分析と量的(統計的)な分析に分けた上で, 量的な分析については細かな方法ごとに行った。以下, 集計の結果をTable2に示すとともに, 多く見られた手法を中心に検討を行った。

Table1 オリジナル版と短縮版の項目数別論文数

	オリジナル版	短縮版
10以下	0	5
11～20	5	22
21～30	10	9
31～40	9	5
41～50	7	3
51～60	7	2
61～70	2	1
71～80	2	0
81～90	0	1
91～100	2	1
101～110	1	0
111～120	1	0
121～130	0	0
131～140	0	0
141～150	2	0
151～160	0	0
161～170	0	0
171～180	0	0
181～190	0	0
191～200	1	0



Table2 短縮版作成方法

	論文数	%
質的分析	22	44.9%
量的分析		
因子分析	42	85.7%
項目-合計相関	10	20.4%
回答の分布の偏り・集中	9	18.4%
$\alpha$ 係数	5	10.2%
G-P分析	3	6.1%
未経験等の回答率	3	6.1%
項目間の相関係数	2	4.1%
IRT	1	2.0%
級内相関係数	1	2.0%
オリジナル尺度との相関	1	2.0%
関連尺度との相関	1	2.0%
欠測値の数	1	2.0%

**因子分析** まず、短縮版の作成方法として最も多くの論文で行われていたのは、因子分析の結果に関する情報をもとに項目選択を行う方法であった（なお、ここでは便宜的に主成分分析も含めて「因子分析」として集計を行うこととした）。Table2 に示した通り、全体の 85.7%にあたる 42 の論文で因子分析の結果（因子負荷量など）を利用して項目選択を行っていた。

具体的な項目選択方法については、「当該因子に.40 以上負荷し、他の因子に.30 以上負荷しない項目を抽出」（有光，1999）など、因子負荷量などについて何らかの数値的な基準を明記している研究も見られたが、「各因子について因子負荷量の多い順に上位 5 項目を残した」（鈴木・依田・越川・杉若・嶋田・瀬戸・上里，1997）や「各因子の負荷量の正負において絶対値の高い項目をそれぞれ 5 項目、計 20 項目を選択し」（白澤・石田・箱田・原口，1999）など、具体的な数値は示さず、因子負荷量の値が大きい順に項目を選択している研究も多く見られた。なお、項目選択の基準としていた因子負荷量などの値も、論文によって.3 から.6 まで幅があり、用いている尺度の特徴等によって一貫してはいなかった。

なお、他の指標や分析結果を項目選択の際に利用した記述が見られず、基本的に因子分析の結果のみから短縮版に用いる項目の選択を行っている論文は全部で 16 本あり、全体の 32.7%を占めていた。85.7%の論文で用いられていたことも含めて、短縮版作成において因子分析の結果が非常に重視されている現状が示唆される。

**質的な分析** 因子分析に次いで多かったのが、統計的な数値などの量的な指標ではなく、項目の内容に関する質的な検討をもとに項目選択を行っている論文であった。これには、「研究者らにより各因子に含まれる項目の内容的妥当性を検討した上で」（石井・井上・大谷・小田切・高宮・下光，2009）や、「逆転項目の数を考慮し」（内藤・鈴木・坂元，2004）等の記述が含まれていた。なお、質的な検討の中には、さまざまな基準（内容的妥当性、逆転項目等の項目の性質など）やさまざまな評価者（著者自身、専門家、回答者）によるものがあり、これらを分けて集計することも検討された。しかしながら、「項目内容を考慮して項目を選出」（清水・川邊・海塚，2006）など、詳細に記述されていない論文も多く見られたため、ここでは質的な分析としてまとめて集計を行うこととした。

短縮版作成の過程で質的な分析を行っている論文は、全体の 44.9%にあたる 22 本が該当した。4 割以上の論文で、内容に関する質的な分析をもとに短縮版の項目選択が行われていることは、単に統計的な数値に関する検討だけでは、十分ではないことが多く認識されていることを示していると言えるだろう。特に、因子分析に基づく検討だけでは、類似した項目のみが選択されてしまう可能性が高くなると考えられる。その場合、内的な一貫性は高くなると考えられるものの、測定される対象の範囲が非常に狭くなり、オリジナル版で測定されていた概念全体が適切に測定されているかどうかは、精査が必要となると考えられる。

なお、これらのうち短縮版に用いる項目選択において、他の量的な指標を併用していないと思われる論文は、佐々木・星野・丹野（2002）、数見・土井・伊藤（2009）の 2 本だけであった。佐々木他（2002）は「逆転項目の数に偏りを生じないように留意しながら、類似した内容の項目を 2 項目ずつ削除」という形で、数見他（2009）は「内容的に類似し重複していると考えられる項目を半分削除」という形で、それぞれ作成していた。いずれも短縮版作成が主たる目的の論文ではなかったことから、短縮版作成を目的とした場合は、質的な分析だけでなく、何らかの統計的な分析と併せて検討されることが一般的であることが示唆される。

**その他の方法** 因子分析や質的な検討ほど多くは見られなかったものの、他にも項目選択の際に行われていた分析や参考にされていた分析・指標がいくつか見られた。まず、項目合計相関（I-T 相関）を計算している論文が 10 本あった。これは全体の 20.4%にあたり、統計的な方法としては、因子分析に次いで多く用いられていた。また、「1 つの選択肢に回答の 90%が集中した 1 項目を削除」（熊野・織井・山内・瀬戸・上里・坂野・宗像・吉永・佐々木・久保木，2000）などのように、各項目への回答の分布をもとに項目選択を行っている論文が 9 本見られた。これは、因子分析や I-T 相関など、項目間の相関に基づく方法とは異なる観点の方法であると言える。また、他には、クロンバックの  $\alpha$  係数を基準にしているものが 5 本、G-P 分析を行っているものが 3 本みられた（Table2）。

**先行研究との比較** 今回の結果を、Goetz et al. (2013) の結果と比較すると、「オリジナル版の尺度との相関」や、「IRT」を使っている研究が今回はそれぞれ一つと少なくなっていた点が異なっていた。「オリジナル版との相関」については、今回の集計には含めていないが、作成された短縮版の妥当性検討の部分では多く用いられており、20以上の論文が該当していた。ただし、その相関をもとに項目の取捨選択は行われていなかったため、これらはTable2にはカウントされなかった。この点については、Goetz et al. (2013) とは単純に比較できない点であると考えられる。

一方、「IRT」については、Table2 でカウントされたのは脇田・小塩・願興寺・桐山 (2007) の1本のみであった。また、作成された短縮版の分析にIRTを用いていた論文も、Table2には含まれていないが、1本(笹川・金井・村中・鈴木・島田・坂野, 2004)のみであった。しかし、これらを含めても今回は合計で2本の論文でしかIRTは利用されていなかった。先行研究(Coste et al., 1997; Goetz et al., 2013)においても、1995年以前についてはIRTによる分析は1本も報告されておらず、95年以降に見られるようになってきていることから考えると、IRTを用いた短縮版作成は今後の重要な検討課題の一つと言えるだろう。

#### 2.2.3.4. 短縮版とオリジナル版との違い

今回検討を行った短縮版の中には、単に項目数を削減した尺度だけではなく、オリジナル版とは異なる因子構造を備えた短縮版尺度が作成されている場合もいくつか含まれていた。そこで、次にオリジナル版と短縮版の相違点の一つとして、両者の因子構造の違いに焦点を当てて検討を行った。

まず、今回の結果では、短縮版を作成する中でオリジナル版とは下位因子数(下位尺度数)が異なる短縮版が作成されている論文が14件(28.5%)抽出された。いずれも短縮版になることによって因子数は削減されていた。オリジナルの尺度が測定していた幅広い測定範囲のうち、特に重要な部分や、より使用頻度の高い側面のみに測定範囲を絞った短縮版を用いることは、調査の目的によっては効率的な測定につながる事が予想されるため、重要な選択枝の一つであると言えるだろう。実際に、短縮版作成の過程において、内容的な検討をもとに(意図的に)特定の因子を削除していた研究(たとえば、上地・宮下, 2009bなど)も見られた。

しかしながら、一定の手続きで短縮版を作成した際に、結果としてオリジナル版とは異なる因子数になっていたという形の研究も複数存在した。因子分析の結果は、分析対象となったデータにも依存する。そのため、たとえば元の因子構造に曖昧な部分があれば、短縮版作成時に新たにデータを収集して探索的な因子分析を行うことによって、オリジナル版とは異なった因子構造が確認されるケースも少なくないと考えられる。その場合、新たに確認された因子構造を採用するのか、それともオリジナル版の因子構造を採用するのか

は、慎重な議論が必要となる。そしてもし新たな因子構造を採用した場合、その短縮版とオリジナル版との対応関係や、そもそもそれがオリジナル版の「短縮版」と呼びうるのかも議論の分かれるところであろう。たとえば、今回検討した短縮版の中には、尺度全体としては項目数がオリジナル版よりも削減されているものの、1因子当たりの項目数でみるとオリジナル版よりも多くなっている岩城（2008）のような尺度も見られた。この岩城（2008）の看護における社会的スキル尺度短縮版では、オリジナル版が55項目6因子だったのに対して、作成された短縮版は39項目4因子であり、1因子当たりの項目数は短縮版になることによって平均9.17から9.75へと0.6項目ほど増加していた。また、作成過程では多義的な項目の修正や新たな項目の追加も行われており、単純な短縮化ではなくオリジナル版改訂や修正といった要素も含まれていた。そのため、オリジナル版とは異なる特徴を持った短縮版が構成されているとも言えるだろう。

また、因子数自体は変わっていない場合でも、探索的な因子分析の結果から、オリジナル版と異なる因子に負荷したため、短縮版ではオリジナルとは別の因子（下位尺度）として解釈される項目が見られる場合もあった（たとえば、福井・原谷・外島・島・高橋・中田・深澤・大庭・佐藤・廣田，2004など）。これらは因子や下位尺度の妥当性という点で、短縮版だけでなくオリジナル版も含めて検討の余地があることを示唆しているとも言える。

これらの結果はいずれも、オリジナル版と短縮版では、それらが必ずしも同一と見なせるわけではないことを示唆していると考えられる。これ自体が即問題であるというわけではないものの、使用する短縮版の選択においては、こういった点についても留意した上で判断を行う必要があることを示していると言えるだろう。

なお、先行研究のGoetz et. al. (2013)では対象となった論文のうち55%が短縮化の過程でオリジナルの下位尺度構造を維持していたのに対し、41%は下位尺度や次元の数を減らしていた。また、短縮化に併せて尺度の改訂を行うなどして、因子・次元数を増やしていたものも4%含まれていた。これらと比較すると、今回の結果はオリジナル版と異なる構造の短縮版が作成されるケースは比較的少なかったと言える。

#### 2.2.3.5. 本研究の限界

今回の研究では条件に合致しなかったため除外されたものの、対象論文選定の過程で確認した論文の中には短縮版を利用して調査を行っているものが多く含まれていた。また、今回は、学会の大会発表の抄録等は記述されている情報が少ないことや他の論文と重複することが予想されたために対象から除外していた。しかしながら、短縮版作成については大会発表のみで報告されるという場合も、実際は少なくなかったと考えられる。さらに、明確に作成方法を記載せずに項目を減らして利用している論文などもあり、今回対象とした49論文だけでは短縮版の作成・利用実態を十分に反映していない可能性がある。特に、近年になるほど論文数が増加していた結果を踏まえると、今後もさらに多くの短縮版が開

発されていくことが推測される。実際、本研究の分析には含めていないが2010年から2013年の4年間で24本の短縮版作成論文が抽出されていた（appendix 2 参照）。そのため、今後も短縮版の作成方法については、状況を細かく見ていく必要がある。

また、今回はJ-STAGEとCiNiiの二つのデータベースを用いて、対象論文を抽出していた。しかし、日本語論文のデータベースは、英語論文のデータベース（たとえばWeb of Science, PsycINFO など）に比べると整備されたのが比較的最近であり、まだ検索で抽出できない論文も実際は多く存在している可能性がある。さらに、今回は検索された候補論文について、抄録や本文を人力で直接確認する方法で対象論文の選定を行ったため、検索の結果候補として抽出されつつも、最終的に見落とされた論文等も存在する可能性は否定できない。これらの点についても今後の課題であると言える。

「短縮版」のとらえ方・定義についても、課題が指摘できる。対象論文を選択する際の条件として、今回は「既存の心理尺度の項目を減らした短縮版を作成していること」を設定していた。しかし、単に尺度の項目数や因子数を減らすだけでなく、項目数はそのままでも選択枝の数を減らすことで回答者の負担軽減を図る短縮版（稲垣・井藤・佐久間・杉山・岡村・栗田，2013）なども見られた。実際、「非常にあてはまる」から「どちらともいえない」や「まったくあてはまらない」等のカテゴリがある7件法の中から回答を選ぶよりも、「はい」か「いいえ」の2件法でいずれかを選ぶ方が、回答にかかる時間は短くなることが推測される。その場合、項目を減らした場合と同様に回答者の負担を軽減することにつながる可能性がある。選択枝の数や、「あてはまる」等の選択枝のカテゴリの表現を変えることは、回答者の反応傾向にも影響を与える可能性があるが（脇田，2004），こういった多様な短縮版の検討も今後重要になってくると考えられる。

また、既存のオリジナル版を短縮するだけでなく、既存の複数の尺度を参考にしながら少数の項目で測定できる「短縮版」を新規に作成している論文（たとえば森・三浦，2007 など）も存在した。それらは今回の分析には含めていないが、目的に合った効率的な測定という意味では、重要な方法の一つであると言える。特に、利用される頻度が高く、類似した尺度も複数存在するような汎用性の高い概念については、このような既存尺度の整理を含めた形で検討が行われることも必要であろう。今後は、これらさまざまな方法を含めて、幅広く短縮版作成のあり方について検討を進めていく必要がある。

### 2.3. まとめと今後の課題

本章では、短縮版がどのように作成されているのかについて英語論文を対象とした先行研究を概観するとともに、日本語の論文を対象にレビューを行った。その結果、短縮版作成研究が近年増加傾向にあることや、作成方法として因子分析結果の利用が多いことなどの現状が示された。最後に、それらの現状を踏まえた上で、今後の短縮版作成における課題について整理を行いたい。

まず、今回の分析では近年、特に 2000 年以降において短縮版作成研究が多くみられることから、今後も同様に短縮版のニーズは多くなることが推測される。しかし、ある特定の調査のニーズに応じて尺度の短縮版を作成しても、それがその調査のみでしか用いられずに埋もれてしまうことも少なくない。また、使用頻度の高い尺度については、用いられる調査ごとに短縮版が作成され、複数の短縮版が存在する場合もある。たとえば、今回抽出された中でもパーソナリティの 5 因子モデルに関する FFPQ の短縮版を作成している論文が 2 つ含まれていた（佐々木・星野・丹野, 2002, 藤島・山田・辻, 2005）。それぞれ別の研究者によって作成されており、選択されている項目も異なり、対応関係も不明確である。藤島他（2005）の研究は、短縮版作成そのものを目的としたものであり、信頼性や妥当性の検討も丁寧に行われている。そのため、FFPQ に関しては今後この短縮版（FFPQ-50）が利用されることが多くなることが予想されるが、同様に複数の短縮版が作成されているケースは複数存在すると思われる。使用頻度の高い尺度については、十分な検討を行った短縮版が、その詳細な特徴やオリジナル版との関係、他の短縮版との関係も含めてきちんと提示されていくことは重要であろう。しかしながら、この FFPQ 等においても、より短い短縮版などのニーズも存在しうると考えられる。こういった多様なニーズにどう対応していくのかも今後の課題の一つであると言えるだろう。

次に、短縮版作成において、因子分析が非常に重視されている点について検討を行う。今回の結果では因子分析を用いているのは全体の 85.7%に上り、Goetz et al. (2013) の 52%という結果と比較しても非常に高い頻度で用いられていた。また、全体の 32.7%の論文では、因子分析の結果のみを用いて短縮版の項目を選択していると考えられることから、因子負荷量の大きい項目から順に選択をするという方法が、一般的な方法として広く用いられていることが示された。

しかしながら、因子分析で得られる因子負荷量は、分析に用いるデータに依存する。実際、Franke, Rapp, & Andzulis (2013) は、SOCO (Selling Orientation- Customer Orientation) 尺度 (Saxe & Weitz, 1982) を例に挙げ、因子分析に用いるデータが異なると因子負荷量が全く異なることを示している。そのばらつきは決して小さくなく、どの調査データを用いるかによって、短縮版として選択される項目が異なってくることも示唆される。そのため、因子負荷量だけではなく、さまざまな指標を用いて短縮版を構成する項目を選択すること

が重要になってくると考えられる。たとえば Stanton, Sinar, Balzer, & Smith (2002) は、5 因子 72 項目から構成される Job Descriptive index (JDI) の短縮版作成において、有用な指標として、因子負荷量、I-T 相関、正答率、外在指標などの 15 の指標を算出し、それらを分類・整理した上で短縮版を作成する試みを行っている。日本においても、こういった試みをもとに短縮版作成を行って行く必要があるだろう。ただし、因子分析が非常に重視されている現状を考慮すると、まずはいくつかの指標や方法を比較検討し、その有用性を日本語で発信していくことも重要であろう。

そういった観点から、今回短縮版作成に用いられる割合が Goetz et al. (2013) に比べて少なかった IRT を適用した分析も、重要な情報を得られる選択枝の一つであると言えるだろう。たとえば、因子分析による方法では、項目の困難度に関する情報が反映されないが、IRT を用いることでその検討を行うことも可能になる。また、古典的テスト理論では尺度の測定精度を表す際に信頼性係数（クロンバックの  $\alpha$  係数など）が用いられるが、IRT ではテスト情報量が用いられる。信頼性係数がテスト全体に対して 1 つの値を用いて精度を表すのに対し、IRT のテスト情報量は潜在特性尺度値によって異なる値を示すことができる。そのため、回答者の潜在特性尺度値ごとに測定精度を推定することができ、その尺度がどのレベルの潜在特性尺度値を持つ回答者に対して適しているのかなど、きめ細かい議論が可能になるという利点も備えている。この点を応用すれば、たとえば潜在特性尺度値が高い人から低い人まで幅広く測定することに適した尺度構成や、特定のカットオフ値周辺に高い測定精度を示す尺度構成など、目的に応じた柔軟な項目選択をすることが可能になる。これらは、短縮版を構成する上で、非常に重要な情報である。また、従来の方法では、測定に用いる項目によって測定次元が定義されるため、項目が少なくなれば測定の結果得られる得点の測定次元は厳密には異なるものになってしまう。しかし、IRT の項目パラメタを用いて回答者の特性尺度値を推定することで、理論上はオリジナル尺度の測定次元を維持したまま測定できる短縮版を作成することも可能となる。実際の分析結果では、IRT と古典的テスト理論の結果は類似するとの指摘もあるが（たとえば Roznowski, 1989）、IRT を用いた短縮版開発を進めることも有効であると考えられる。ただし、IRT を用いた短縮版作成の試みはまだ少なく、IRT を用いた場合と、因子分析を用いた場合では、短縮版として採用される項目がどのように異なるのかも、十分検討されているとは言えない。これらの点についても、今後十分な議論が必要である。

## 第3章 抑うつ尺度（DSRS-C）短縮版に関する検討



### 3.1. はじめに

前章では、短縮版作成の現状について概観し、短縮版作成の際に用いる情報としては因子分析によるものが最も多いことなどが示された。特に、日本語の論文を対象とした調査では、短縮版作成の9割近くが因子分析を用いており、約3割の論文では基本的に因子分析結果のみを用いて短縮版の項目選択を行っていた。しかし、短縮版作成においては、因子分析以外の方法も併用してさまざまな情報をもとに項目選択を行う必要があると考えられる。その際、前章でも指摘したように、近年さまざまな領域で注目されているテスト理論である Item Response Theory（以下、IRT とする）を用いることも、重要な選択枝の一つである。Goetz et al. (2013) による報告においても、近年 IRT を用いて短縮版を作成する研究が新たに見られるようになっていることが指摘されており、今後さらなる検討が必要な点である。

そこで、本研究では IRT を適用した分析を通して短縮版作成を行い、短縮版作成における課題等を整理することとする。本章以降、複数の尺度を取り上げて検討を行うが、まず本章では、IRT を用いた短縮版作成の試みの一つとして、Birlleson (1981) の子ども用抑うつ自己評価尺度（Depression Self-Rating Scale for Children; 以下、DSRS-C とする）のデータを取り上げ、分析・検討を行う。次節でもあらためて触れるが、DSRS-C は子どもを対象とした抑うつ測定に多く用いられる尺度の一つである。この尺度のように、子どもを対象とした測定においては、一般成人を対象とした場合に比べて、長時間にわたって回答を求めることは困難になる場合が多い。そのため、短縮版のニーズが多い分野の一つであると考えられる。また、抑うつのように臨床心理学や精神医学領域で用いられる尺度は、スクリーニング等を目的とし、一定の精度は確保しつつも短時間での効率的に測定できる特徴を備えていることが必要とされる場合も多い。よって、本研究では、短縮版作成研究の事例として、まず DSRS-C のデータを取り上げ、IRT を用いた分析を行い、短縮版を構成する場合に、どのような項目の組み合わせが考えられるかを示し、その信頼性・妥当性について検討することとする。

### 3.2. 問題と目的

これまでは大人の問題として扱われていた抑うつが、近年児童期などの子どもの問題としても注目されている。我が国でもさまざまな領域で子どもの抑うつ状態をとらえようとする研究が行われており（傳田・賀古・佐々木・伊藤・北川・小山，2004；黒田・櫻井，2001；村田，1993；佐藤・下津・石川，2008），それらによって一定程度の割合で抑うつ状態の子どもが存在するということが指摘されている（谷・吉橋・神谷・宮地・野村・

伊藤・辻井, 2010)。そのため、子どもを対象とした抑うつのアセスメントや治療法、予防法の研究に対するニーズは高く、今後それらの研究がさらに進められることが期待されている。

抑うつ状態の測定は、研究場面においては心理尺度が多く用いられており（奥村・亀山・勝谷・坂本, 2008）、子どもを対象とした研究においても、同様に心理尺度が用いられる可能性が高い。しかし、子どもを対象に測定を行う際には、大人を対象とした場合に比べ、回答に要する負担はできる限り少なくする必要がある。たとえば、回答すべき項目数が多すぎると、一定時間内ですべての項目に回答することができなかつたり、欠測値が多く生じたりするなど測定に支障が生じる可能性が、子どもの場合は大人に比べて特に高くなることが指摘できる。そのため、回答する子どもにとっても、できる限り負担が少なくなるように、項目数の少ない短縮版の開発が強く望まれる。

現在、児童期の抑うつ状態に関する調査においては、自己報告式の尺度である Birlerson (1981) の子ども用抑うつ自己評価尺度 (Depression Self-Rating Scale for Children; DSRS-C) がよく用いられている。この尺度は 18 項目から構成されており、単独で用いる際には項目数が多い尺度というわけではない。しかし、複数の尺度を組み合わせで実施したり、臨床現場で短時間に測定を行ったりする必要のある場合には、回答者にとって負担となる可能性もある。そこで本研究では、子どもを対象とした抑うつ尺度として、DSRS-C を取り上げ、その短縮版の作成を試みる。

短縮版作成の方法には、前章でも概観したように因子分析における因子負荷量を基準にする等、いくつかの方法があるが、その方法論は十分に確立されていない。本研究では、前章での結果などを踏まえ、近年さまざまな領域で注目されているテスト理論である IRT を用いて短縮版作成の検討を行う。IRT を適用することにより、従来の古典的テスト理論に基づいた方法に比べ、当該尺度や項目についてより多くの情報を得ることが可能になり、

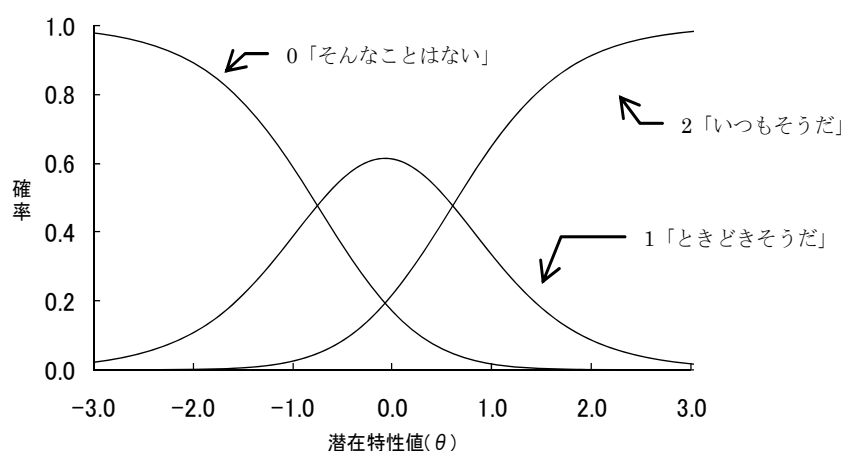


Figure 1 IRTCC の例

短縮版作成においてもより精緻な検討が可能になる。

IRT では、テスト（尺度）によって測定される特性（抑うつの程度など）を潜在特性尺度値（ $\theta$ ）で表現する。そして項目の特徴を、潜在特性尺度値（ $\theta$ ）の関数として表現する。たとえば、本研究で扱う DSRS-C の回答選択枝は、「そんなことはない (0)」「ときどきそうだ (1)」「いつもそうだ (2)」の 3 件法であり、得られるデータは 0, 1, 2 といった多値型のデータである。これに多値型のモデルである Generalized Partial Credit Model (Muraki, 1992; 以下 GPCM とする) を適用すると、各項目について Figure1 のような項目反応カテゴリ特性曲線 (Item Response category Characteristic curve; 以下 IRCCC とする) が得られる。

IRCCC は、特定の潜在特性尺度値（ $\theta$ ）を持つ回答者が、それぞれの選択枝をどの程度の確率で選択するかを表わしている。例えば、Figure1 では、平均的な抑うつの程度（ $\theta = 0.0$ ）をもつ人がこの項目に「ときどきそうだ」と答える確率は 0.6 程度であり、「いつもそうだ」「そんなことはない」と答える確率はそれぞれ 0.2 程度であるという情報を得ることができる。

この IRCCC は、slope（識別力）、location（位置）、category の 3 種類のパラメタで形状が決定される。このうち、slope パラメタは、各曲線の勾配に関係し、この値が大きい項目ほど曲線の傾きが急になり、その項目が回答者間での潜在特性尺度値  $\theta$  の違いを明確に識別できる項目であることを示す。また、location パラメタは項目の潜在特性尺度上における位置を表し、この値が大きいほど曲線の位置は右側に、小さいほど左側になる。そのため、location パラメタが大きい項目ほど、「いつもそうだ」と答えるのが難しい項目ということになる。category パラメタは、各曲線の交点に関連したパラメタである。

また、古典的テスト理論ではテスト（尺度）の測定精度を表す際に信頼性係数が用いられるが、IRT ではテスト情報量が用いられる。信頼性係数がテスト全体に対して 1 つの値を用いて精度を表すのに対し、テスト情報量は潜在特性尺度値により異なる値を示す。そのため、回答者の潜在特性尺度値ごとに測定精度が推定でき、そのテストがどのレベルの潜在特性尺度値を持つ回答者に対して適しているのかなど、きめ細かい議論が可能になるという利点も備えている。（IRT やそれを適用した測定についてのより詳細な説明は Embretson & Reise (2000), Fayers & Machin (2007) などに詳しい。）

以上のことから、本研究では IRT を適用した分析を通して、DSRS-C 短縮版を作成することを目的とする。

### 3.3. 方法

#### 調査対象者

調査協力市における全小学校の3学年から6学年の児童、および全中学校の1学年から2学年の普通学級に所属する生徒4683名(男性2414名, 女性2269名)を対象とした。

#### 手続き

調査は浜松医科大学の倫理委員会の承認を得て行われた。研究の目的と方法、およびプライバシー保護に関する説明を、A県X市の教育委員会に対して行い、調査への協力を依頼した。その結果、X市の全校(小学校, 中学校)から調査への同意が得られた。

そこで、X市の小・中学校に在籍する児童・生徒を対象として、質問紙による一斉調査を行った。調査は、授業時間内にクラス単位で担任教師によって実施された。実施に際し、調査への協力は各自の意志で拒否できること、拒否をしても不利益は生じないことを教示した。調査の結果は、各個人に文書でフィードバックされた。

#### 質問項目

村田・清水・森・大島(1996)のBirleson Depression Self-Rating Scale for Children (DSRS-C)の日本語版を用いた。DSRS-Cは合計18項目から構成され、3件法(「いつもそうだ」、「ときどきそうだ」、「そんなことはない」)の尺度であった。

また、DSRS-C短縮版の構成概念妥当性について検討するために、谷他(2010)と同様に攻撃性および不安の測定を行った。攻撃性の尺度としてはBuss-Perry Aggression Questionnaire (Buss & Perry, 1992)を日本の小学生用に改定したthe Hostility-Aggression Questionnaire for Children (HAQ-C) (坂井・山崎・曾我・大芦・島井・大竹, 2000)を、不安の尺度としては曾我(1983)の日本語版Stait-Trait Anxiety Inventory for Children (STAIC)より特性不安尺度を用いた。HAQ-Cは、「敵意(6項目)」、「短気(5項目)」、「身体的攻撃(6項目)」、「言語的攻撃(5項目)」の四つの下位尺度の計22項目から構成される4件法の尺度であり、日本語版STAICの特性不安は20項目から構成される3件法の尺度であった。

#### 分析方法

IRTを適用した分析に際しては、まず因子分析(主因子法)を行い、その際に計算される固有値の減衰状況をもとに各下位尺度の1次元性を確認した。また、IRT分析で用いる具体的なモデルとしてはGPCM (Muraki, 1992)を採用し、パラメタ推定用の計算機プログラムにはPARSCALE (Muraki & Bock, 2003)を用いた。

Table1 各項目の IRT パラメタおよび短縮版採用項目

		slope (SE)	location (SE)	c1 (SE)	c2 (SE)
活動性および楽しみの減衰					
*	1 楽しみにし	0.847 (0.021)	1.298 (0.022)	1.461 (0.024)	-1.461 (0.060)
	2 とても良く	0.490 (0.016)	1.101 (0.034)	1.067 (0.040)	-1.067 (0.062)
*	4 遊びに出か	0.527 (0.021)	1.890 (0.055)	0.513 (0.041)	-0.513 (0.080)
*	7 元気いっぱ	1.201 (0.034)	1.238 (0.019)	0.794 (0.019)	-0.794 (0.040)
	8 食事が楽し	0.824 (0.025)	1.309 (0.025)	0.838 (0.026)	-0.838 (0.050)
	9 いじめられ	0.405 (0.015)	0.977 (0.039)	0.809 (0.048)	-0.809 (0.066)
*	11 やろうと思	0.769 (0.016)	0.170 (0.025)	1.745 (0.034)	-1.745 (0.039)
*	12 いつものよ	1.277 (0.031)	0.636 (0.015)	0.955 (0.018)	-0.955 (0.028)
	13 家族と話す	0.600 (0.020)	1.287 (0.032)	0.842 (0.033)	-0.842 (0.058)
	16 落ち込んで	0.737 (0.021)	0.802 (0.023)	0.847 (0.028)	-0.847 (0.041)
抑うつ気分					
*	3 泣きたいよ	1.038 (0.031)	1.403 (0.022)	0.812 (0.022)	-0.812 (0.047)
	5 逃げ出した	1.051 (0.032)	1.396 (0.023)	0.663 (0.022)	-0.663 (0.045)
	6 おなかが痛	0.414 (0.011)	1.512 (0.039)	2.031 (0.046)	-2.031 (0.089)
*	10 生きていて	0.898 (0.032)	1.780 (0.037)	0.320 (0.030)	-0.320 (0.060)
	14 こわい夢を	0.473 (0.016)	1.497 (0.039)	1.285 (0.040)	-1.285 (0.075)
*	15 独りぼっち	1.097 (0.035)	1.371 (0.023)	0.485 (0.023)	-0.485 (0.041)
*	17 とても悲し	1.750 (0.054)	1.383 (0.016)	0.586 (0.016)	-0.586 (0.034)
	18 とても退屈	0.523 (0.017)	1.192 (0.034)	0.958 (0.037)	-0.958 (0.060)

\*短縮版採用項目

なお、著作権等に配慮し、本論文では尺度の項目については原則的に一部のみを表記する形にした。

Table2 各下位尺度のパラメタの平均値および標準偏差

		slope		location	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
活動性および楽しみの減衰	オリジナル版	0.768	0.274	1.071	0.440
	短縮版	0.924	0.279	1.046	0.591
抑うつ気分	オリジナル版	0.906	0.413	1.442	0.157
	短縮版	1.196	0.561	1.484	0.613

### 3.4. 結果

#### 尺度の1次元性の確認

DSRS-C は 2 因子構造であるという指摘 (佐藤・新井, 2002) に従い, 谷他 (2010) を参考に「活動性および楽しみの減衰」10 項目, 「抑うつ気分」8 項目の二つを下位尺度とし, それぞれで因子分析 (主因子法) を行った。固有値の減衰状況は, 「活動性および楽しみの減衰」で 3.58, 0.94, 0.87, …, 「抑うつ気分」3.16, 0.96, 0.83, …, であり, い

ずれも第1固有値が第2以下の固有値に比べて極めて大きな値を示し、1次元性が満たされていると判断された。

#### 項目パラメタの推定と短縮版項目の検討

下位尺度ごとに項目パラメタ推定を行った（Table 1）。また、下位尺度ごとに項目パラメタ推定値の平均および標準偏差を求め Table 2 に示した。続いて各項目のパラメタ推定

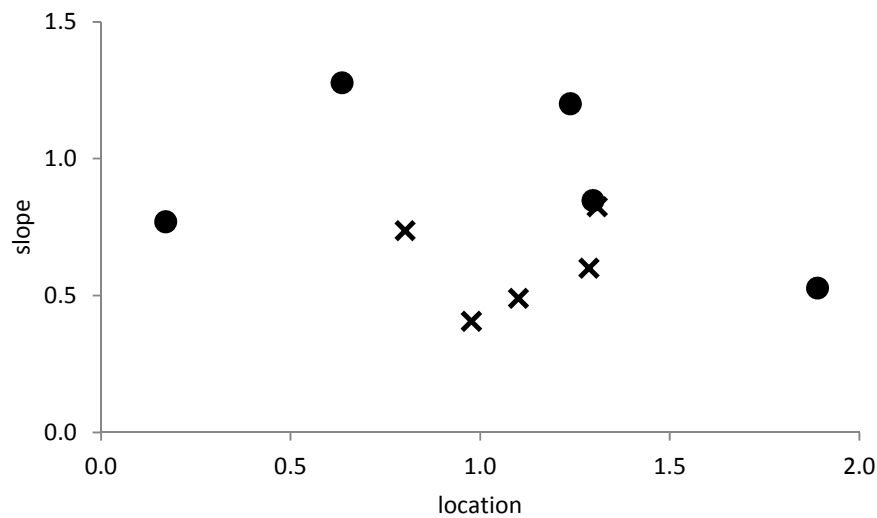


Figure2 「活動性および楽しみの減衰」の項目パラメタ

● 短縮版採用項目, ×短縮版不採用項目

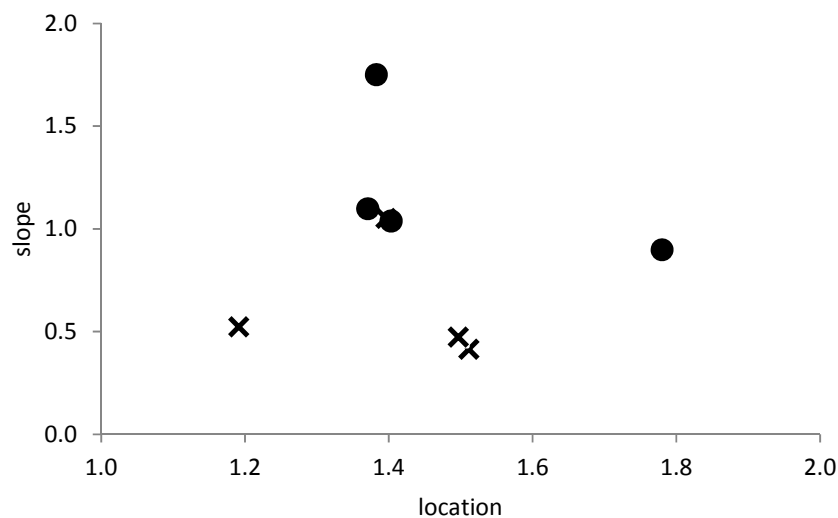


Figure3 「抑うつ気分」の項目パラメタ

● 短縮版採用項目, ×短縮版不採用項目

値を参考に短縮版の項目選択を行った。選択の際には、項目の slope パラメタの推定値が大きいことや、幅広い潜在特性尺度値の推定が可能になるように location パラメタの推定値が高いものから低いものまでバランスよく散らばるように配慮した。また、パラメタ推定値が類似している項目があった場合は、他の項目とできるだけ内容的に意味が異なるものを選ぶ等、項目内容にも配慮した。最終的に、「活動性および楽しみの減衰」から 5 項目、「抑うつ気分」から 4 項目が短縮版の項目として選択された (Figure2, Figure3)。短縮版に採用した項目は Table 1 に\*を記す形で示した。また、短縮版での各パラメタ推定値の平均値も Table 2 に示した。

なお、選択された計 9 項目については、20 年以上の診療経験のある 3 名の児童精神科医によって内容的妥当性の検討が行われ、児童の抑うつ状態を測定する尺度としてこの 9 項目が適切であることが確認された。

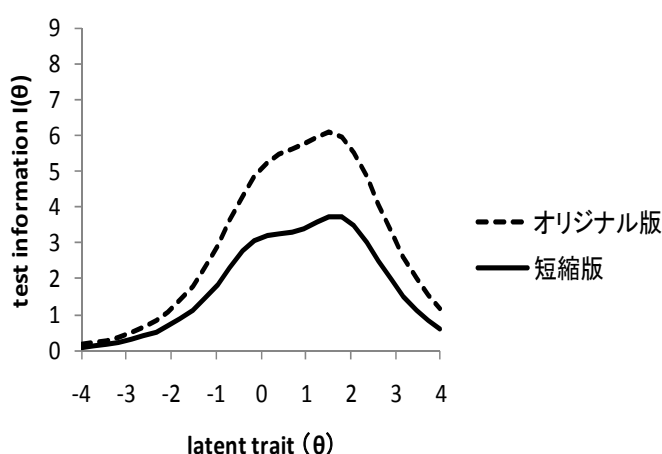


Figure 4 活動性および楽しみの減衰のテスト情報量

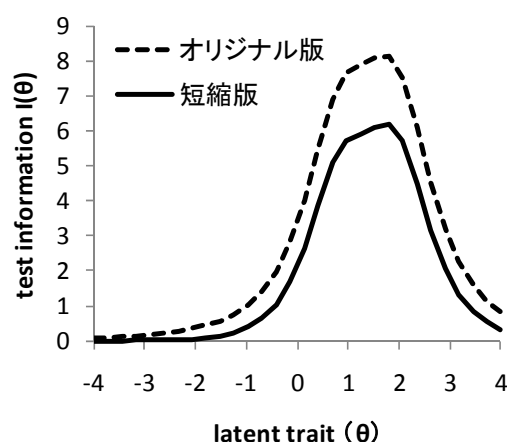


Figure 5 抑うつ気分のテスト情報量

Table 3 DSRS-C 短縮版の因子分析結果

	I	II
抑うつ気分		
17 とても悲し	.82	-.05
3 泣きたいよ	.64	-.06
15 独りぼっち	.62	.07
10 生きていて	.47	.11
活動性および楽しみの減衰		
12 いつものよ	.01	.73
7 元気いっぱい	.02	.64
1 楽しみにし	-.04	.59
11 やらうと思	.04	.46
4 遊びに出か	.00	.43
因子間相関		.46

さらに、短縮版尺度の測定精度を検討するために、信頼性係数の推定値として Cronbach の  $\alpha$  係数を求めた。「活動性および楽しみの減衰」のオリジナル版では.79 であり、短縮版では.71 であった。また「抑うつ気分」では、オリジナル版で.77、短縮版では.73 であった。次に、オリジナル版、短縮版それぞれのテスト情報量  $I(\theta)$  を求め、Figure 4 および Figure 5 に示した。短縮版における最大情報量は「活動性および楽しみの減衰」では 3.75、「抑うつ気分」では 6.20 であった。

#### 短縮版の因子的妥当性の検討

次に、短縮版においても、オリジナル版と同様の 2 因子構造が保持されるかについて検討を行った。短縮版 9 項目を用いて因子分析（主因子法・プロマックス回転）を行ったところ、固有値は第 1 固有値から順に 3.10, 1.47, 0.84, 0.74, 0.71, …, であり、短縮版においてもオリジナル版と同様の 2 因子構造（「抑うつ気分」4 項目、「活動性および楽しみの減衰」5 項目）が確認された（Table 3）。

#### 短縮版の構成概念妥当性の検討

短縮版の得点とオリジナル版の得点間の相関係数は、「抑うつ気分」で.89、「活動性および楽しみの減衰」で.90、総合得点で.92 であった。また、オリジナル版の IRT 尺度得点との相関係数は、短縮版の「抑うつ気分」で.83、「活動性および楽しみの減衰」で.90 であり、オリジナル版の「抑うつ気分」で.92、「活動性および楽しみの減衰」で.90 で、全体として極めて高い相関が得られた（Table 4）。



Table 4 DSRS-C オリジナル版と短縮版、関連尺度間の相関

	短縮版			オリジナル版			不安	攻撃性
	抑うつ気分	活動性減衰	抑うつ総得点	抑うつ気分	活動性減衰	抑うつ総得点		
短縮版(9項目)								
抑うつ気分		.36	.78	.89	.42	.72	.61	.35
活動性および楽しみの減衰			.86	.41	.90	.80	.39	.14
抑うつ総得点				.76	.82	.92	.59	.29
オリジナル版(18項目)								
抑うつ気分					.47	.82	.68	.39
活動性および楽しみの減衰						.89	.46	.17
抑うつ総得点							.65	.31
不安総得点								.38
攻撃性総得点								
オリジナル版(18項目)IRT尺度値								
抑うつ気分	.83			.92				
活動性および楽しみの減衰		.90			.90			

※ すべて 5%水準で有意

Table 5 DSRS-C のカットオフ値と関連指標

短縮版	感度	特異性
5	1.00	.70
6	.98	.83
7	<b>.93</b>	<b>.92</b>
8	.79	.97
9	.59	.99
10	.40	1.00

攻撃性との関連では、HAQ-C と「抑うつ気分」の相関はオリジナル版が.39 であるのに対して、短縮版が.35, 「活動性および楽しみの減衰」との相関はオリジナル版が.17 であるのに対して、短縮版が.14, 「総得点」との相関はオリジナル版が.31 であるのに対して、短縮版が.29 と極めて近い値を示した。

不安との関連では、日本語版 STAIC と「抑うつ気分」との相関はオリジナル版が.68 であるのに対して、短縮版が.61, 「活動性および楽しみの減衰」との相関はオリジナル版が.46 であるのに対して、短縮版が.39, 「総得点」との相関はオリジナル版が.65 であるのに対して、短縮版が.59 と極めて近い値を示した。

### 短縮版でのカットオフ値の検討

次に、オリジナル版 18 項目の得点を用いて抑うつ者を判定した後に ROC 分析を行い、DSRS-C 短縮版 9 項目について対象者の受信者動作特性曲線 (ROC 曲線) を作成し、関連指標を求めた (Table 5)。結果を元にカットオフ値を検討した結果、感度、特異性のいずれについても十分な値が得られたため、短縮版については 7 点をカットオフ値として用いることとした。なお感度は、(短縮版でカットオフ値を上回った人数) / (オリジナル版でカットオフ値を超えた人数) によって算出し、特異性は (短縮版でカットオフ値を下回った人数) / (オリジナル版でカットオフ値を下回った人数) から算出した。

## 3.5. 考察

本研究では、IRT を用いて DSRS-C の特徴を検討し、その結果をもとに 9 項目からなる短縮版の作成を試みた。今回作成された短縮版での location パラメタの平均値は、オリジナル版とほぼ等しい値を示しており、尺度の困難度としてはオリジナル版も短縮版も同程度であると考えられる。テスト情報量を見ると、「活動性および楽しみの減衰」の短縮版では最大情報量が 3.75 であり、情報量としては比較的低い値となった。ただし、IRT のテスト情報量において一定の条件の下では、信頼性係数  $\rho = .70$  に相当するのが 3.33 であり、 $\rho = .80$  に相当するのが 5.00 であるため、短縮版としては一定程度の精度が確認されたとと言える。また、オリジナル版においても最大情報量が 6.09 と必ずしも高い値を示すわけではなく、項目内容も含めて更に検討が必要である。一方、「抑うつ気分」では、短縮版の最大情報量が 6.20 であり、短縮版の精度としては許容できるものと考えられる。

また、短縮版でもオリジナル版と同様の 2 因子構造が確認され、因子の妥当性が示されるとともに、短縮版においても「活動性および楽しみの減衰」と「抑うつ気分」の 2 つの下位尺度によって抑うつ測定が可能であることが示された。さらに、DSRS-C 短縮版の得点と DSRS-C オリジナル版の得点の相関や IRT 尺度得点との相関も非常に高く、DSRS-C 短縮版の得点化が妥当なものであることが示された。また、DSRS-C 短縮版の得点と HAQ-C 並びに STAIC との相関が、DSRS-C オリジナル版とそれらの相関と極めて類似した傾向を示すことが明らかにされた。さらに、カットオフ値に関する検討でも、十分な感度と特異性が得られ、オリジナル版と同程度の高いスクリーニング特性が認められた。

これらの結果から、今回作成された短縮版は、臨床場面などで必要に応じてオリジナル版に替えて利用できる可能性が示された。また、従来多く行われていた因子分析を用いる方法ではなく、IRT を用いる方法によっても、一定の信頼性・妥当性を持つ短縮版を作成することが可能であることが、あらためて示唆されたと言える。今後は、他の尺度につい

ても同様に IRT を用いた検討を行い、手続きを洗練させていくことも重要である。

## 第4章 Big Five 尺度短縮版に関する検討

## 4.1. はじめに

第4章では、前章に引き続き一般的に多く用いられている尺度について、IRTを適用した分析に基づいて短縮版を作成する試みを行い、短縮版作成方法についての検討を行う。本章では、尺度を用いた研究が多いと考えられるパーソナリティ心理学領域において、広く用いられているBig Five尺度(和田, 1996)を取り上げる。Big Five尺度は、5因子モデルに基づいた尺度であり、一般的なパーソナリティの指標として多く用いられている。また、他の概念との関連が検討されることも多く、言い換えれば、他の尺度と併用されることが多い尺度の一つであると言えるだろう。そのため、短縮版のニーズも高いと考えられる。実際、第2章のレビューにおいても、5因子モデルに関する尺度であるFFPQの短縮版が、複数の論文で作成・利用されていた。また、ここで取り上げる和田(1996)のBig Five尺度自体も、第2章のレビューでは最終的に対象として抽出されなかったものの、内田(2005)などの複数の研究の中で短縮版が作成・利用されていた。そこで、本研究では、Big Five尺度(和田, 1996)について、IRTを用いた検討を行うこととする。

## 4.2. 問題と目的

パーソナリティを情緒不安定性(Neuroticism)、外向性(Extraversion)、開放性(Openness)、調和性(Agreeableness)、誠実性(Conscientiousness)の五つからとらようとするパーソナリティ特性の5因子モデル(Five-Factor model; 以下FFMとする)は、1990年代後半以降、「心理学」に限らず多くの領域で採用されている。関連する尺度もNEO-PI-R(Costa & McCrae, 1985)をはじめとして多数開発されており、日本においてもそれぞれ尺度作成のアプローチは異なるがNEO-PI-R日本語版(下仲・中里・権藤・高山, 1999)やFFPQ(FFPQ研究会, 1998)、主要5因子性格検査(村上・村上, 1999)などが開発されている。それらの中でも日本において特に多く用いられている尺度に、和田(1996)のBig Five尺度がある。この尺度は、各因子12項目合計60の形容詞により構成されており、文章を用いた尺度よりも構造が比較的安定して抽出されやすいことや、他の尺度よりも項目数が少ないといった特徴を備えている。そのため、今後も様々な場で活用されていくことが予想される。

しかし、実際に研究等に利用していく上では、Big Five尺度にもいくつかの検討すべき課題が存在する。中でも特に重要と考えられるのが、回答者の負担軽減である。例えば、研究や臨床等多くの場合では一つの尺度が単独で用いられることは少なく、他の尺度や実験課題などと併用される場合が多い。その際、Big Five尺度の60項目という項目数は、他の5因子モデルの尺度よりは少ないものの回答者にとって大きな負担となりうる。また、

質問紙に不慣れな子どもや高齢者などが回答する際には、60 項目のみでも実施が困難になる場合もある。これらの負担軽減は研究倫理的な面でも改善が望まれる点であろう。

そこで期待されるのが短縮版尺度の開発である。他の FFM の尺度でも、NEO-PI-R の短縮版である NEO-FFI や、藤島他（2005）による FFPQ 短縮版（FFPQ-50）など、すでにいくつかの短縮版が開発されているものの、Big Five 尺度と比べて必ずしも負担が少ないわけではない。また、Big Five 尺度に関しても、これまでさまざまな研究の中で短縮版が開発され、用いられてきた（たとえば内田, 2005）。しかし、短縮版に用いられる項目は、研究によって異なっている場合も多く、それぞれの研究の知見が有効に活用されているとは言い切れないのが現状である。

これには、短縮版作成の方法に関する検討が、日本においてはまだ十分になされていないことが影響していると考えられる。海外では、たとえば Stanton, Sinar, Balzer, & Smith（2002）は、5 因子 72 項目から構成される Job Descriptive Index（JDI）の短縮版作成において、有用な指標として、因子負荷量、I-T 相関、正答率、外在指標などの 15 の指標を算出しそれらを分類・整理した上で短縮版を作成する試みを行っている。しかしながら、日本においては少ない情報のみから短縮版が作成されることが多い。たとえば、従来短縮版を作成する際によく用いられる方法の一つに、因子分析を行い各因子の負荷量の高い項目から順に選択する方法がある。しかし、この方法では項目の困難度に関する情報が反映されていないことなど、必ずしも十分な検討が可能であるとは言い切れない。そのため、より測定精度が高く、妥当性も高い短縮版の開発に向けた多様な議論が重要である。特に、テスト理論の観点も含め、詳細な情報をもとにした検討が行われる必要があるだろう。

近年、重要度が増しているテスト理論として、Item Response Theory（以下 IRT とする）が挙げられる。IRT は、従来の古典的テスト理論に基づく方法に比べて、尺度や項目に関する精緻な情報が得られる。たとえば、測定精度にかかわる指標であるテスト情報量は、古典的テスト理論における信頼性係数の推定値（例えばクロンバックの  $\alpha$ ）とは異なり、単一の値ではなく、能力値の関数として測定精度を表わすことができる。また、IRT では項目の困難度と回答者の特性尺度値とが同一尺度上に位置づけられる。そのため、IRT を適用することによって、幅広い能力を測定することに適した尺度構成や特定のカットオフ値周辺に高い測定精度を示す尺度構成など、目的に応じた柔軟な項目選択をすることが可能になる。また、従来の方法では、短縮版にすることにより、測定次元が変わるが、IRT の項目パラメータを用いて回答者の特性尺度値を推定することで、オリジナル尺度の測定次元を維持したまま短縮版を作成することも可能となる。そこで、本研究では IRT を適用して Big Five 尺度の特徴を検討し、短縮版の開発を行うことを目的とする。

### 4.3. 研究1 短縮版の作成

#### 4.3.1. 目的

和田（1996）の Big Five 尺度に関して、IRT を適用した分析を行い、短縮版の作成を試みることを目的とする。

#### 4.3.2. 方法

**調査協力者** 高校生および大学生、専門学校生 2099 名（男性 906 名、女性 1165 名、不明 28 名）を対象とした。年齢は平均 19.10 歳（ $SD=3.19$ ）であった。

**質問項目** 和田（1996）の Big Five 尺度を用いた。回答選択枝は、“非常にあてはまる” “かなりあてはまる” “ややあてはまる” “どちらともいえない” “あまりあてはまらない” “ほとんどあてはまらない” “まったくあてはまらない” の 7 件法であった。

**分析方法** まず IRT の前提である 1 次元性の確認を行った。また IRT モデルは、Generalized Partial Credit Model（Muraki, 1992）を用い、項目パラメタの推定には PARSCALE 4.1（Muraki & Bock, 2003）を用いた。なお、Generalized Partial Credit Model は、尺度で一貫した category パラメタを推定でき、これはリッカート形式のデータに適していると考えられる。しかし、脇田（2004）でも指摘されたように、逆転項目と非逆転項目ではカテゴリ間の心理的距離が異なる。そのため、本研究では非逆転項目と逆転項目それぞれで category パラメタを推定するモデルを選択した。

**項目選択** 短縮版作成においては、以下の点を基準に各下位尺度の項目数の決定を行った。まず、短縮版では、項目数の減少に伴い、オリジナルと比べて測定精度が低下することは予想される。しかし、いずれの下位尺度においても一定程度の信頼性を確保する必要がある。IRT のテスト情報量  $I(\theta)$  において、クロンバックの  $\alpha=.7$  に相当するのが 3.33 であり、クロンバックの  $\alpha=.8$  に相当するのが 5.00 である。そこで、本研究では特性尺度値  $\theta$  が -2.00 から 2.00 の範囲における平均テスト情報量 ( $I_m$ ) を求めて、それがこの基準を超えることを条件として項目数を決定した。 $I_m$  は、以下の式で求めた。

$$I_m = \int I(\theta) \cdot g(\theta) d\theta$$

ただし、 $g(\theta)$  は標準正規分布とする。

Table 1  
各下位尺度の項目パラメタ推定値

	項目	slope	(SE)	location	(SE)		c1	c2	c3	c4
N	悩みがち	0.97	(0.04)	-0.74	(0.03)	非逆転	0.61	0.26	0.07	-0.94
	a) 不安になりやすい	1.44	(0.06)	-0.68	(0.03)	(SE)	(0.02)	(0.02)	(0.02)	(0.01)
	a) 心配性	0.97	(0.04)	-0.84	(0.03)					
	気苦労の多い	0.48	(0.02)	-0.47	(0.04)					
	a) 弱気になる	0.85	(0.03)	-0.42	(0.03)					
	傷つきやすい	0.65	(0.03)	-0.70	(0.03)					
	動揺しやすい	0.55	(0.02)	-0.71	(0.04)					
	神経質な	0.36	(0.02)	-0.20	(0.04)					
	悲観的な	0.41	(0.02)	0.06	(0.04)					
	a) 緊張しやすい	0.33	(0.02)	-1.50	(0.07)					
	a) 憂鬱な	0.42	(0.02)	0.08	(0.04)					
	くよくよしない	0.49	(0.02)	-0.25	(0.04)	逆転	0.85	0.24	0.22	-1.3
						(SE)	(0.10)	(0.09)	(0.08)	(0.08)
E	a) 話好き	0.53	(0.02)	-1.25	(0.04)	非逆転	1.18	0.26	-0.26	-1.18
	a) 陽気な	0.75	(0.03)	-0.76	(0.03)	(SE)	(0.03)	(0.02)	(0.02)	(0.02)
	a) 外向的	1.05	(0.04)	-0.22	(0.03)					
	a) 社交的	1.08	(0.04)	-0.32	(0.03)					
	活動的な	0.64	(0.02)	-0.49	(0.03)					
	積極的な	0.62	(0.02)	-0.31	(0.03)					
	a) 無口な	0.62	(0.02)	-0.51	(0.03)	逆転	1.18	0.24	-0.70	-0.72
	暗い	0.65	(0.02)	-0.66	(0.03)	(SE)	(0.05)	(0.04)	(0.03)	(0.03)
	無愛想な	0.41	(0.02)	-0.52	(0.04)					
	人嫌い	0.39	(0.02)	-0.89	(0.04)					
	意思表示しない	0.30	(0.01)	-0.51	(0.05)					
	地味な	0.36	(0.02)	-0.02	(0.04)					
O	a) 独創的な	0.40	(0.02)	-0.47	(0.04)	非逆転	1.27	0.70	-0.67	-1.30
	a) 多才の	0.66	(0.03)	0.43	(0.03)	(SE)	(0.03)	(0.03)	(0.03)	(0.03)
	a) 進歩的	0.67	(0.03)	-0.24	(0.03)					
	洞察力のある	0.38	(0.02)	-0.55	(0.04)					
	想像力に富んだ	0.37	(0.02)	-0.68	(0.04)					
	美的感覚の鋭い	0.32	(0.01)	0.14	(0.04)					
	a) 頭の回転の速い	0.45	(0.02)	0.03	(0.04)					
	臨機応変な	0.45	(0.02)	-0.32	(0.04)					
	a) 興味の広い	0.29	(0.02)	-1.22	(0.07)					
	a) 好奇心が強い	0.34	(0.02)	-1.62	(0.07)					
	独立した	0.27	(0.01)	0.05	(0.05)					
	呑み込みの速い	0.33	(0.02)	-0.14	(0.04)					



Table 1

(つづき)

		slope	(SE)	location	(SE)		c1	c2	c3	c4
A	a) 温和な	0.36	(0.02)	-1.36	(0.05)	category	1.65	1.17	-0.63	-2.19
	a) 寛大な	0.31	(0.01)	-0.81	(0.05)	(SE)	(0.11)	(0.07)	(0.05)	(0.05)
	a) 親切的な	0.34	(0.02)	-1.20	(0.07)					
	良心的な	0.34	(0.02)	-1.34	(0.06)					
	協力的な	0.21	(0.01)	-1.37	(0.09)					
	素直な	0.17	(0.01)	-1.36	(0.10)					
	a) 短気	1.00	(0.05)	0.01	(0.03)	逆転	1.09	0.10	-0.48	-0.70
	a) 怒りっぽい	1.00	(0.05)	-0.03	(0.03)	(SE)	(0.03)	(0.03)	(0.03)	(0.03)
	とげがある	0.33	(0.02)	-0.12	(0.04)					
	かんしゃくもち	0.42	(0.02)	-0.51	(0.04)					
C	a) 自己中心的	0.34	(0.02)	0.22	(0.04)					
	反抗的	0.43	(0.02)	-0.05	(0.04)					
	a) 計画性のある	0.27	(0.02)	0.16	(0.05)	非逆転	0.95	0.51	-0.14	-1.32
	勤勉な	0.20	(0.01)	0.47	(0.07)	(SE)	(0.10)	(0.10)	(0.09)	(0.10)
	a) 几帳面な	0.23	(0.01)	-0.26	(0.05)					
	a) いい加減な	0.73	(0.04)	0.78	(0.03)	逆転	1.38	0.44	-0.76	-1.06
	a) ルーズな	0.39	(0.02)	0.39	(0.04)	(SE)	(0.03)	(0.03)	(0.03)	(0.04)
	a) 怠惰な	0.55	(0.02)	0.46	(0.04)					
	a) 成り行きまかせ	0.39	(0.02)	0.88	(0.05)					
	不精な	0.42	(0.02)	0.06	(0.04)					
	無頓着な	0.35	(0.02)	0.01	(0.05)					
	a) 軽率な	0.39	(0.02)	-0.17	(0.04)					
	無節操	0.35	(0.02)	-0.50	(0.05)					
	飽きっぽい	0.22	(0.01)	1.08	(0.08)					

## a) 短縮版選択項目

注) N : 情緒不安定性 (Neuroticism), E : 外向性 (Extraversion), O : 開放性 (Openness), A : 調和性 (Agreeableness), C : 誠実性 (Conscientiousness)

## 4.3.3. 結果

IRT モデルを適用する前提である尺度の 1 次元性の検討を行った。下位尺度ごとに因子分析(主因子法)を行ったところ, 第 1 固有値と第 2 固有値の比は情緒不安定性では 6.31, 外向性では 5.28, 開放性では 3.16, 調和性では 2.24, 誠実性では 3.25 であった。そのため, いずれの下位尺度においても一定の 1 次元性が満たされていると考えられた。そこで, 下位尺度ごとに IRT を適用した分析を行い, 項目パラメタの推定を行った。

その結果, 調和性の非逆転項目では, c2 と c3 の category パラメタに逆転が認められた。これは, 選択枝の順序性が崩れていることを示しており, 望ましいことではない。この結

果は、七つの選択枝が十分に機能していない可能性を示唆している。同様の現象が情緒不安定性においても認められた。そこで、本研究では、7件法の第1選択枝と第2選択枝、第6選択枝と第7選択枝を同一の選択枝として扱い、改めて5件法のデータとしてIRT分析を行った。Table 1に5件法の場合の項目パラメタを示した。5件法の場合、いずれの下位尺度においてもcategoryパラメタの逆転は生じていなかった。したがって、短縮版作成の際には5件法により推定された項目パラメタの情報をを用いるのが適当と考えられる。

次に、推定されたパラメタの結果などから、以下の手続きに従って短縮版に用いる項目の選定を行った。まず、slopeパラメタが高い、すなわち識別力が高いと思われる項目を短縮版に用いる候補として選択した。更に幅広い範囲の能力値の測定が可能になるように、既に選択された項目とはlocationパラメタができるだけ異なる項目を候補として選択した。

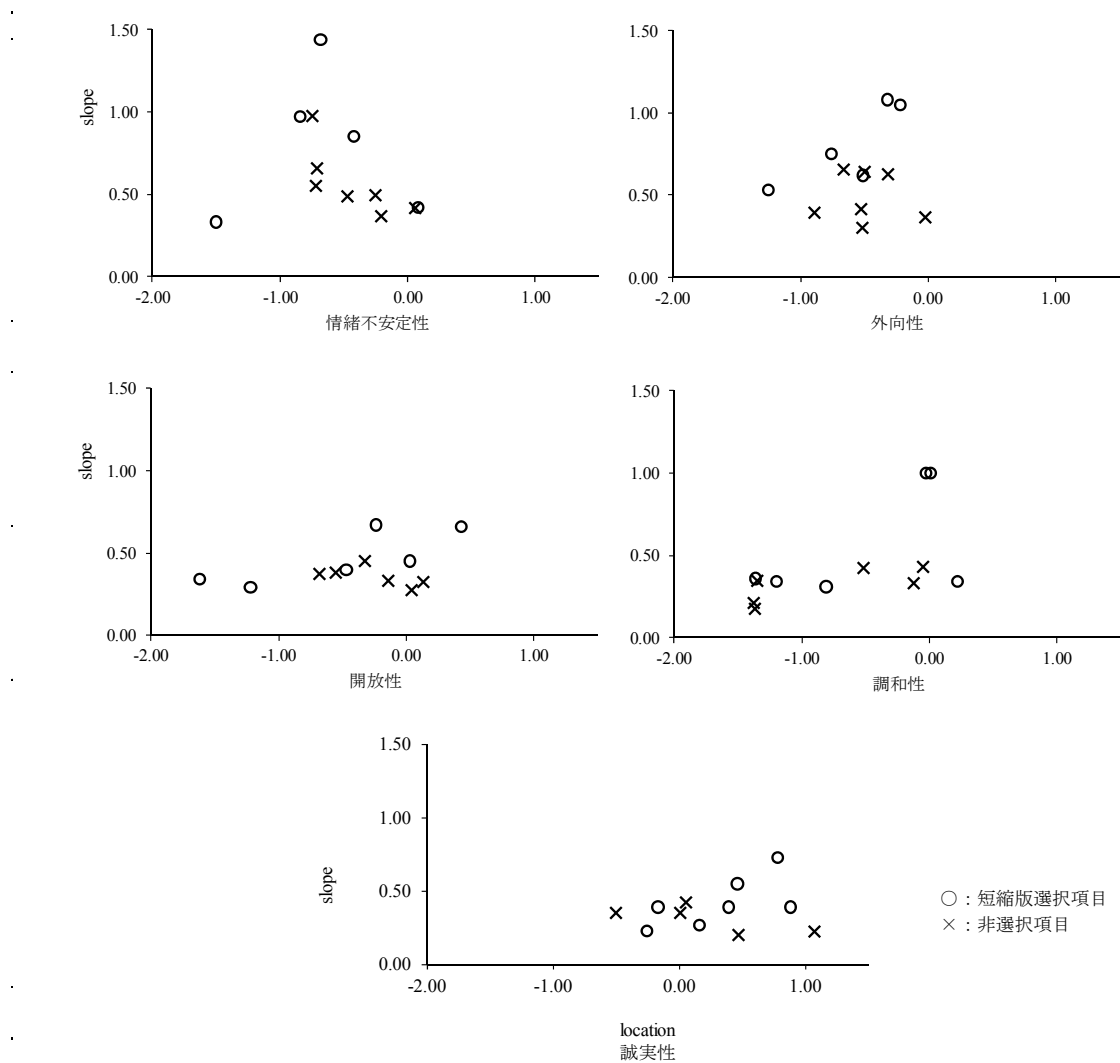


Figure 1 各下位尺度の項目パラメタ

その上で他の下位尺度の項目数とのバランスも考慮しながら、 $-2.0 < \theta < 2.0$  の範囲で、 $Im$  が 3.3 以上になるように項目数を決定した<sup>2</sup>。

その結果、各下位尺度の項目数は外向性と情緒不安定性が 5 項目、開放性と調和性は 6 項目、誠実性は 7 項目となった。選択された項目は Table1 の右列に a) を付ける形で示した。また、オリジナル版と、本研究で作成された短縮版の情報量を Figure2 に示した。

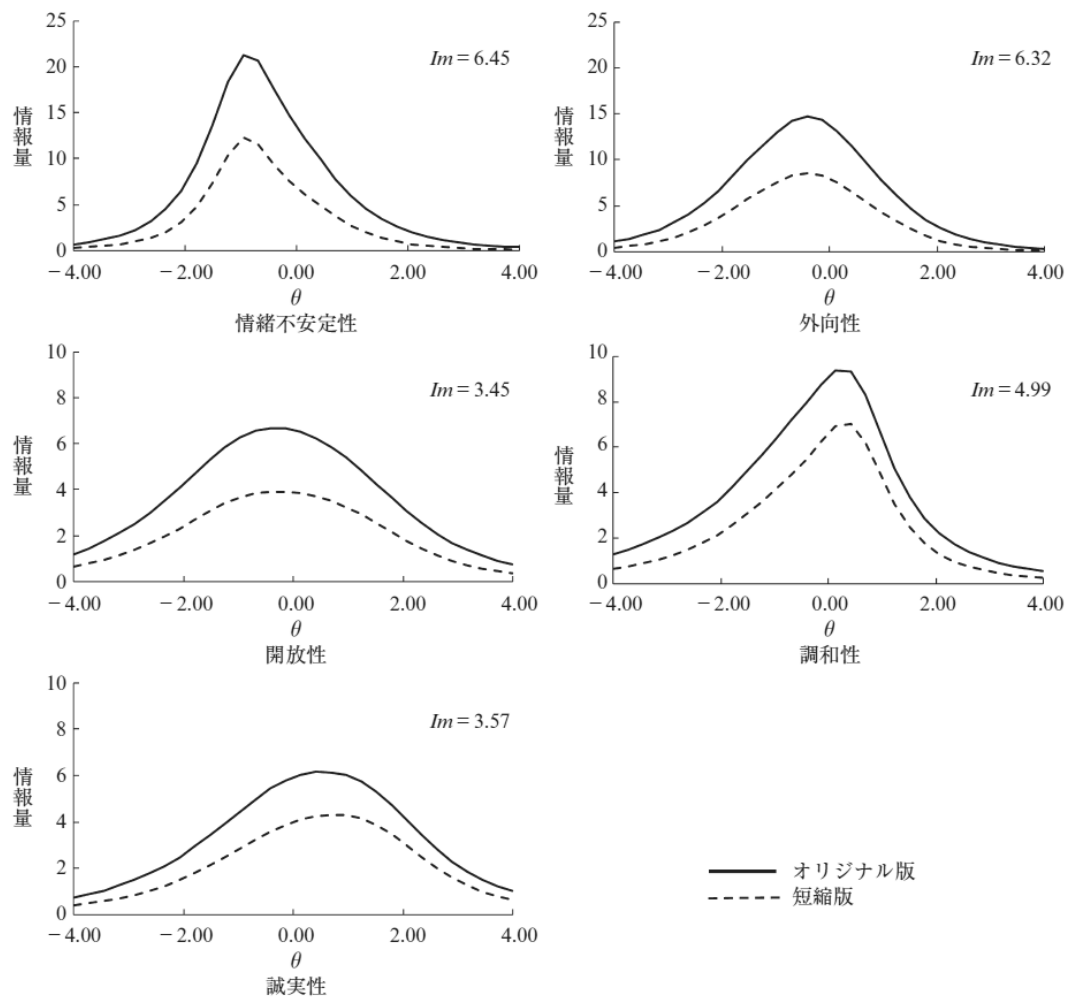


Figure 2. 各下位尺度のテスト情報量および短縮版の平均テスト情報量 ( $Im$ )

<sup>2</sup>外向性を例に挙げると、slope が高い項目としてまず「外向的」「社交的」が選択され、それらと location が大きく異なる (低い) 項目として「話好き」が選択された。さらに、残りの 9 項目のうち、slope が相対的に高く、location がすでに選択された 3 項目から離れた (中間に位置する) 項目として「陽気な」「無口な」の 2 項目を選択した。

Table2  
Big Five 短縮版の因子パターン行列

	I	II	III	IV	V
外向性 ( $\alpha=.86$ )					
無口な	-0.82	0.08	-0.06	-0.22	-0.07
社交的	0.74	0.02	0.00	0.09	0.09
話好き	0.73	-0.05	0.12	-0.03	-0.04
外向的	0.71	0.01	-0.06	0.15	-0.03
陽気な	0.68	-0.12	0.00	0.11	0.02
誠実性 ( $\alpha=.78$ )					
いい加減な	-0.02	0.80	0.03	0.01	-0.01
ルーズな	-0.01	0.66	-0.02	-0.05	0.00
成り行きまかせ	0.01	0.63	0.05	-0.01	-0.13
怠惰な	0.08	0.61	-0.12	0.01	0.05
計画性のある	-0.03	0.53	0.08	0.29	-0.04
軽率な	-0.01	0.47	-0.04	-0.10	0.10
几帳面な	-0.13	0.45	0.20	0.21	0.02
情緒不安定性 ( $\alpha=.82$ )					
不安になりやすい	0.06	0.04	0.84	-0.09	-0.02
心配性	0.12	0.14	0.83	-0.06	-0.02
弱気になる	-0.02	-0.08	0.75	-0.17	0.02
緊張しやすい	-0.01	0.00	0.52	-0.06	0.05
憂鬱な	-0.32	-0.07	0.51	0.09	-0.03
開放性 ( $\alpha=.76$ )					
多才の	-0.10	0.09	-0.07	0.74	0.03
進歩的	0.04	0.12	-0.07	0.63	0.03
独創的な	-0.09	-0.10	-0.02	0.59	-0.04
頭の回転の速い	-0.02	0.13	-0.10	0.52	-0.04
興味の広い	0.10	-0.02	-0.10	0.50	-0.02
好奇心が強い	0.17	-0.09	-0.03	0.46	-0.04
調和性 ( $\alpha=.78$ )					
短気	-0.08	0.07	-0.11	-0.16	0.77
怒りっぽい	-0.07	0.03	-0.15	-0.16	0.75
温和な	0.05	-0.13	0.18	0.13	-0.70
寛大な	-0.03	-0.17	0.04	0.31	-0.60
自己中心的	0.04	0.24	-0.04	-0.20	0.45
親切的な	0.19	0.07	0.20	0.24	-0.42
因子間相関					
	I	-0.01	-0.20	0.46	0.10
	II		-0.12	-0.02	0.20
	III			-0.03	-0.23
	IV				0.15

Table3  
オリジナル版と短縮版の各下位尺度間相関係数  
(右上：調査1，左下：調査2)

	オリジナル版					短縮版				
	N	E	O	A	C	N	E	O	A	C
オリジナル版										
N	—	-0.38	-0.16	-0.31	-0.05	<b>0.95</b>	-0.23	-0.16	-0.23	-0.01
E	-0.36	—	0.41	0.32	0.16	-0.35	<b>0.93</b>	0.43	0.13	0.09
O	-0.26	0.44	—	0.19	0.11	-0.21	0.38	<b>0.94</b>	0.16	0.11
A	-0.19	0.21	0.09	—	0.31	-0.22	0.17	0.16	<b>0.94</b>	0.24
C	-0.09	0.11	-0.03	0.33	—	-0.1	0.04	0.09	0.25	<b>0.95</b>
短縮版										
N	<b>0.94</b>	-0.37	-0.26	-0.21	-0.17	—	-0.33	-0.19	-0.23	-0.04
E	-0.27	<b>0.92</b>	0.39	0.14	0.02	-0.28	—	0.41	0.08	0.02
O	-0.29	0.5	<b>0.94</b>	0.13	-0.03	-0.27	0.45	—	0.17	0.02
A	-0.17	0.1	0.1	<b>0.95</b>	0.3	-0.19	0.04	0.12	—	0.19
C	-0.04	0.02	-0.04	0.28	<b>0.95</b>	-0.12	-0.06	-0.06	0.27	—

次に、短縮版の項目として選択された 29 項目について因子分析（主因子法・プロマックス回転）を行った。その結果、固有値の減衰状況（5.21, 3.71, 2.83, 2.42, 1.86, 1.14, 0.89, 0.83, …）から 5 因子の状態が保たれていることが確認された。また、各項目は特定の因子に.40 以上の高い負荷を示し、複数の因子に.35 以上の負荷を示す項目は見られなかった。また、各因子の信頼性係数の推定値として、クロンバックの  $\alpha$  係数を算出したところ、.76 から.86 の値が得られた（Table2）。

また、各下位尺度について、オリジナル版の 12 項目で算出した得点と短縮版で算出した得点間の相関係数を算出した。その結果、今回作成された短縮版 5 件法はいずれの得点とも  $r=.93$  以上の相関が認められた（Table3）。

#### 4.3.4. 考察

本研究では、IRT モデルに基づく項目パラメタが求められた。調和性と誠実性、開放性の slope パラメタは全体的に低かったものの、短縮版を構成する上ではそれらの情報も有用であった。また、作成された短縮版は、いずれの下位尺度も一定程度の情報量が確保されているため、十分な精度での測定が可能であると考えられる。

また category パラメタに関する議論から、5 件法が望ましいことが示唆された。しかし、これは 5 件法で得られたデータによるものではなく 7 件法で得られたデータを変換したものであるため、この方法の妥当性については更なる検討が必要である。ただし、調査協力者の負担を軽減するという短縮版本来の目的を考慮すると、項目を減らすだけではなく、選択枝数を減らすことも考慮すべきだろう。

因子分析の結果からは、短縮版の項目でも 5 因子構造が確認され、因子的な妥当性は十分有していると考えられる。さらに、いずれの下位尺度も  $\alpha$  が.75 以上と十分な値を示し

ていることから、内的な一貫性も確認された。また、オリジナル版と短縮版との間の相関係数は、いずれも  $r=.93$  以上と高い値を示していた。以上のことから、今回作成された短縮版はオリジナル版と同程度に十分な信頼性や因子的妥当性を備えていることが示された。しかしながら、外在基準との関連などは十分に検討できていないことなどもあり、更なる検討が必要である。

#### 4.4. 研究2 短縮版の妥当性検討

##### 4.4.1. 目的

研究1において作成された Big Five 尺度短縮版の併存的妥当性について検討することを目的とする。本研究では、外在基準として同じ5因子モデルの尺度である NEO-FFI を取り上げ、Big-Five 尺度オリジナル版と短縮版で同様の相関が得られることを確認する。

##### 4.4.2. 方法

**調査協力者** 大学生 238 名 (男性 122 名, 女性 116 名), 年齢は平均 20.35 歳 ( $SD=3.44$ ) であった。

**質問紙構成** Big Five 尺度 (和田, 1996) のオリジナル版 60 項目および NEO-FFI (下仲他, 1999) の 60 項目を用いた。

##### 4.4.3. 結果

Big Five 尺度について、60 項目すべてを用いる形でオリジナル版の得点を算出するとともに、研究1で選択された 29 項目の部分のみを用いる形で短縮版の得点も算出した。オリジナル版と短縮版の相関を求めたところ、研究1と同様に同一の下位尺度間ではいずれも高い相関関係 ( $r_s=.92-.95$ ) を示していた (Table3)。

次に、NEO-FFI も含めた各下位尺度間の相関係数を求めた (Table4)。まず Big Five 尺度オリジナル版の情緒不安定性をみると、対応する NEO-FFI の神経症傾向と最も高い相関 ( $r=.73$ ) を示していた。また、外向性においても同様に NEO-FFI の外向性と最も高い .79 の相関を示していた。調和性と誠実性においても、それぞれ対応する尺度と .54, .52 の中程度の相関を示しており、もっとも高い値であった。開放性では対応する NEO-FFI の開放性との相関が .25 と相対的に低くなっていたが、NEO-FFI の開放性からみると、Big Five 尺度の中で最も高い相関を示していたのは開放性の .25 であった。

Table 4

Big Five 尺度と NEO-FFI の相関係数および  $\alpha$  係数

	Big Fiveオリジナル版					短縮版					$\alpha$
	N	E	O	A	C	N	E	O	A	C	
NEO-FFI											
N	<b>.73</b> **	-.29 **	-.20 **	-.27 **	-.20 **	<b>.70</b> **	-.23 **	-.23 **	-.27 **	-.15 *	.79
E	-.25 **	<b>.79</b> **	.43 **	.21 **	.05	-.26 **	<b>.74</b> **	.46 **	.11	-.02	.78
O	.01	.10	<b>.25</b> **	.07	-.16	.06	.08	<b>.21</b> **	.06	-.17 **	.56
A	-.19 **	.23 **	-.07	<b>.54</b> **	.13 *	-.17 **	.18 **	-.03	<b>.46</b> **	.07	.69
C	-.23 **	.27 **	.33 **	.16	<b>.52</b> **	-.24 **	.22 **	.33 **	.15 *	<b>.52</b> **	.60
$\alpha$	.87	.88	.84	.83	.80	.79	.84	.75	.77	.77	

\*\* $p < .01$ , \* $p < .05$ 

Big Five 尺度短縮版と NEO-FFI の相関をみると、対応する尺度間の相関は情緒不安定性で.70、外向性で.74 と、いずれもオリジナル版と同様の傾向を示していた。他の尺度でも同様にオリジナル版と近い値を示しており、短縮版とオリジナル版の相関係数の差の平均値は 0.04 ( $SD=0.03$ ) であった。

また、信頼性係数の推定値として各尺度の  $\alpha$  係数を求めた (Table4)。その結果、Big Five 尺度のオリジナル版ではいずれの下位尺度も.80 以上、短縮版でも.75 以上の値を示していた。

#### 4.4.4. 考察

本研究では Big Five 尺度オリジナル版および短縮版と NEO-FFI との相関を検討した。いずれも 5 因子モデルに基づく尺度であるため、対応する下位尺度の間には、すべて有意な相関が認められていた。また、研究 1 で作成された短縮版は、いずれの下位尺度においてもオリジナル版との相関が高く、NEO-FFI との相関も同様の値を示していた。そのため、この短縮版は併存的妥当性の観点からもオリジナル版と同様の測定機能を有していると考えられる。さらに、いずれの尺度も  $\alpha$  係数が.75 以上と十分な値を示していたことから、信頼性についてもあらためて確認された。

なお、Big Five 尺度の開放性については、オリジナル版においても短縮版においても NEO-FFI の開放性との相関が相対的に低いことが示されていた。開放性の因子については、同じ 5 因子モデルの尺度間でも相関が低い場合もあることが報告されており、例えば大野木 (2004) でも NEO-PI-R の開放性と主要 5 因子性格検査 (村上・村上, 1999) の知性との相関が.29 と低く、知性は神経症傾向や誠実性との相関が高いことが示されている。本研究では Big Five 尺度の開放性と相関が高いのは、外向性や誠実性であり、大野木 (2004) の結果とは必ずしも一致しないが、この点については和田 (1996) の Big Five 尺度や今回

作成された短縮版だけの問題でなく、大野木（2004）も指摘するように5因子モデル全体での総合的な議論が、今後必要な点である。

また、本研究ではオリジナル版と短縮版の相関を求めるため、オリジナル版の60項目版のみを実施し、そのうちの一部を用いて短縮版の得点を算出して分析を行った。そのため、同一のデータを基に相関を計算しているという問題は残る。今後は短縮版29項目を単独で実施しての検討なども必要である。

## 4.5. 総合考察

本研究では、IRTを適用した分析結果に基づいて、Big Five 尺度の短縮版の作成を試みた。作成された短縮版は外向性と情緒不安定性が5項目、開放性と調和性は6項目、誠実性は7項目の計29項目であり、オリジナル版の60項目と比べて、項目数が半分未満になった。これは回答時間の短縮や、回答者の負担の軽減に大きく寄与し、データの質の向上にもつながると考えられる。

また、短縮版の因子構造、信頼性係数、および外在基準との相関係数から、オリジナル版と遜色のない短縮版であることも確認された。情報量などを考慮すると、より精緻な測定を目指すにはオリジナル版を用いる方が望ましいのは変わらないが、今回作成された短縮版もテストバッテリーや回答者の状況等を勘案し、必要に応じてオリジナル版に代えて用いることができると考えられる。

なお、今回は一般的なパーソナリティを測定する尺度であったため、短縮版の項目選択においては幅広い対象に対して一定程度のテスト情報量が確保されることを基準としていた。特にlocationパラメタは、特定の数値に偏らないように配慮し、項目の選択を行った。しかしながら、IRTを用いて短縮版を作成する場合、一定の特性値周辺の情報量が高くなるように項目を選択することも可能であり、項目選択の基準は多様である。そのため、今後はさまざまな尺度における分析を通して、短縮版作成の際の手続きや項目選択の基準等について検討を重ねて行く必要がある。



## 第5章 継時的比較志向性尺度の作成と短縮版に関する検討

## 5.1. はじめに

第3章および第4章では既存の尺度を用いて、心理尺度短縮版の作成方法に関する検討を行ってきた。本章では、新たに一つの尺度を作成した上で、前章までと同様に短縮版の作成を行うことで、尺度の作成から短縮化までのプロセス全体を検討する。

検討に用いる尺度は、継時的比較に関するものであり、社会心理学（もしくはパーソナリティ心理学）領域に関連する尺度であり、具体的には、人が継時的比較をどの程度行うかの個人差を測定する尺度である。継時的比較自体は、次節でも詳しく述べるが、まだ発展途上のテーマであるものの、継時的比較や社会的比較といったテーマを扱う社会心理学は短縮版尺度のニーズの高い領域の一つであると言える。たとえば並川・脇田・野口（2010）ではRosenberg（1965）の自尊感情尺度の使用状況をまとめているが、10項目のRosenbergの自尊感情尺度であっても短縮化して利用されているケースとして社会心理学領域の研究が複数報告されている。これは、社会心理学領域では大規模な社会調査や、複数の変数を同時に扱う調査が行われる場合も多く、それらの場合はできるだけ少ない項目で測定が可能であることが望ましいためであると考えられる。そのため、ここでは継時的比較に関する尺度を取り上げることとした。また、次章で追加の分析を行うことも考慮し、前章までの尺度と異なり1次元の尺度を取り上げる。

## 5.2. 研究1 継時的比較志向性尺度の作成

自分自身の能力やパーソナリティなどを評価する際に、人はさまざまな対象と比較を行う。たとえば自己と身近な他者とを比べることによって、自己の能力が高いと言えるか否かを判断することがある。これは社会的比較（Festinger, 1954）と呼ばれ、社会心理学の領域を中心に多くの研究がなされてきた現象である。また、このような他者との比較だけでなく、人は異なる時点における自己間の比較を行うこともある。たとえば、“過去の自己”と“現在の自己”とを比べることによって、過去よりも成長した自己を認識したり、一貫した自己の姿を確認したりする。この“過去の自己”と“現在の自己”のように異なる時点における自己間の比較のことは継時的比較と呼ばれている（Albert, 1977）。これらの比較は、いずれも自己をとらえていく上で重要な役割を担っている（Suls & Wheeler, 2000；高田, 1992）。

### 社会的比較志向性に関する研究

しかしながら、これらの比較をどの程度行うのかには個人差があり、必ずしも全ての人がそれを積極的に行っているわけではない（Bunnk & Gibbons, 2006）。社会的比較に関

する情報を多く収集し、他者との比較を積極的に行おうとする人もいれば、あまりそれを意識しない人も存在する。

近年、社会的比較におけるこの個人差、すなわち社会的比較志向性 (Social comparison orientation) に注目する研究がいくつかみられる。たとえば, Gibbons & Buunk (1999) は、社会的比較志向性を測定する尺度 (Iowa-Netherlands comparison orientation measure ; 以下 INCOM とする) を作成し、その可能性について論じている。

INCOM は, Gibbons & Buunk (1999) によると社会的比較の情報への注意 (Lennox & Wolfe, 1984) やセルフモニタリング (Snyder, 1974), Big-Five (Costa & McCrae, 1992), など多くの変数と、必ずしも強くはないものの有意な相関がみられることが指摘されている。また, Buunk & Gibbons (2006) は、社会的比較志向性の高さに関連する特徴として、(a) 自己への意識の活性、(b) 他者への関心の高さ (c) ネガティブな感情の高さや自己評価の低さの3点を挙げている。実際, Gibbons & Buunk (1999) でもその指摘に対応する形で、(a) 公的自己意識との正の相関 ( $r = .38 - .49$ ) や、(b) 対人志向性との正の相関 ( $r = .45$ ) がみられているほか、強くはないものの (c) Beck の抑うつ尺度 (Beck, 1967) や CES-D (Radloff, 1977) と有意な正の相関 ( $r = .13 - .25$ )、自尊感情とも有意な負の相関 ( $r = -.09 - -.23$ ) などがみられている。

また、この社会的比較志向性の高低は、比較によって生じる感情や評価にも影響を与えていることも示唆されている。例えば、社会的比較志向性が高い人は、自分より優れた他者との比較である上方比較によってネガティブな感情を、自分よりも劣った他者との比較である下方比較によってポジティブな感情をそれぞれ強く経験する傾向があることが報告されている (Buunk, Zurriaga, Peiro, Nauta, & Gosalvez, 2005)。

さらに、INCOM を用いた他の研究からも社会的比較志向性の高さが、リスク認知や危険行動 (Gibbons & Gerrard, 1995; Gibbons, Lane, Gerrard, Pomery, & Lautrup, 2002)、関係性や集団の満足度 (Buunk, Oldersma, & de Dreu, 2001 ; Buunk, Nauta, & Molleman, 2005)、バーンアウト (Buunk, Ybema, Gibbons, & Ipenburg, 2001)、大学生活の不安 (友寄・佐藤, 2006)、学業的援助要請 (野崎, 2006) など様々な変数と関連していることが報告されている。たとえば Buunk, Ybema, Gibbons et al. (2001) は、社会的比較志向性の高い人では、バーンアウトの傾向が高い人ほど下方比較によってネガティブな感情が喚起される傾向が見られるが、社会的比較志向性の低い人ではそのような傾向は見られないことを指摘している。また, Buunk, Oldersma et al. (2001) は、社会的比較志向性が高い場合に、下方比較を行うことが関係性における不満を和らげることを示している。

これらの知見は、比較のプロセスや比較による自己評価をとらえる上で、社会的比較志向性が重要な観点となりうることを示唆するものである。社会的比較志向性は必ずしも直接的に他の変数に影響を与えるわけではないが、 Buunk, Zurriaga et al. (2005) など

も示唆されるように、社会的比較を行うことによって生じる感情など、比較過程について詳細な検討をしていく上で特に重要な指標となると考えられる。また、Wheeler (2000) でも INCOM は新しい尺度として社会的比較過程の研究における重要性が指摘されている。

### 継時的比較と継時的比較志向性

社会的比較志向性についてはこのように実証的な研究が重ねられている一方で、継時的比較における個人差については、まだ実証的な研究が重ねられてはいない。たとえば、継時的比較が用いられる程度の集団差、特に発達のな差異については、Suls & Mullen (1982) などでも社会的比較と併せて議論されており、日本でも高田 (1993) などでも検討がなされている。また、過去を振り返るという点で共通する回想研究では、長田・長田 (1994) や野村・橋本 (2001) などでも回想の頻度を測定する形で検討されている。しかしながら、継時的比較において、社会的比較志向性のような個人差を扱う実証的研究は、尺度の開発を含めて確認されていない。

しかし、社会的比較や回想において個人差が存在するように、継時的比較においても個人差は存在すると考えられる。すなわち、継時的比較に関する情報を多く集め、継時的比較を多く行おうとする人もいれば、逆に継時的比較をまったく意識しない人も存在するのではないだろうか。この個人差は、やはり社会的比較志向性と同様に、比較の結果生じる自己評価や感情に少なからぬ影響を与えていると考えられる。したがって、比較過程の詳細を検討する際に、継時的比較志向性は有用な指標となると考えられる。たとえば、継時的比較を行うことによって感情がどのような影響を受けるのかを検討する際には、これらの志向性が重要な役割を担っていると推測される。

また、継時的比較は、Wilson & Ross (2000) も指摘しているように、自己をポジティブにとらえようとする際には社会的比較よりも多く用いられるなど、自己高揚とも密接に関連しているものであると考えられる。そのため、継時的比較について、継時的比較志向性を含めて詳細に検討していくことは、単に比較の過程に関する現象をとらえる上で重要なだけでなく、自己評価の全体像をとらえる上でも大きな役割を果たすのではないだろうか。

さらに“過去”をとらえるという点からも継時的比較志向性は重要なものとなると考えられる。従来から、過去を想起することや過去をどのようにとらえるかは、さまざまな形で自己評価や精神的健康と関連していることが示唆されている。例えば、高齢者の回想法に関する研究では、過去を振り返ることが適応的な効果をもたらす場合があることが示されていたり (野村, 1998)、過去の他者との肯定的なエピソードを想起することが、その他者との関係性の評価にも肯定的に作用することなども示唆されている (Alea & Bluck, 2007)。こういった“過去”を扱う研究は、これまで回想研究や自伝的記憶研究、時間的展望研究等の領域で主に行われてきた。しかし、本研究で扱うような継時的比較という視点

でこの“過去”に対してアプローチすることは、従来とは異なった視点から“過去”をとらえることにつながる。また、比較過程としてとらえることは、これまで社会的比較研究で提案されてきたさまざまな理論やモデルを援用することにもつながり、“過去”についての新たな理論の構築に寄与できる可能性がある。

たとえば Wilson & Ross (2001) や Ross & Wilson (2002) では、社会的比較における自己評価維持モデル (Tesser, 1988) における他者との親密度を“過去の自己”との主観的な時間距離に置き換えるなどの形で継時的自己評価理論が提案されている。そこでは、人には自己高揚動機が存在するということを前提に、継時的比較において比較対象となる“過去の自己”が、主観的な時間距離などに関連し、“現在の自己”を相対的に高く評価できるように用いられていると指摘されている。例えば、良くない“過去の自己”との比較（下方比較）においては、その良くない“過去の自己”を時間的に遠くに感じることによって、より成長などを実感でき、“現在の自己”にとってポジティブに作用するのに対し、良くない“過去の自己”を近くに感じるとそれが現在にも持続しているものとなり、“現在の自己”にとってネガティブに作用する。このような理論をもとに、継時的比較過程について検討していくことは自己評価のメカニズムを考える上でも重要な位置を占めるものと考えられる。しかしながら、Wilson & Ross (2001) や Ross & Wilson (2002) などの知見は、日本などでは必ずしも再現されないという指摘もなされている（工藤・遠藤, 2007; 佐藤, 2008）。しかし、これらの議論においても継時的比較志向性といった個人差を加えることによって、より詳細な検討が可能になると考えられる。

そこで本研究では、継時的比較志向性を定量的にとらえていくために、継時的比較志向性尺度を作成することを主な目的とする。まず予備調査において、項目収集を行い尺度の予備的検討を行う。また調査 1 ではその結果を踏まえ、継時的比較志向性尺度の信頼性、妥当性の確認を行う。さらに調査 2 においては、尺度の適用可能性について検討するために、測定された継時的比較志向性が比較によって生じる感情にどのような影響を与えるかについて検討を行う。

なお、社会的比較志向性を測定する INCOM では、意見比較と能力比較の 2 因子が見出されている。これらの因子は Festinger (1954) が社会的比較について論じる中で“人には意見や能力について評価する動機が存在する”などの形で指摘した二つの側面に対応しているものであり、意見や能力について他者との比較を志向する程度を示すものである。しかし、Gibbons & Buunk (1999) では、両因子の相関も.61 と高いことなどから、得点算出においてはこれら異なる因子の得点を合計し、1 因子の尺度として分析を行っている。しかしながら、異なる下位尺度の得点を合計することは、測定している概念を曖昧にし、下位尺度ごとに本来測定されていた特徴を打ち消してしまう恐れがあると考えられる。また“意見比較”と“能力比較”の両者はその生起プロセスも異なる現象であると考えられることや、“能力比較”を社会的比較志向性の概念として捉えた方が妥当ではないかという

指摘も存在している（外山，2002）。そこで，ここでは“能力比較”に主に焦点を当てることによって，1因子構造の概念として，継時的比較志向性をとらえることを試みる。

### 5.2.1. 予備調査

#### 5.2.1.1. 目的

継時的比較志向性の個人差を測定するための尺度作成における項目収集と尺度の予備的検討を行うことを目的とする。

#### 5.2.1.2. 方法

**調査協力者** 大学生 205 名（男性 63 名，女性 140 名，不明 2 名），平均年齢は 18.73 歳（ $SD=1.21$ ）であった。

**質問紙構成** 本調査で新たに作成された継時的比較志向性尺度 12 項目を使用した。項目は以下の手続きで作成された。まず INCOM (Gibbons & Bunck, 1999 ; 外山, 2002) の，“能力比較”因子の項目内容をもとに，“他者”を“過去の自己”に変更する形で 7 項目を作成した（e.g. “物事がうまく進んでいるかどうか評価する時には，昔の状況と比べてみることが多い”）。また，項目の補充や修正の参考にすることを目的に，大学生 52 名（男性 29 名，女性 23 名）を対象にした調査を行った。調査では，最近 1 ヶ月間に過去の自分との比較を行ったことがあるかどうか，あればどういった場面でどういった内容について比較を行ったのか，思いつくものをできるだけ多く記述するように教示した。得られた記述は，類似するものをまとめて行く形でグループ化され，整理された。その結果から，継時的比較においては過去の自分と比べて“成長したかどうか”や“変わったかどうか”といった表現が多くみられたため，これらを参考に 5 項目を作成し，追加した（e.g. “以前よりも自分が成長しているかどうか考えることが多い”）。

なお，尺度への回答は“とてもあてはまる”“ややあてはまる”“どちらとも言えない”“あまりあてはまらない”“全くあてはまらない”の 5 件法を用いた。

#### 5.2.1.3. 結果と考察

はじめに，継時的比較志向性尺度について各項目の平均値を算出したところ，1 項目だけ平均値が 4.52（ $SD=0.71$ ）と高い値を示す項目が見られた。当該項目は，尺度全体との相関も.20 を下回っていたため不適切であると判断し，その項目を除いた 11 項目を用いてその後の分析を進めた。各項目の平均値と標準偏差，I-T 相関などは Table1 に示した。

Table1  
継時的比較志向性尺度の項目分析結果

Item	<i>M</i>	<i>SD</i>	I-T相関	因子負荷
1	3.81	0.99	.71	.77
2	2.75	1.21	.64	.66
3	3.80	1.11	.55	.60
4	3.41	1.15	.73	.78
5	2.77	1.15	.51	.54
6	2.97	1.16	.57	.59
7	3.60	0.96	.47	.51
8	3.41	1.15	.74	.79
9	3.61	1.05	.69	.74
10	2.61	1.07	.48	.50
11	2.09	1.00	.69	.74

$\alpha = .89$

次に、因子分析（主因子法）を行い、尺度の因子構造の検討を行った。その結果、固有値の減衰状況を見ると第 1 固有値から順に 5.44, 1.11, 0.85, 0.75（以下省略）であり、1 因子構造であると判断された<sup>3</sup>。因子負荷量もすべての項目で.50 を超えており、1 次元性の強い尺度であると判断された。また、信頼性係数の推定値として  $\alpha$  係数を算出したところ.89 と高い値が得られた。

## 5.2.2. 調査 1

### 5.2.2.1. 目的

予備調査で作成された継時的比較志向性尺度について、より詳細に信頼性、妥当性の検討を行うことを目的とする。妥当性の検討では、社会的比較志向性（Bunnk & Gibbons, 2006）において指摘されている（a）自己への意識の活性、（b）他者への関心の高さ（c）ネガティブな感情の高さの 3 点を参考に、（a）公的・私的自己意識、（b）過去に対する評価傾向、（c）自己に対するネガティブな評価や感情、の 3 点との関連が見られるかを基準とする。継時的比較を多く志向することは、自己に対する意識が高いことと関連しているとともに、比較対象となる過去に対しても関心を持ち、多くの評価を行うこととも関連していると考えられる。また、比較への志向性は自己の不確かさと密接に関連していると考えられ、自己の不確かさと関連する低い自己評価や自己に対するネガティブな感情とも関連していると考えられる。そのため、INCOM 同様に、これらの変数とは強くはないもの

<sup>3</sup> なお、念のため 2 因子構造を仮定し、プロマックス回転を行ったが解釈可能な因子は抽出されなかった。

Table2  
各尺度の基本統計量

	<i>N</i>	<i>M</i>	<i>SD</i>
継時的比較志向性	475	36.92	8.89
私的自己意識	252	30.71	5.06
公的自己意識	254	32.31	5.68
否定的回想	269	21.44	5.1
肯定的回想	267	42.93	10.45
再評価傾向	268	42.41	8.75
抑うつ	261	43.73	7.07
自尊感情	397	28.98	7.06
社会的比較志向性	303	34.78	7.11

の有意な相関が見られると考えられる。

さらに、(d) 社会的比較志向性との関連も検討する。Wilson & Ross (2000) でも指摘されているように、継時的比較と社会的比較は、用いられる際の動機なども異なると考えられるため、それらに対する志向性も完全には一致しないと考えられる。しかし、(現在の)自己と異なるものと比較を行うという点では社会的比較も継時的比較も同じであるため、両者の間には有意な相関が見られると考えられる。

#### 5.2.2.2. 方法

**調査協力者** 大学生および専門学校生 481 名 (男性 252 名, 女性 229 名), 平均年齢は 19.81 歳 ( $SD=1.73$ ) であった。

**質問紙構成** (a) 公的自己意識・私的自己意識に関しては、押見・渡辺・石川 (1985) による自己意識尺度から、公的自己意識および私的自己意識の 2 つの下位尺度を用いた。

(b) 過去に対する評価傾向は、ネガティブな過去の出来事を再評価する傾向を測定する再評価傾向尺度 (野村・橋本, 2001) を用いた。また、回想される過去の評価も探索的に検討するため、回想がポジティブな感情や認知を伴うかどうかを測定する肯定的回想尺度および、回想がネガティブな感情や認知を伴うかどうかを測定する否定的回想尺度 (いずれも野村・橋本, 2001) も併せて用いた。(c) 自己に対する感情や評価に関しては、抑うつ傾向と自尊感情の 2 点を取り上げた。抑うつについては SDS (Zung, 1965; 福田・小林, 1973) を、自尊感情については Rosenberg (1965) による自尊感情尺度の邦訳版 (山本・松井・山成, 1982) を用いた。(d) 社会的比較志向性は、Gibbons & Buunk (1999) による INCOM の邦訳版 (外山, 2002) を用いた。(e) 継時的比較志向性については、予備調査で新たに作成された 11 項目を用いた。回答は予備調査と同様に 5 件法を用いた。

**手続き** 質問紙を授業時に一斉に配布し、回収する形で実施した。ただし、調査協力者の負担を軽減するため、項目数を勘案のうえ、上記の尺度から 3 尺度もしくは 4 尺度を含む



Table3  
継時的比較志向性尺度の因子分析結果

	F1
1 今の自分と昔の自分を比べることがよくある	.77
2 何かをするときに、昔の自分だったらどうしたかを考えることが多い	.63
3 自分の過去をよく思い出す	.62
4 今の自分がどんな性格であるかを、昔の自分と比べて考えることが多い	.73
5 昔の自分について考えることはほとんどない <sup>a)</sup>	.66
6 私は、昔の自分だったら同じ状況でどうするのかなどと考えたりすることはない <sup>a)</sup>	.51
7 以前よりも自分が成長しているかどうか考えることが多い	.49
8 昔の自分の状況と、今の自分の状況を比べることがよくある	.73
9 自分が、昔と変わったかどうかを考えることがよくある	.74
10 物事がうまく進んでいるかどうか評価する時には、昔の状況と比べてみるが多い	.45
11 今の自分の境遇と、昔の自分の境遇との違いを考えることはほとんどない <sup>a)</sup>	.65

<sup>a)</sup> 逆転項目

冊子を6種類作成し、ランダムに配布した。そのため尺度によって回答した人数が異なっていた。なお、各尺度の最終的な有効回答者数は、Table2に示した。

### 5.2.2.3. 結果と考察

まず継時的比較志向性尺度11項目について、主因子法による因子分析を行ったところ、固有値の減衰状況が第1固有値から順に5.11, 1.07, 0.93, 0.78であったため、1因子構造であると判断された。因子分析の結果はTable3に示した<sup>4</sup>。さらに、信頼性係数の推定値としてクロンバックの $\alpha$ 係数を算出したところ.88と高い値が得られた。これらのことから、継時的比較志向性尺度は、1次元性も高く、内的に一貫した尺度であることが示され、十分な信頼性を備えていると判断された。そこで、継時的比較尺度11項目の得点を合計することで、継時的比較志向性の得点とした。他の尺度についても同様に合計を求めることで尺度得点を算出した。各尺度の平均値、標準偏差はTable2に示した。継時的比較志向性の平均値は、36.92 ( $SD=8.89$ )であった。また、性差についても検討したところ、男性( $M=35.59$ ,  $SD=9.00$ )の方が女性( $M=38.38$ ,  $SD=8.56$ )に比べて有意に平均値が低いことが示された( $t(473)=3.45$ ,  $p<.05$ )。社会的比較志向性(Gibbons & Buunk, 1999)でも、同様に男性の方が有意に低いことが示されており、この傾向は社会的、

<sup>4</sup> 予備的に男女別でも因子分析を行ったところ、同様の1因子構造が確認されたため、以下では男女のデータを合わせて分析を行った。

継時的いずれにおいても共通することが示唆された。

次に、妥当性の検討のため、各尺度間の相関係数を算出した。その結果、まず公的自己意識尺度とは  $r=.42$  ( $p<.05$ )、私的自己意識尺度とは  $r=.43$  ( $p<.05$ ) の有意な正の相関が見られた。また、過去に関連する尺度との相関では、肯定的回想尺度 ( $r=.25$ ,  $p<.05$ )、否定的回想尺度 ( $r=.22$ ,  $p<.05$ )、再評価傾向尺度 ( $r=.19$ ,  $p<.05$ ) のいずれとも、弱いものの有意な正の相関が見られた。さらに抑うつ尺度とは有意な正の相関 ( $r=.19$ ,  $p<.05$ )、自尊感情尺度とは有意な負の相関 ( $r=-.22$ ,  $p<.05$ ) が、社会的比較志向性尺度とは有意な正の相関 ( $r=.45$ ,  $p<.05$ ) が見られた。

これらは、いずれも仮説と一致した傾向を示すものであった。すなわち、継時的比較志向性は、自己に対する意識の高さや過去に対する関心、ネガティブな情動と関連があり、社会的比較志向性とも関連があった。そのため、継時的比較志向性尺度は一定の妥当性を有していると判断された。

しかしながら、再評価傾向尺度との相関係数は、有意ではあったものの、 $r=.19$  と低い値であった。この点に関しては、今回用いた再評価傾向尺度が、単に過去の自己について関心がありそれについて評価を行うことだけではなく、ネガティブな過去を回想しそれを再評価する傾向を捉えようとするものであったことが関連していると考えられる。継時的比較には必ずしもネガティブな過去を対象とする下方比較のみではなく、ポジティブな過去との比較を行う上方比較も含まれていると考えられる。そのため、継時的比較志向性の高い人の中には、再評価傾向尺度で測定されるようなネガティブな過去の出来事を意識するだけでなく、ポジティブな過去を対象とした比較を志向する人も含まれているはずである。そのため仮説通り両者に関連はあるものの、高い相関ではなかったと考えられる。この点に関しては、純粋に回想量を測る尺度を用いるなどして再度検討する必要がある。なお、抑うつ尺度や自尊感情尺度との相関も決して高くはなかった。しかし、これらは社会的比較志向性の場合 (Gibbons & Buunk, 1999) も同程度の相関係数 (例えば、アメリカのサンプルで  $rs=-.09$  -  $-.23$ ) が報告されており、今回の検討においては妥当な値であると判断される。

### 5.2.3. 調査2

#### 5.2.3.1. 目的

社会的比較志向性は、比較後の感情などに関連していることが示唆されていた (Buunk, Ybema, Van der Zee, Schaufeli, & Gibbons, 2001)。そのため継時的比較においても、継時的比較志向性の高さは比較後に感じる感情の強さなどに影響を与えているものと思われる。そこで継時的比較志向性の有用性について議論するために、継時的比較志向性の高さが比較を行った後の感情と関連しているか

どうかを探索的に検討する。

また、調査1では、継時的比較志向性と過去を再評価する傾向との相関が必ずしも高くなかった。そこで、本調査では回想量および上方比較、下方比較の頻度との関連を検討することによって、測定された継時的比較志向性が、比較対象となる過去への関心と関連していることを確認する。

### 5.2.3.2. 方法

**調査協力者** 大学生 200 名（男性 72 名，女性 128 名），平均年齢は 18.95 歳（ $SD=1.06$ ）であった。

**質問紙構成** (a) 継時的比較志向性は、調査1と同一の継時的比較志向性尺度を用いた。(b) 回想量については、長田・長田（1994）による回想尺度を用いた。尺度は8項目からなり、“よく考える”から“考えない”の4件法で尋ねた。(c) 比較の頻度では、継時的な上方比較と下方比較をどの程度行うかについて、“今より良かった（悪かった）頃の‘過去の自分’と、‘今の自分’とを比べて考えること”がどの程度あるか“まったくない”から“よくある”までの5件法で尋ねた。(d) 比較後の感情として、継時的な上方比較や下方比較をした後に、どのくらいポジティブな感情やネガティブな感情を感じるかについて Buunk, Ybema, Van der Zee et al. (2001)などを参考にして各1項目を新たに作成し用いた。上方比較によるポジティブ（ネガティブ）な感情については“今と比べてもっと良かった頃のことを思い出して、よい（不快な）気分になること”がどのくらいあるかを、“まったくない”から“よくある”までの5件法で尋ねた。同様に、下方比較によるポジティブ（ネガティブ）な感情についても“今よりも悪かった頃のことを思い出して、よい（不快な）気分になること”がどのくらいあるか尋ねた。

**手続き** 質問紙を授業時に一斉に配布し、回収する形で実施した。

### 5.2.3.3. 結果と考察

まず、継時的比較志向性尺度の妥当性検討のため、比較対象となる過去への関心と関連しているかを検討した。継時的比較志向性と、回想量の相関を求めたところ、 $r=.52$  ( $p<.01$ ) の有意な相関が認められた。また、比較の頻度との関連では、上方比較とは  $r=.43$  ( $p<.01$ )、下方比較とは  $r=.47$  ( $p<.01$ ) の有意な相関がみられた。そのため、調査1で不十分であった過去への関心の高さとの関連は確認されたと考えられ、継時的比較志向性尺度の妥当性があらためて確認された。

さらに、継時的比較志向性が比較後の感情とどのように関連しているかについて検討を行った。その結果、継時的比較志向性の高さは、上方比較後にネガティブな感情を経験する程度 ( $r=.33$ ,  $p<.01$ )、及び下方比較後にネガティブな感情を経験する程度 ( $r=.33$ ,  $p<.01$ ) の両方と有意な相関がみられた。また、上方比較の後にポジティブな感情を経験する程度

とも弱いながらも有意な相関がみられた ( $r=.16$ ,  $p<.05$ )。

これらのことから、継時的比較によってネガティブな感情を経験する程度は、継時的志向性が高い人の方が強いことが示唆された。すなわち過去の良かった頃の自分と比較を行った場合も、過去の悪かった頃の自分と比較を行った場合も、比較志向性が高い人の方が一様にネガティブな感情を強く経験する傾向が見られた。社会的比較志向性を扱った Buunk, Zurriaga, Gonzalez-Roma, & Subirats (2003) などでも、今回と同様に比較後のネガティブ感情との相関が見られていた。そのため、継時的比較においても、比較への志向性が重要な役割を担っていることが考えられ、本尺度を適用した研究の可能性が示唆される。

しかしながら、継時的比較と社会的比較では、同じ比較過程であってもその構造は必ずしも同じではないと考えられる。また、社会的比較志向性に関する研究間でも、比較志向性と比較後の感情との関連については、必ずしも一貫した結果が得られているわけではない。実際に、Buunk, Zurriaga, et al. (2005) では、Buunk, et al. (2003) とは異なり、社会的比較志向性と下方比較後のポジティブ感情との間にも相関が見られている。そのため、継時的比較過程に関するモデルに基づいたより詳細な検討が必要である。

例えば、継時的自己評価理論 (Ross & Wilson, 2002) では、比較される“過去の自己”は“現在”との主観的な時間距離などによって“現在の自己”に対してさまざまな影響を与えるとされていた。そのため、継時的比較志向性が高い人の上方比較および下方比較におけるネガティブな感情の経験は、継時的比較志向性が過去との主観的な距離に影響を与えることによって生じているとも考えられる。佐藤 (2008) では、現在の自己概念や自己評価と一致する“過去”は近く、一致しない“過去”は遠く感じられることが示されている。調査1で示唆されたように、継時的比較志向性が高い人は、過去を頻繁に意識すると同時に、現在の自己に対する評価は低い傾向がある。そのため継時的比較志向性が高く、自己評価が低い人ほど、悪い過去は近く、良い過去は遠くに感じられるようになったのではないだろうか。これらはいずれも現時点に対してネガティブな影響を与えるものである。しかしながら、この点については自伝的記憶などの関連する領域の知見も併せてさらに検討していく必要があるだろう。

#### 5.2.4. 総合考察

今回の研究では、継時的比較志向性尺度を作成し、信頼性、妥当性の検討を行った。その結果、作成された尺度は、十分な信頼性、妥当性を備えていることが確認された。また、調査2では継時的比較志向性が比較後の感情とも関連していることが示唆された。継時的比較志向性は、自尊感情との相関が決して高くはなかったことから示唆されるように、自己評価に対して直接的に影響を与えているわけではない。しかしながら、社会的比較志向性 (Gibbons & Buunk, 1999) の場合と同様に、比較過程について詳細に検討していく

上では、重要な変数であると考えられる。特に継時的比較志向性の高い人ほど、比較後のネガティブ感情を強く感じる傾向がみられた点などは、継時的比較や“過去”が自己の感情や評価に与える影響を検討する上で重要であろう。たとえば“ネガティブな反すう”のように、悪い過去を反すうすることは抑うつなどの不適応とも関連していることが指摘されている（伊藤・上里，2001）。しかし、本研究の結果からは、継時的比較志向性の高い人ほど、悪い過去との比較（下方比較）だけでなく、良い過去との比較（上方比較）でもネガティブな感情を強く経験していた。これは“過去”とのかかわりについて、従来の研究とは異なる視点を与えるものである。そのため、今回作成された継時的比較志向性尺度を用いて更なる検討を進めていくことは、自己に対する評価や感情を研究する上で意義のあることであると考えられる。

しかしながら、比較の方向性をとらえることができていない点など、今回作成された尺度には限界も存在する。調査1では、継時的比較志向性尺度は、肯定的回想尺度とも否定的回想尺度とも正の相関がみられていた。これは、継時的比較志向性の高さによって志向される“過去の自己”には、ポジティブなものである場合もあれば、ネガティブなものである場合もあることを示唆している。比較対象が現在よりも優れているか劣っているか、すなわち上方比較なのか下方比較なのかは、比較過程やそれによる自己評価などを捉える上では非常に重要な視点である（Tesser, 1988 ; Wills, 1981 ; Ross & Wilson, 2002）。今回の尺度は、これらの方向性は考慮に入れず、純粹に比較への志向性の高さを捉えようとするものであったが、これらの比較の方向性も今後捉えていく必要があるだろう。

今回焦点を当てた比較志向性に限らず、継時的比較に関する研究は、社会的比較に比べてまだ少ないのが現状である。そのため継時的比較に関する研究が、今後様々な視点からさらに進められることが期待される。

### 5.3. 研究 2 継時的比較志向性尺度短縮版の作成

#### 5.3.1. 問題

前節では、新たに継時的比較志向性尺度を作成するとともに、測定された継時的比較志向性と比較を行った後の感情との関連等を検討した。継時的比較志向性は、自己評価や、感情とも密接に関連していると考えられることから、今後もさまざまな質問紙調査や実験的研究が行われることが期待される。しかし、その際には、他の尺度や実験課題と継時的比較志向性尺度が併用されることになる。その場合、一つの概念を測定するために 11 項目に回答するのは、研究の計画を行う上でやや多く感じられることが予想される。実際、前章の Big Five 尺度も 1 下位尺度あたり 12 項目であり、継時的比較志向性尺度とほぼ同じであったが、短縮版のニーズは多く存在していた。

そこで、第 3 章（並川他，2011）の自己記入式抑うつ評価尺度短縮版に関する研究および第 4 章（並川他，2012）による Big-Five 尺度短縮版作成の作成に関する研究同様に、継時的比較志向性尺度についても IRT を適用した分析を行い、その結果をもとに短縮版の作成を試みる。

#### 5.3.2. 調査 3

##### 5.3.2.1. 目的

前節で作成された継時的比較志向性尺度について、IRT を適用して尺度の特徴について分析を行う。また、IRT のパラメタを参考にし、継時的比較志向性尺度短縮版の作成を試みる。

##### 5.3.2.2. 方法

**調査協力者** 調査協力者は、大学生および専門学校生 934 名（男性 395 名、女性 537 名、不明 2 名）、年齢は平均 19.35 歳（ $SD=1.29$ ）であった。

**質問紙構成** 継時的比較志向性尺度：前節で作成された継時的比較志向性尺度（並川，2011）を用いた。尺度は 11 項目からなり、回答選択枝は“全くあてはまらない”“あまりあてはまらない”“どちらともいえない”“ややあてはまる”“とてもあてはまる”の 5 件法であった。

Table 4

継時的比較志向性尺度の項目パラメタ値 (SE)

item	slope	(SE)	location	(SE)		c1	c2	c3
1 *	1.13	(0.06)	-0.79	(0.05)	category	1.68	-0.18	-1.50
2 *	0.78	(0.04)	0.58	(0.06)	(SE)	(0.03)	(0.02)	(0.02)
3 *	0.74	(0.05)	-0.95	(0.06)				
4	0.92	(0.05)	-0.22	(0.05)				
7	0.70	(0.04)	-0.69	(0.06)				
8 *	1.05	(0.05)	-0.37	(0.05)				
9 *	1.24	(0.07)	-0.52	(0.05)				
10	0.61	(0.03)	0.54	(0.06)				
5	0.74	(0.04)	-0.93	(0.06)	category	1.11	0.83	-1.94
6	0.50	(0.03)	0.05	(0.07)	(SE)	(0.08)	(0.05)	(0.05)
11	0.69	(0.03)	-0.46	(0.06)				

\*短縮版項目

## 5.3.2.3. 結果と考察

**継時的比較志向性尺度の分析** まず継時的比較志向性尺度の1次元性の確認のために因子分析(主因子法)を行った。固有値の減衰状況は、第1固有値から順に 5.86, 1.04, 0.82, 0.61, 0.59, …, であり、第1固有値が第2以下の固有値に比べて極めて大きな値を示したため、1次元性が満たされていると判断された。また、信頼性係数の推定値としてクロンバックの $\alpha$ 係数を算出したところ.89と高い値を示した。

次に、継時的比較尺度11項目での合計得点を算出したところ、平均値は36.92( $SD=8.88$ )であった。男女別にも平均値を算出したところ、男性では 35.96 ( $SD=9.23$ )、女性では 37.60 ( $SD=8.56$ ) であり、女性の方が有意に高い値を示した ( $t(805.52)=2.75, p<.01$ )。なお、女性の方が高くなる傾向は、調査1で示された結果と同様であった。

次にIRTを適用し、項目パラメタの推定を行った。モデルにはGPCM (Muraki,1992)を採用し、パラメタ推定用の計算機プログラムにはPARSCALE (Muraki & Bock, 2003)を用いた。

まず、推定されたパラメタを確認したところ、カテゴリパラメタにおいて、c1が1.47, c2が-0.09, c3が0.26, c4が-1.65と一部で逆転していることが示唆された。そこで、本研究では第4章のBig Five尺度と同様に一部のカテゴリを統合して分析を行うこととした。カテゴリパラメタの値等をもとに2番目(あまりあてはまらない)と3番目(どちらともいえない)のカテゴリを統合し、4件法のデータとして再分析を行った。その結果、推定された項目パラメタをTable4に示した。

推定されたパラメタ値からは、項目10(物事がうまく進んでいるかどうか評価する時に

は、昔の状況と比べてみることが多い) や項目 6 (私は、昔の自分だったら同じ状況でどうするのかなどと考えたりすることはない) の slope パラメタがそれぞれ 0.61, 0.50 と相対的に低いことが示された。

次に、11 項目全体を用いてテスト情報量を算出した (Figure1)。テスト情報量の最大値は 9.0 を超えており、 $\theta$  が -2.0 から 2.0 までの範囲で平均情報量も 8.10 と十分な値を示した。また、情報量のピークは 0.0 周辺にあるものの、平均値よりも若干低い特性尺度値の回答者の測定に適していることが示唆された。

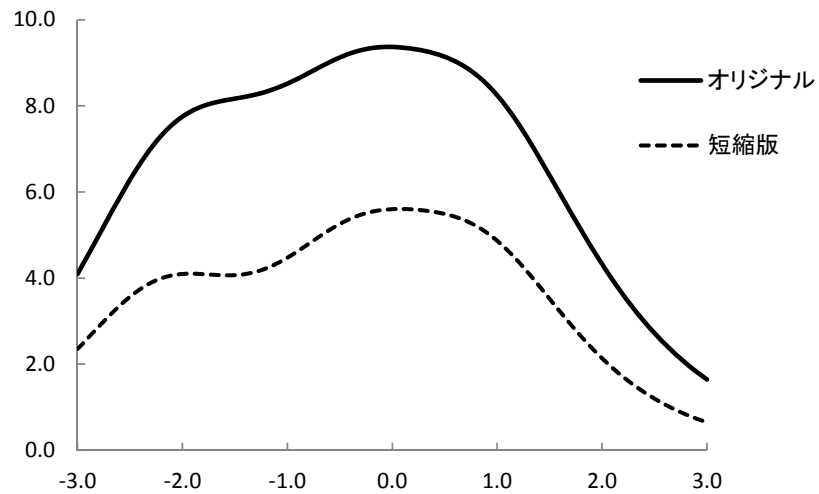


Figure1

継時的比較志向性尺度のテスト情報量

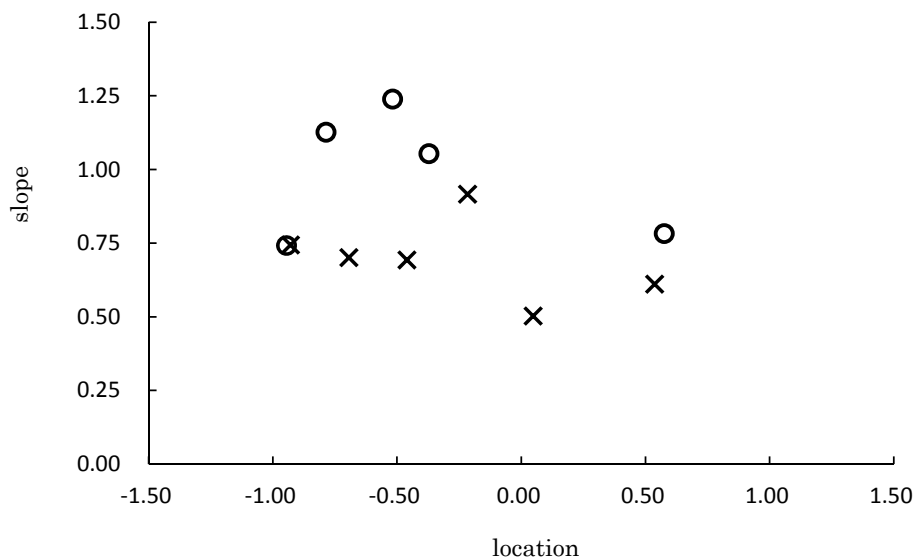


Figure2 継時的比較志向性尺度の項目パラメタ

○ : 短縮版選択項目, × : 短縮版非選択項目



**継時的比較志向性尺度短縮版の作成** 次に、推定された各項目のパラメタ推定値を参考に、短縮版の項目選択を行った。選択の際には、前章までの方法を参考に、まず項目の識別力に関連する slope パラメタの推定値が高いことを基準に、3 項目（項目 1, 8, 9）を選択した。また、それら 3 項目の location パラメタが比較的近い値を示していたため、幅広い特性尺度値の測定に適用できるように location パラメタの値も参考に、追加する項目を検討した。その結果、Slope パラメタの低かった項目 6, 10 を除き、特性尺度値の高い回答者に適した項目として項目 2 を、特に低い回答者に適した項目として項目 3 を選択した（Figure2）。最終的に、5 項目（項目 1, 2, 3, 8, 9）を短縮版の項目として選択した。なお、選択された項目は Table 4 に\*を付ける形で示した。

また、短縮版 5 項目からなる短縮版のテスト情報量も Figure1 に併せて示した。短縮版においても、テスト情報量の最大値は 5.0 を超えており、 $\theta$  が -2.0 から 2.0 の範囲の平均情報量は 4.53 あった。このことから、短縮版でも一定程度の測定精度が確保されていると考えられる。

### 5.3.3. 調査 4

#### 5.3.3.1. 目的

調査 3 で作成された継時的比較志向性尺度短縮版について、さらに信頼性および妥当性の検討を行う。調査 1 では、妥当性検討の際に、社会的比較志向性尺度に関する研究（Gibbons & Buunk, 1999）を参考に、継時的比較志向性尺度と（a）公的・私的自己意識、（b）過去に対する評価傾向、（c）自己に対する評価や感情などとの関係を検討している。そこで、本研究においても継時的比較志向性尺度短縮版の妥当性検討として、短縮版とこれらの尺度との相関が、オリジナル版 11 項目と同様にみられるかどうかを確認する。

#### 5.3.3.2. 方法

**調査協力者** 大学生および専門学校生 136 名（男性 72 名、女性 64 名）、平均年齢は 20.02 歳（ $SD=2.31$ ）であった。

Table5  
継時的比較志向性尺度と各尺度の相関係数

	私的自己意識	公的自己意識	否定的回想	肯定的回想	再評価傾向
継時的比較志向性	0.39 *	0.51 *	0.19 *	0.28 *	0.24 *
継時的比較志向性（短縮版）	0.43 *	0.54 *	0.22 *	0.23 *	0.15

	自尊感情	抑うつ
継時的比較志向性	-0.20 *	0.26 *
継時的比較志向性（短縮版）	-0.24 *	0.28 *

\*  $p < .05$

**質問紙構成** (a) 公的自己意識・私的自己意識：押見・渡辺・石川（1985）による自己意識尺度から、公的自己意識および私的自己意識の2つの下位尺度を用いた。(b) 過去に対する評価：野村・橋本（2001）による再評価傾向尺度、肯定的回想尺度、否定的回想尺度を用いた。(c) ネガティブな評価：自己に対する評価・感情：抑うつ傾向と自尊感情の2点を取り上げた。抑うつについては SDS (Zung, 1965; 福田・小林, 1973) を、自尊感情については Rosenberg (1965) による自尊感情尺度の邦訳版 (山本・松井・山成, 1982) を用いた。(d) 継時的比較志向性：継時的比較志向性尺度の11項目を用いた。

### 5.3.3.3. 結果と考察

継時的比較志向性尺度について、信頼性係数の推定値としてクロンバックの $\alpha$ 係数を算出したところ.88と高い値が得られた。また、短縮版5項目でも同様に $\alpha$ 係数を算出したところ.81と高い値が得られた。そのため、継時的比較志向性尺度短縮版は、内的に一貫した尺度であることが示され、十分な信頼性を備えていると判断された。

次に、妥当性の検討のため、継時的比較志向性尺度および短縮版と各尺度間の相関係数を算出した (Table5)。その結果、継時的比較志向性尺度短縮版は、私的自己意識尺度、公的自己意識尺度とそれぞれ.43 ( $p < .05$ )、.54 ( $p < .05$ ) の有意な正の相関が見られた。また、過去に関連する尺度との相関では、肯定的回想尺度 ( $r = .22, p < .05$ )、否定的回想尺度 ( $r = .23, p < .05$ ) とは有意な相関がみられたが、再評価傾向尺度は  $r = .15$  と有意な相関は認められなかった。自尊感情尺度 ( $r = -.24, p < .05$ ) および抑うつ尺度 ( $r = .28, p < .05$ ) とはいずれも有意な相関が認められた。これらは、いずれもオリジナル版と同様の結果であった。再評価傾向尺度との相関係数は有意ではなかったものの、調査1で報告されているオリジナル版での値も  $r = .19$  と低く、概ね同じ傾向であると考えられる。そのため、継時的比較志向性尺度短縮版は、オリジナル版と同様の特徴を備えており、一定の妥当性を持つ尺度であると考えられる。

### 5.3.4. 総合考察

本研究では、継時的比較志向性尺度について IRT を用いた分析を行い、短縮版の作成を行った。まず、テスト情報量を示した Figure1 からは、継時的比較志向性尺度（オリジナル版）が平均値よりも低い特性尺度値を持つ回答者の測定に適していることが示唆された。尺度作成において、IRT を用いた分析を行う研究はまだあまり多くは見られないが、本来このような情報についても十分に検討される必要があるものと考えられる。また、本研究で作成された短縮版は、オリジナル版の持つ特徴を活かせるように項目選択を行った。そのため、オリジナル版と同様に平均値より低い特性尺度値の方が、テスト情報量は高くなっていた。たとえば、Figure1 に示したように短縮版の情報量は  $-1 < \theta < 1$  の範囲では 4.8 以上が維持されているものの、 $\theta = -2$  では 4.6 であるのに対し、 $\theta = 2$  では 2.1 となっていた。今回は分析にあたって複数の選択枝のカテゴリを統合した上で分析を行っているため、解釈は慎重を期す必要があるが、短縮版尺度の特徴を理解する上で重要な情報であると考えられる。

また、短縮版尺度を用いる目的によっては、能力の高い回答者について特に高い精度で測定を行いたい場合等もあると考えられる。その場合、短縮版の項目の構成を、今回のものとは異なるものによって、テスト情報量のピークがプラスの側に来るように工夫する必要があるだろう。しかしながら、継時的比較志向性尺度では、Location パラメタが正の値を示す項目が 11 項目中 3 項目しかなく、十分な精度を保つのは難しい可能性がある。このことから、さまざまな目的に合わせた短縮版の尺度構成を行うためには、ある程度多様な特性を持った項目がオリジナル版や項目プールにおいて用意されている必要があることが示唆される。

## 5.4. 本章のまとめ

本章では、新たに 11 項目からなる継時的比較志向性尺度（オリジナル版）を作成し、今後の継時的比較や自己評価研究における意義や限界について示すとともに、そこから 5 項目の短縮版の作成を行った。まだ一つの事例ではあるものの、前章までのように既存尺度から短縮版を作成するだけでなく、新規のオリジナル版尺度の作成からその短縮版作成までの一連のプロセスを確認できたことは、意義のあることと考えられる。

また、今回のオリジナル版作成においては、基本的に因子分析や I-T 相関、 $\alpha$  係数などにに基づき、項目の取舍選択を行う方法を採用していた。これは尺度作成研究で多く行われている方法であると考えられる。しかし、短縮版作成研究において IRT を適用することで、5 件法の選択枝がうまく機能していない可能性も指摘されるなど、尺度の持つ特徴についての検討も可能になった。これらの IRT に基づいた検討は、短縮版作成のみならず、オリ

ジナル版の作成においても、本来は必要なものであると言えるだろう。

本章だけでなく、第 3 章から第 5 章までの間では、計 3 種類の尺度を取り上げ、IRT を用いた短縮版作成を行ってきた。同様の研究はまだ少ない上、今回行った研究においても課題はいくつか残されているものの、短縮版の作成において IRT を用いることについて一定の意義を示すことはできたと考えられる。そのため、今後も同様の検討が、多様な尺度を対象に行われていくことが期待される。

また、ここまでの研究では、IRT を用いた方法と、従来多く用いられてきた因子負荷量を用いた方法との間で、作成される短縮版の項目がどのように異なるのかなどについてはまだ十分に検討されていない。そのため、作成の方法によって、短縮版の持つ測定機能がどのように異なるのかを含め、今後もさまざまな尺度やデータを用いて検討を行っていく必要があるだろう。なお、ここまでの研究では、短縮版の項目選択に IRT のパラメタを活用するだけにとどまっていたが、回答者の得点算出においても IRT を活用することも重要になってくると考えられる。そこで、章をあらためてこれらの点について検討を行う。

## 第6章 短縮版作成方法による差異に関する検討

## 6.1. はじめに

第3章から第5章では、IRTを適用した分析をもとに、短縮版の作成を試みた。いずれも一定の信頼性・妥当性が確認され、短縮版作成においてIRTを用いることは有効であることが示唆された。しかしながら、従来多く行われている短縮版作成方法と比較して、IRTを用いた方法がどのように異なるのかについて等は、まだ明らかになっていない。そこで本章では、短縮版作成に関してIRTを用いることによって得られる結果は、他の方法による結果とどのように異なるのかを明らかにすることを目的とし、以下の2点について検討を行う。

1点目は、項目の選択方法に関する検討である。第3章から第5章までで行ったIRTをもとにした方法の特徴を明らかにするために、従来最も多く用いられている因子分析の結果を用いた方法によって項目選択を試み、それらが前章までで作成された短縮版の項目とどのように異なるのかを検討する。また、それら異なる方法で作成された短縮版の信頼性や妥当性についても比較を行い、両者の違いを明らかにする。

2点目は、得点算出方法に関する検討である。IRTを用いて短縮版を作成することのメリットの一つに、推定された項目パラメタを利用して回答者の能力値の推定を行うことによって、オリジナル版と同一の次元上で評価が行えることが挙げられる。しかし、実際の調査などで尺度が利用される際には、各項目への反応の合計値（もしくは平均値）が算出されることが多いと考えられる。そのため、前章までではそのような形で算出された得点をもとに、妥当性の検討などを行ってきた。しかし、IRTを適用するメリットを活かすためには、回答者の得点算出においてもIRTのパラメタを活用する必要があると考えられる。そこで、両者を比較し、相違点などを検討することで、IRT活用の可能性について検討を行う。

## 6.2. 項目の選択方法に関する検討

### 6.2.1. 問題と目的

第3章から第5章にかけて、三つの尺度について短縮版作成を試みた。そこでは、IRTを用いた分析を行い、従来多く用いられている因子分析の結果をもとに作成する方法とは異なる方法を採用して項目選択を行った。しかし、そこで作成された短縮版が、因子分析を基にした従来の方法が採用された場合とどのように異なるのかについては明らかになっていない。

そこで、本章では、これまでの検討を踏まえつつ、短縮版作成において項目選択の方法が異なることによって作成される短縮版がどのように異なるのかについて検討を行う。具

体的には、まず因子分析を用いて短縮版作成の試みを行う。そしてその結果と、前章までで行った IRT を用いた場合の結果とでは、選択される項目がどのように異なるのか、また、それらによって作成された短縮版の信頼性や妥当性等の特徴はどのように異なるのかを明らかにする。

## 6.2.2. 方法

### 6.2.2.1. 調査協力者および使用尺度

前章までで実施された DSRS-C, Big Five 尺度, 継時的比較志向性尺度の 3 尺度に関する調査データを対象に、再分析を行った。使用データは、基本的に第 3 章から第 5 章でそれぞれの尺度の短縮版を作成した際と同一のデータであった。

以下に尺度ごとに分析に用いたデータの調査協力者の人数のみを掲載した。データ及び尺度等の詳細については、第 3 章から第 5 章を参照されたい。

**DSRS-C** 第 3 章で行われた研究のデータを用いた。調査協力者は、小学校 3 年生から 6 年生の児童、および中学校 1 年から 2 年の生徒 4683 名（男性 2414 名、女性 2269 名）であった。

**Big Five 尺度** 第 4 章研究 1 のデータを用いた。調査協力者は、高校生および大学生、専門学校生 2099 名（男性 906 名、女性 1165 名、不明 28 名）であった。なお、妥当性検討の分析には第 4 章研究 2 のデータを用いた。調査協力者は、大学生 238 名（男性 112 名、女性 116 名）であった。

**継時的比較志向性尺度** 第 5 章調査 3 のデータを用いた。調査協力者は、大学生および専門学校生 934 名（男性 395 名、女性 537 名、不明 2 名）であった。また、妥当性検討の分析には、第 5 章調査 4 の大学生および専門学校生 136 名（男性 72 名、女性 64 名）のデータを用いた。

### 6.2.2.2. 分析方法

まずそれぞれの尺度について、あらためて因子分析を行った。そして、その結果をもとに項目選択を行い、暫定的な短縮版を作成した（以下 FA 版と表記する）。その上で、前章までで作成された IRT を基にした短縮版（IRT 版とする）と比較を行った。なお、因子分析の結果をもとに作成する際は、第 2 章の調査において多く見られた因子負荷量の高い順に項目を選択する方法を採用した。また、短縮版の項目数については、FA 版も IRT 版と同一になるように設定した。

Table1 DSRS-C の因子分析結果

	F1	F2
12 いつものよ	0.70	0.00
7 元気いっぱ	0.66	0.00
8 食事が楽し	0.62	-0.07
1 楽しみにし	0.57	-0.05
13 家族と話す	0.56	-0.10
11 やろうと思	0.48	0.04
16 落ち込んで	0.46	0.19
4 遊びに出か	0.44	-0.03
2 とても良く	0.42	0.06
9 いじめられ	0.33	0.12
17 とても悲し	-0.08	0.78
3 泣きたいよ	-0.12	0.71
15 独りぼっち	0.05	0.62
5 逃げ出した	0.06	0.61
10 生きていて	0.10	0.49
14 こわい夢を	-0.06	0.44
6 おなかが痛	0.06	0.31
18 とても退屈	0.29	0.30
因子間相関		
F1		0.53

### 6.2.3. 結果および考察

#### 6.2.3.1. 項目の相違について

DSRS-C, Big Five 尺度, 継時的比較志向性尺度の 3 尺度について, それぞれ下位尺度ごとに因子分析を行った。まず DSRS-C オリジナル版 18 項目に対する因子分析 (主因子法・プロマックス回転) の結果を Table1 に示した。この結果をもとに, 各因子で負荷量の高い項目から順に項目を選択し, 短縮版 (FA 版) を構成した。なお, 選択する項目数は, 3 章で作成した IRT 版と同一としたため, 「活動性および楽しみの減衰」では 5 項目, 「抑うつ気分」では 4 項目であった。

各項目の負荷量と選択された項目の一覧を Table2 に示した。また, Table2 には 3 章で作成された短縮版 (IRT 版) の項目と Slope, Location パラメタも併記した。そこに示した通り, 短縮版 (FA 版) と短縮版 (IRT 版) の間では, 「活動性および楽しみの減衰」では 5 項目中 3 項目が, 「抑うつ気分」では 4 項目中 3 項目がいずれの短縮版でも共通して選択されていた。

なお, 因子負荷量と IRT の Slope パラメタの相関係数を算出したところ, 「活動性および楽しみの減衰」では.90, 「抑うつ気分」では.91 であった。今回, IRT 版の作成においては, 「Slope パラメタの推定値が大きいことや幅広い潜在特性尺度値の推定が可能になる



Table2 DSRS-C の短縮版の項目および因子負荷量，項目パラメタ

	因子分析		IRT		
	負荷	短縮版	Slope	Location	短縮版
1 楽しみにし	0.57	○	0.85	1.30	○
2 とても良く	0.42		0.49	1.10	
4 遊びに出か	0.44		0.53	1.89	○
7 元気いっぱ	0.66	○	1.20	1.24	○
8 食事が楽し	0.62	○	0.82	1.31	
9 いじめられ	0.33		0.41	0.98	
11 やろうと思	0.48		0.77	0.17	○
12 いつものよ	0.70	○	1.28	0.64	○
13 家族と話す	0.56	○	0.60	1.29	
16 落ち込んで	0.46		0.74	0.80	
3 泣きたいよ	0.71	○	1.04	1.40	○
5 逃げ出した	0.61	○	1.05	1.40	
6 おなかが痛	0.31		0.41	1.51	
10 生きていて	0.49		0.90	1.78	○
14 こわい夢を	0.44		0.47	1.50	
15 独りぼっち	0.62	○	1.10	1.37	○
17 とても悲し	0.78	○	1.75	1.38	○
18 とても退屈	0.30		0.52	1.19	

ように Location パラメタの推定値が高いものから低いものまでバランスよく散らばる」ことを基準に項目の選択を行っていた。そのため，9 項目中 6 項目が共通する形にとどまったが，IRT を適用した場合でも Slope パラメタの高さのみを基準にして項目を選択した場合は，因子分析による結果と同一になる場合が多くなると考えられる。実際，今回の DSRS-C のデータにおいても Slope パラメタの大きさだけを基準に項目選択を行った場合は，IRT 版と FA 版で全 9 項目中 8 項目が一致する結果となった。

次に，Big Five 尺度についても DSRS-C と同様の分析を行った。まず因子分析（主因子法・プロマックス回転）を行ったところ，Table3 に示すような結果が得られた。この結果をもとに，各因子で負荷量の高い項目から順に選択し，短縮版（FA 版）とした。なお，各下位尺度で選択する項目数は，4 章で作成した IRT 版と同一としたため，全体で 5 因子 29 項目であった。Table4 に，下位尺度ごとの因子負荷量と IRT のパラメタをそれぞれ示した。FA 版と IRT 版で共通する項目も多く，全 29 項目中 21 項目は共通の項目が選ばれていた。ただし，いずれの下位尺度においても，まったく同一の項目の組み合わせになることはなかった。

Table3 Big Five 尺度の因子分析結果

	F1	F2	F3	F4	F5		F1	F2	F3	F4	F5
n02 不安になりやすい	0.83	0.06	-0.14	0.01	-0.01	c01 いい加減な	0.06	-0.06	0.06	0.78	-0.02
n03 心配性	0.79	0.10	-0.09	0.10	-0.04	c04 成り行きまかせ	0.10	0.03	0.01	0.64	-0.14
n01 悩みがち	0.76	0.05	-0.02	0.06	-0.06	c02 ルーズな	-0.01	-0.05	0.03	0.64	-0.02
n06 傷つきやすい	0.75	0.05	0.05	0.03	0.02	c03 怠惰な	-0.08	0.08	0.06	0.62	0.03
n05 弱気になる	0.74	-0.06	-0.20	-0.11	0.03	c08 軽率な	-0.01	-0.02	-0.09	0.52	0.12
n07 動揺しやすい	0.70	0.15	-0.21	-0.13	0.04	c06 計画性のある	0.08	-0.03	0.32	0.50	-0.03
n04 気苦労の多い	0.63	-0.06	0.10	0.11	-0.04	c05 不精な	-0.09	0.20	-0.02	0.50	-0.03
n09 くよくよしない	0.60	-0.01	-0.23	0.10	-0.02	c07 無頓着な	0.07	0.16	0.00	0.47	-0.01
n10 悲観的な	0.54	-0.18	0.00	-0.03	-0.08	c11 几帳面な	0.20	-0.18	0.27	0.46	0.01
n08 神経質な	0.52	-0.23	0.21	0.20	-0.16	c12 飽きっぽい	-0.05	-0.07	-0.01	0.37	0.19
n11 緊張しやすい	0.51	0.01	-0.11	-0.05	0.07	c10 無節操	0.00	0.15	-0.15	0.36	0.12
n12 憂鬱な	0.51	-0.36	0.08	-0.08	-0.02	c09 勤勉な	0.12	-0.10	0.31	0.34	0.10
e02 無口な	0.01	0.83	-0.18	0.04	-0.11	a01 温和な	0.12	-0.04	0.15	-0.19	0.73
e07 社交的	0.06	0.73	0.13	-0.04	0.07	a02 短気	-0.17	-0.21	-0.04	0.08	0.71
e04 外向的	0.00	0.73	0.17	-0.05	-0.06	a03 怒りっぽい	-0.22	-0.19	-0.06	0.04	0.70
e05 暗い	-0.19	0.71	-0.09	0.14	-0.03	a04 寛大な	-0.02	-0.09	0.33	-0.22	0.62
e01 話好き	0.17	0.70	-0.01	-0.12	-0.07	a12 反抗的	-0.05	-0.03	-0.17	0.14	0.57
e03 陽気な	0.03	0.69	0.11	-0.20	0.01	a08 とげがある	-0.02	0.16	-0.28	0.11	0.51
e06 無愛想な	0.03	0.65	-0.19	0.15	0.18	a09 かんしゃくもち	-0.25	-0.06	-0.09	0.03	0.50
e08 人嫌い	-0.07	0.62	-0.20	0.09	0.17	a05 親切的な	0.22	0.16	0.27	0.00	0.50
e09 活動的な	-0.03	0.58	0.30	-0.02	-0.08	a06 良心的な	0.21	0.13	0.31	0.07	0.48
e11 積極的な	-0.03	0.56	0.35	-0.01	-0.13	a10 自己中心的	-0.04	-0.04	-0.14	0.26	0.46
e12 地味な	-0.16	0.54	-0.03	0.15	-0.16	a07 協力的な	0.23	0.34	0.13	0.04	0.36
e10 意思表示しない	-0.05	0.51	0.03	0.09	-0.09	a11 素直な	0.13	0.18	0.19	-0.03	0.35
o02 多才の	-0.08	-0.02	0.69	0.06	0.04	因子間相関					
o07 頭の回転の速い	-0.12	-0.02	0.60	0.12	-0.06	F1		-0.26	0.04	-0.12	-0.21
o03 進歩的	-0.06	0.11	0.60	0.08	0.05	F2			0.35	0.11	0.24
o01 独創的な	-0.03	-0.04	0.58	-0.16	-0.02	F3				-0.02	0.09
o05 想像力に富んだ	-0.01	-0.04	0.58	-0.16	0.07	F4					0.26
o04 洞察力のある	0.01	-0.05	0.56	0.13	0.01						
o08 臨機応変な	-0.19	0.12	0.53	0.02	0.02						
o06 美的感覚の鋭い	0.02	-0.05	0.52	0.07	-0.01						
o12 呑み込みの速い	-0.16	0.05	0.50	0.12	-0.03						
o11 独立した	-0.12	-0.14	0.48	0.07	-0.04						
o09 興味の広い	-0.08	0.19	0.44	-0.06	0.00						
o10 好奇心が強い	0.00	0.28	0.39	-0.13	-0.02						

Table4 Big Five 尺度短縮版の項目および因子負荷量, 項目パラメタ

	因子分析		IRT				因子分析		IRT		
	負荷	短縮	Slope	Location	短縮		負荷	短縮	Slope	Location	短縮
n01 悩みがち	0.76	○	0.97	-0.74		a01 温和な	<b>0.73</b>	○	<b>0.36</b>	<b>-1.36</b>	○
n02 不安になりやすい	<b>0.83</b>	○	<b>1.44</b>	<b>-0.68</b>	○	a02 短気	<b>0.71</b>	○	<b>1.00</b>	<b>0.01</b>	○
n03 心配性	<b>0.79</b>	○	<b>0.97</b>	<b>-0.84</b>	○	a03 怒りっぽい	<b>0.70</b>	○	<b>1.00</b>	<b>-0.03</b>	○
n04 気苦労の多い	0.63		0.48	-0.47		a04 寛大な	<b>0.62</b>	○	<b>0.31</b>	<b>-0.81</b>	○
n05 弱気になる	<b>0.74</b>	○	<b>0.85</b>	<b>-0.42</b>	○	a05 親切的な	0.50		0.34	-1.20	○
n06 傷つきやすい	0.75	○	0.65	-0.70		a06 良心的な	0.48		0.34	-1.34	
n07 動揺しやすい	0.70		0.55	-0.71		a07 協力的な	0.36		0.21	-1.37	
n08 神経質な	0.52		0.36	-0.20		a08 とげがある	0.51	○	0.33	-0.12	
n09 くよくよしない	0.60		0.49	-0.25		a09 かんしゃくもち	0.50		0.42	-0.51	
n10 悲観的な	0.54		0.41	0.06		a10 自己中心的	0.46		0.34	0.22	○
n11 緊張しやすい	0.51		0.33	-1.50	○	a11 素直な	0.35		0.17	-1.36	
n12 憂鬱な	0.51		0.42	0.08	○	a12 反抗的	0.57	○	0.43	-0.05	
e01 話好き	<b>0.70</b>	○	<b>0.53</b>	<b>-1.25</b>	○	e01 いい加減な	<b>0.78</b>	○	<b>0.73</b>	<b>0.78</b>	○
e02 無口な	<b>0.83</b>	○	<b>0.62</b>	<b>-0.51</b>	○	e02 ルーズな	<b>0.64</b>	○	<b>0.39</b>	<b>0.39</b>	○
e03 陽気な	0.69		0.75	-0.76	○	e03 怠惰な	<b>0.62</b>	○	<b>0.55</b>	<b>0.46</b>	○
e04 外向的	<b>0.73</b>	○	<b>1.05</b>	<b>-0.22</b>	○	e04 成り行きまかせ	<b>0.64</b>	○	<b>0.39</b>	<b>0.88</b>	○
e05 暗い	0.71	○	0.65	-0.66		e05 不精な	0.50	○	0.42	0.06	
e06 無愛想な	0.65		0.41	-0.52		e06 計画性のある	<b>0.50</b>	○	<b>0.27</b>	<b>0.16</b>	○
e07 社交的	<b>0.73</b>	○	<b>1.08</b>	<b>-0.32</b>	○	e07 無頓着な	0.47		0.35	0.01	
e08 人嫌い	0.62		0.39	-0.89		e08 軽率な	<b>0.52</b>	○	<b>0.39</b>	<b>-0.17</b>	○
e09 活動的な	0.58		0.64	-0.49		e09 勤勉な	0.34		0.20	0.47	
e10 意思表示しない	0.51		0.30	-0.51		e10 無節操	0.36		0.35	-0.50	
e11 積極的な	0.56		0.62	-0.31		e11 几帳面な	0.46		0.23	-0.26	○
e12 地味な	0.54		0.36	-0.02		e12 飽きっぽい	0.37		0.22	1.08	
o01 独創的な	<b>0.58</b>	○	<b>0.40</b>	<b>-0.47</b>	○						
o02 多才の	<b>0.69</b>	○	<b>0.66</b>	<b>0.43</b>	○						
o03 進歩的	<b>0.60</b>	○	<b>0.67</b>	<b>-0.24</b>	○						
o04 洞察力のある	0.56	○	0.38	-0.55							
o05 想像力に富んだ	0.58	○	0.37	-0.68							
o06 美的感覚の鋭い	0.52		0.32	0.14							
o07 頭の回転の速い	<b>0.60</b>	○	<b>0.45</b>	<b>0.03</b>	○						
o08 臨機応変な	0.53		0.45	-0.32							
o09 興味の広い	0.44		0.29	-1.22	○						
o10 好奇心が強い	0.39		0.34	-1.62	○						
o11 独立した	0.48		0.27	0.05							
o12 呑み込みの速い	0.50		0.33	-0.14							

次に継時的比較志向性尺度についても同様の分析を行った。因子分析（主因子法）の結果および、IRT のパラメタを Table5 に示した。両者で共通して短縮版として選択された項目は短縮版の 5 項目中 3 項目であった。

Table5 継時的比較志向性尺度の項目および因子負荷量, 項目パラメタ

	因子分析		IRT		
	負荷	短縮版	Slope	Location	短縮版
1 今の自分と昔の自分を比べることがよくある	0.76	○	1.13	-0.79	○
2 何かをするときに、昔の自分だったらどうしたかを考えることが多い	0.61		0.78	0.58	○
3 自分の過去をよく思い出す	0.60		0.74	-0.95	○
4 今の自分がどんな性格であるかを、昔の自分と比べて考えることが多い	0.75	○	0.92	-0.22	
5 昔の自分について考えることはほとんどない	0.57		0.74	-0.93	
6 私は、昔の自分だったら同じ状況でどうするのかなどと考えたりすることはない	0.62		0.50	0.05	
7 以前よりも自分が成長しているかどうか考えることが多い	0.52		0.70	-0.69	
8 昔の自分の状況と、今の自分の状況を比べることがよくある	0.73	○	1.05	-0.37	○
9 自分が、昔と変わったかどうかを考えることがよくある	0.73	○	1.24	-0.52	○
10 物事がうまく進んでいるかどうか評価する時には、昔の状況と比べてみるが多い	0.40		0.61	0.54	
11 今の自分の境遇と、昔の自分の境遇との違いを考えることはほとんどない	0.69	○	0.69	-0.46	

Table6 各尺度の $\alpha$ 係数

	Original	FA	IRT
DSRS-C			
活動性および楽しみの減衰	.79	.74	.71
抑うつ気分	.77	.76	.73
Big Five			
情緒不安定性	.91	.88	.82
外向性	.90	.85	.86
開放性	.85	.78	.76
調和性	.84	.80	.78
誠実性	.82	.80	.78
継時的比較志向性	.89	.84	.83

### 6.2.3.2. 各短縮版の信頼性についての検討

それぞれの尺度について、IRT 版と FA 版とで信頼性や妥当性にどのような差がみられるかを検討した。まず、3 尺度のオリジナル版、短縮版 (IRT 版)、短縮版 (FA 版) それぞれについて、信頼性係数の推定値としてクロンバックの  $\alpha$  係数を算出し、Table6 に示した。その結果、Big Five 尺度の「外向性」を除きいずれの尺度も FA 版の  $\alpha$  係数の方が値は大きくなっていったものの、IRT 版と FA 版ではほぼ同程度の信頼性が認められた。

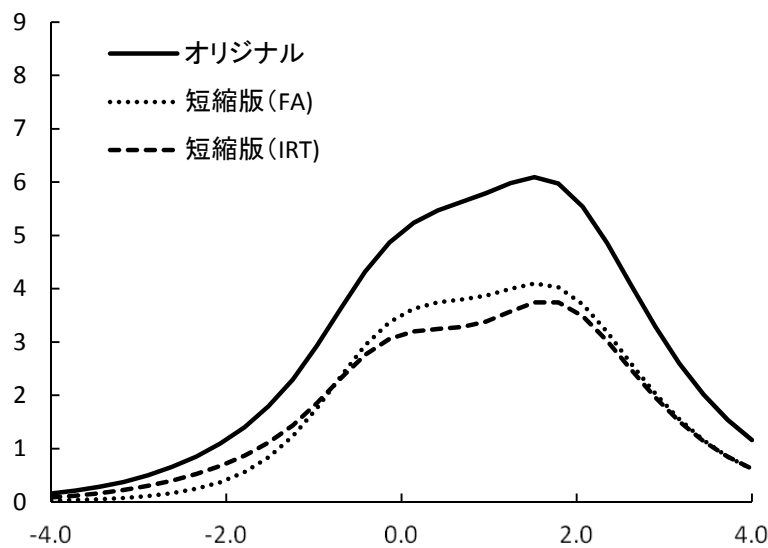


Figure1 「活動性および楽しみの減衰」のテスト情報量

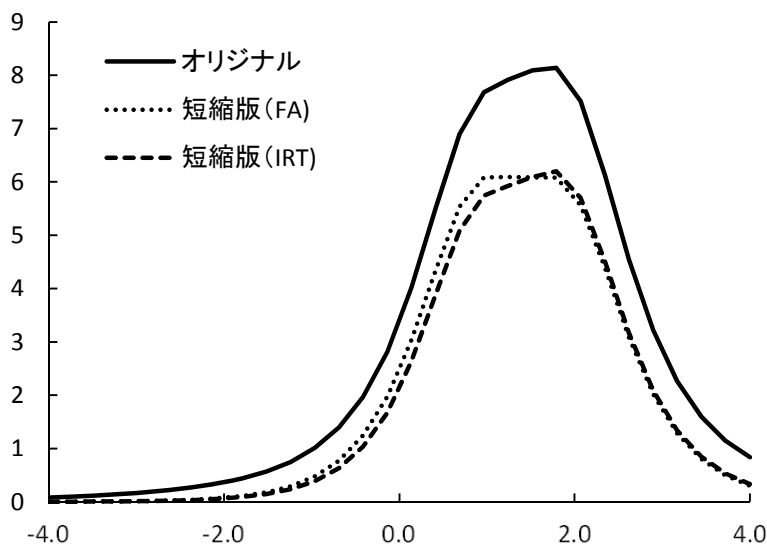


Figure2 「抑うつ気分」のテスト情報量

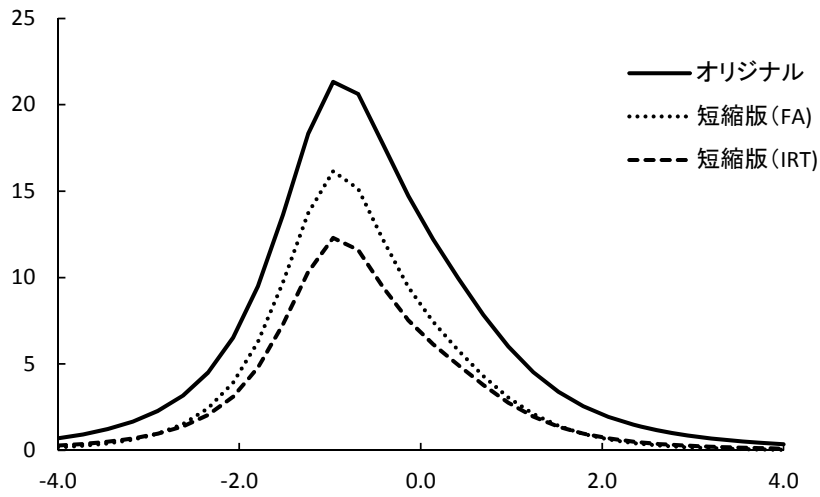


Figure3 「情緒不安定性」のテスト情報量

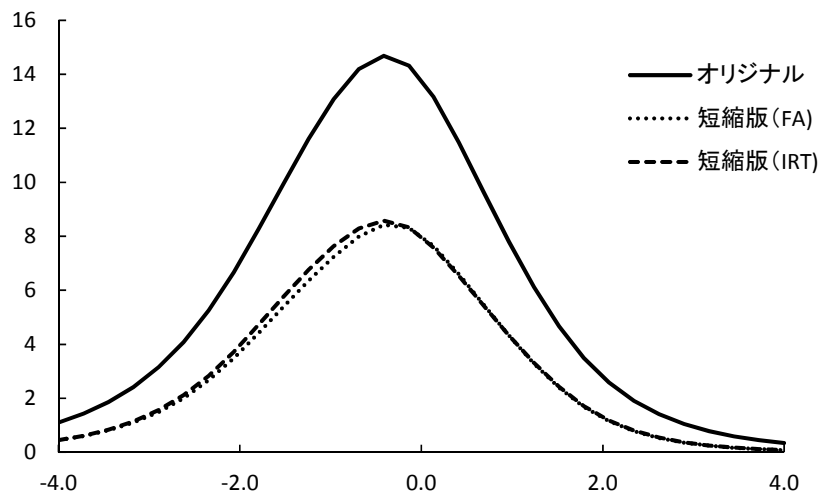


Figure4 「外向性」のテスト情報量

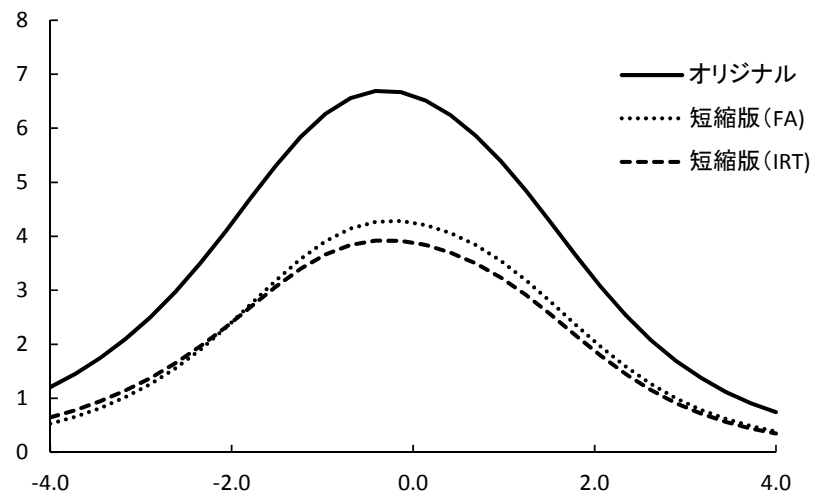


Figure5 「開放性」のテスト情報量

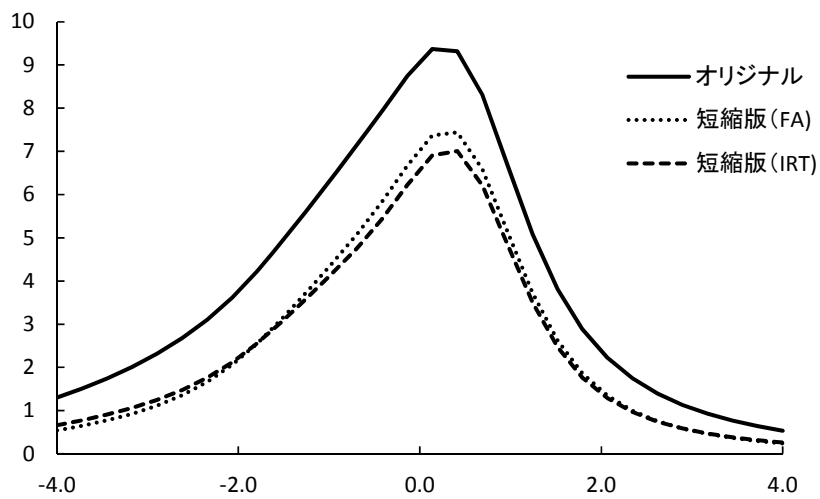


Figure6 「調和性」のテスト情報量

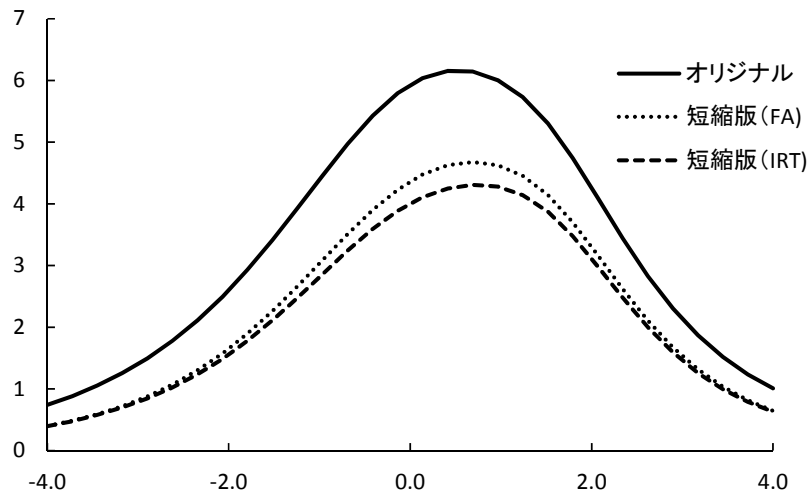


Figure7 「誠実性」のテスト情報量

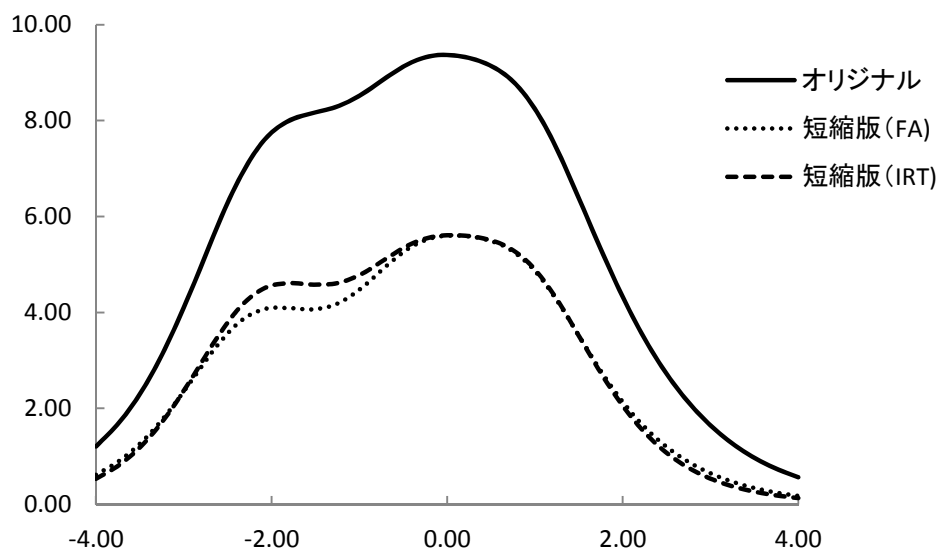


Figure8 「継時的比較志向性」のテスト情報量



また、IRT のテスト情報量を、オリジナル版と IRT 版、FA 版の 3 種類でそれぞれ推定し、Figure1～8 に示した。FA 版と IRT 版は、項目数も同一であり、共通の項目も多いためいずれも同程度のテスト情報量を示していた。しかし、「情緒不安定性」の Figure3 に代表されるように、平均値付近では FA 版の方が高くなる傾向が多く見られた。一方で、IRT 版は必ずしも常に FA 版よりも低いわけではなく、平均から離れた潜在特性値のところなどでは、情報量が FA 版よりも高くなる場合も多く見られた。これらは、IRT 版の項目選択基準が影響していると考えられる。IRT 版では、識別力だけでなく、location パラメタができるだけ幅広くなるように項目選択を行っていた。そのため、情報量のピークは必ずしも高くはならなかった一方で、平均値 ( $\theta = 0$ ) から離れたところでの情報量がやや高くなったと考えられる。

なお、6.2.3.1 でも少し触れたが、IRT 版の項目を slope パラメタの高さのみを基準にして選択した場合、FA 版との類似性が高くなることが予想される。実際、今回扱った 3 尺度（下位尺度の数では 8 尺度）を見てみると、短縮版として選択された全 43 項目中、IRT 版と FA 版は 30 項目が共通だったのに対し、slope パラメタが高い項目から順に項目を選択した場合は FA 版と 36 項目が共通する形になった。そのうち 2 尺度（「抑うつ気分」「情緒不安定性」）は、slope パラメタのみを基準にした場合と FA 版とは、全く同一の組み合わせになった。IRT を用いた場合には、Location パラメタをどのように活用し、項目選択を行うかが重要であることが示唆される。

#### 6.2.3.3. 各短縮版の妥当性についての検討

次に、各尺度について、オリジナル版と二つの短縮（FA 版、IRT 版）との相関係数を求めた。また、第 3 章から第 5 章で妥当性検討に用いた関連尺度との相関係数もそれぞれ求め、Table7～10 に示した。

まず、DSRS-C については、オリジナル版と 2 種類の短縮版（FA 版、IRT 版）、そして第 3 章と同様に「不安（曾我，1983）」と、「攻撃性（坂井他，2000）」との相関を検討した（Table7）。その結果、FA 版も IRT 版もオリジナル版との相関が .89 以上と高くなっており、両短縮版の間にオリジナル版との相関において有意な差は認められなかった。また、FA 版、IRT 版と「不安」との相関は、「抑うつ気分」では  $r = .62$  と  $r = .61$ 、「活動性および楽しみの減衰」では  $r = .35$  と  $r = .39$  と同程度であった。攻撃性についても「抑うつ気分」で  $r = -.36$ 、 $r = -.35$ 、「活動性および楽しみの減衰」で  $r = -.16$ 、 $r = -.15$  と、同程度の相関が見られた。

Table7 DSRs-C のオリジナル版, 短縮版と関連尺度との相関

	不安	攻撃性	抑うつ(FA)	抑うつ(IRT)	活動性(ori)	活動性(FA)	活動性(IRT)
抑うつ気分(original)	.68 *	-.39 *	.90 *	.89 *	.47 *	.40 *	.41 *
抑うつ気分 (FA)	.62 *	-.36 *		.94 *	.41 *	.34 *	.36 *
抑うつ気分 (IRT)	.61 *	-.35 *			.42 *	.34 *	.36 *
活動性および楽しみの減衰(original)	.46 *	-.17 *				.91 *	.90 *
活動性および楽しみの減衰 (FA)	.35 *	-.16 *					.87 *
活動性および楽しみの減衰 (IRT)	.39 *	-.15 *					

\*  $p < .05$ 

Table8 Big Five 尺度オリジナル版, 短縮版と, NEO-FFI の相関

	N (NEO-FFI)	E (NEO-FFI)	O (NEO-FFI)	A (NEO-FFI)	C (NEO-FFI)
N (original)	<b>.73</b> *	-.25 *	.01	-.19 *	-.23 *
N (FA)	<b>.71</b> *	-.15 *	.08	-.10	-.20 *
N (IRT)	<b>.70</b> *	-.26 *	.06	-.17 *	-.24 *
E (original)	-.29 *	<b>.79</b> *	.10	.23 *	.27 *
E (FA)	-.25 *	<b>.74</b> *	.07	.18 *	.25 *
E (IRT)	-.23 *	<b>.74</b> *	.08	.18 *	.22 *
O (original)	-.20 *	.43 *	<b>.25</b> *	-.07	.33 *
O (FA)	-.19 *	.33 *	<b>.17</b> *	-.10	.26 *
O (IRT)	-.23 *	.46 *	<b>.21</b> *	-.03	.33 *
A (original)	-.27 *	.21 *	.07	<b>.54</b> *	.16 *
A (FA)	-.31 *	.08	.05	<b>.50</b> *	.11
A (IRT)	-.27 *	.11	.06	<b>.46</b> *	.15 *
C (original)	-.20 *	.05	-.16 *	.13	<b>.52</b> *
C (FA)	-.21 *	.02	-.17 *	.10	<b>.49</b> *
C (IRT)	-.15 *	-.02	-.17 *	.07	<b>.52</b> *

\*  $p < .05$ 

Big Five 尺度については, 第4章と同様に NEO-FFI (下仲他, 1999) との相関を算出した (Table8)。FA 版, IRT 版いずれも NEO-FFI の対応する因子との相関は同程度であった。また, Table9 にオリジナル版と短縮版 (IRT 版, FA 版) の相関を示した。いずれの短縮版も, オリジナル版の当該尺度と .92 以上の高い相関を示していた。

最後に継時的比較志向性尺度について検討を行った。ここでは, 第5章の調査4と同様に私的自己・公的自己意識 (押見他, 1985), 回想 (否定的回想, 肯定的回想, 再評価傾向; 野村・橋本, 2001), 自尊感情 (Rosenberg, 1965; 山本他, 1982), 抑うつ (Zung, 1965; 福田・小林, 1973) の7種類の尺度を用いた。その結果, 再評価傾向については, FA 版でのみ .26 と有意な相関が得られ, IRT 版では .15 と有意ではなかったものの, その他はいずれの尺度においても FA 版と IRT 版では同程度の相関が得られた。

Table9 Big Five 尺度オリジナル版と短縮版との相関

	N (ori)	N (FA)	N (IRT)	E (ori)	E (FA)	E (IRT)	O (ori)	O (FA)	O (IRT)
N (original)	—	.93 *	.94 *	-.36 *	-.31 *	-.27 *	-.26 *	-.22 *	-.29 *
N (FA)		—	.90 *	-.22 *	-.18 *	-.16 *	-.19 *	-.15 *	-.21 *
N (IRT)			—	-.37 *	-.33 *	-.28 *	-.26 *	-.22 *	-.27 *
E (original)				—	.94 *	.92 *	.44 *	.35 *	.50 *
E (FA)					—	.96 *	.40 *	.32 *	.45 *
E (IRT)						—	.39 *	.32 *	.45 *
O (original)							—	.93 *	.94 *
O (FA)								—	.89 *
O (IRT)									—
A (original)									
A (FA)									
A (IRT)									
C (original)									
C (FA)									
C (IRT)									

	A (ori)	A (FA)	A (IRT)	C (ori)	C (FA)	C (IRT)
N (original)	-.19 *	-.26 *	-.17	-.09	-.11	-.04
N (FA)	-.16 *	-.23 *	-.16	-.08	-.10	-.05
N (IRT)	-.21 *	-.27 *	-.19	-.17 *	-.19 *	-.12
E (original)	.21 *	.12	.10	.11	.09	.02
E (FA)	.16 *	.06	.05	.10	.09	.01
E (IRT)	.14 *	.04	.04	.02	.00	-.06
O (original)	.09	-.02	.10	-.02	-.04	-.04
O (FA)	.07	-.03	.09	-.02	-.04	-.02
O (IRT)	.13 *	.02	.12	-.03	-.04	-.06
A (original)	—	.93 *	.95 *	.33 *	.31 *	.28 *
A (FA)		—	.92 *	.30 *	.28 *	.24 *
A (IRT)			—	.30 *	.30 *	.27 *
C (original)				—	.96 *	.95 *
C (FA)					—	.96 *
C (IRT)						—

\*  $p < .05$ 

Table10 継時的比較尺度オリジナル版、短縮版と、関連尺度との相関

	継時的比較(FA)	継時的比較(IRT)	私的自己	公的自己
継時的比較志向性 (original)	.94 *	.94 *	.39 *	.51 *
継時的比較志向性 (FA)		.92 *	.42 *	.52 *
継時的比較志向性 (IRT)			.43 *	.54 *

	否定的回想	肯定的回想	再評価	自尊感情	抑うつ
継時的比較志向性 (original)	.19 *	.28 *	.24 *	-.20 *	.26 *
継時的比較志向性 (FA)	.21 *	.24 *	.26 *	-.22 *	.24 *
継時的比較志向性 (IRT)	.22 *	.23 *	.15	-.24 *	.28 *

\*  $p < .05$

今回分析を行った **DSRS-C**, **Big Five** 尺度, 継時的比較志向性尺度においては, いずれの短縮版も同程度の信頼性や妥当性が確認されたと言えるだろう。従来から多く行われている因子分析を基にした場合も, **IRT** を基にした場合も, 同程度の精度を持つ短縮版の作成につながることを示された。しかし, これには重複する項目が多いことが影響していると考えられる。また, オリジナル版が一定の信頼性を備えていたということも, 重要であろう。もしオリジナル版に質の異なる項目が含まれていたとしたら, それが含まれるか否かで短縮版の性能は変わってくることも予想される。

### 6.3. 得点算出方法に関する検討

#### 6.3.1. 問題と目的

本研究のこれまでの分析では、IRT を適用した方法で短縮版作成を行った場合においても、項目の選択時のみに IRT に関する情報を用いており、個人の得点算出においては、基本的に各項目の反応を合計する形を採用してきた。しかし、IRT を用いた場合、短縮版作成時に推定された項目パラメタを用いて回答者の特性尺度値も推定し、それを利用することも可能である。この方法を用いれば、理論上はオリジナル版と同一の次元で特性尺度値が推定可能であり、項目を減らすことによって測定次元が変容してしまう可能性を減らすことができる。これは IRT を用いる上でのメリットの一つであると考えられる。そのため、本研究においても、たとえば妥当性検討で他の尺度との相関を検討する際などにこの値を利用することも検討された。しかし、現状の短縮版尺度の利用状況を鑑みると、項目パラメタを利用して特性尺度値を推定することはややハードルが高く、現状においては合計得点を算出する方法が一般的であると考えられることから、より実際の使用に即した形で検討を行った。しかし、IRT のメリットを活かすためには、個人の得点算出においても IRT を適用することが有効であると考えられる。そこで、IRT を適用し特性尺度値 ( $\theta$ ) を推定する形で得点算出を行い、合計得点の場合との比較を試みる。なお、前節までの分析では3尺度のデータを用いていたが、いずれも同様の結果が得られていたことから、ここでは継時的比較志向性尺度のみを用いて議論を行うこととした。

#### 6.3.2. 方法

第5章研究2の調査1で推定された項目パラメタ（第5章 Table4）を用いて、回答者の  $\theta$  を推定した。推定には最尤推定法を用いた。調査協力者のデータは、第5章の研究2調査2と同じ大学生および専門学校生136名（男性72名、女性64名）のデータを用いた。（データの詳細は5章参照のこと。）なお、全項目「5 とてもあてはまる」（もしくは「1 全くあてはまらない」）の最大（もしくは最少）の選択枝を一貫して選んだ調査協力者のデータについては、最尤推定値が求まらないため、以下では除外して分析を進めた。

まず、オリジナル版と、前節までで作成された FA 版、IRT 版の2種類の短縮版の計3種類について、IRT の  $\theta$  もしくは反応の合計値を算出する形でそれぞれ得点を算出した。算出されたのは、i) オリジナル版11項目の  $\theta$ 、ii) 短縮版5項目（IRT 版）の  $\theta$ 、iii) オリジナル版11項目の合計得点、iv) 短縮版5項目（FA 版）の合計得点、v) 短縮版5項目（IRT 版）の合計得点、の5種類であった。

Table11 継時的比較志向性尺度の記述統計

	平均値	(平均値／ 項目数)	標準偏差	最小値	最大値
i) オリジナル版( $\theta$ )	-0.03	—	1.05	-2.61	2.60
ii) 短縮版 (IRT版, $\theta$ )	-0.13	—	0.99	-2.41	2.23
iii) オリジナル版(合計)	37.00	(3.36)	8.95	14	55
iv) 短縮版 (FA版, 合計)	17.29	(3.46)	4.73	7	25
v) 短縮版 (IRT版, 合計)	17.22	(3.44)	4.39	8	25

Table 12 継時的比較志向性尺度の各得点間の相関係数

	ii) 短縮版 (IRT, $\theta$ )	iii) オリジナル版 (合計)	iv) 短縮版 (FA, 合計)	v) 短縮版 (IRT, 合計)
i) オリジナル版( $\theta$ )	.95	.97	.93	.94
ii) 短縮版 (IRT版, $\theta$ )		.91	.91	.96
iii) オリジナル版(合計)			.94	.94
iv) 短縮版 (FA版, 合計)				.92
v) 短縮版 (IRT版, 合計)				

### 6.3.3. 結果と考察

算出された 5 種類の得点について、それぞれの平均値、標準偏差、最小値、最大値を Table11 に示した。合計得点の iii) ～ v) については、参考のため平均値を項目数で除した値も併記した。また、オリジナル版 11 項目で推定した  $\theta$  と、短縮版 (IRT 版) 5 項目で推定した  $\theta$  のヒストグラムを Figure9 に示した。また、それら 2 つの  $\theta$  について、散布図を Figure10 に示した。(なお、参考のため Figure では、全項目で「5 ととてもあてはまる」を選択し、 $\theta$  が推定できなかった調査協力者を 3.00 として表示した。)

短縮版では、オリジナル版に比べて項目数を半減させているため、全項目に「5あてはまる」と回答する人数が増えており、結果として最大値 ( $\theta=3.00$ ) を取る人数も多くなっていた。そのため、 $\theta=3.0$ の人数を含めて、Figure9では両者の得点分布にやや違いは見られた。しかし、Figure10に示したように、両者の相関は非常に高くなっていた。また、Table12に5種類の得点間の相関係数を示したが、いずれの得点間の相関係数も、 $r=.9$ 以上であり、非常に高い相関となっていた。オリジナル版の合計得点 (Table12の iii) と比べてみても、短縮版の  $\theta$  (Table12の ii) は $r=.91$ と高い相関になっていた (Figure11)。

Table13に継時的比較志向性尺度の各得点と関連尺度との相関係数を示した。一部で有意な相関が認められない形になっているものの、全体的には概ねいずれの得点算出方法でも同様の相関が得られており、一定の妥当性が確認された。

Table13 継時的比較志向性尺度の各得点と関連尺度との相関係数

	自尊感情	否定的回想	肯定的回想	再評価	私的自己	公的自己	抑うつ
i) オリジナル版( $\theta$ )	-.16	.19 *	.25 *	.24 *	.38 *	.50 *	.24 *
ii) 短縮版 (IRT, $\theta$ )	-.18 *	.24 *	.20 *	.16	.42 *	.52 *	.21 *
iii) オリジナル版(合計)	-.20 *	.19 *	.28 *	.24 *	.39 *	.51 *	.26 *
iv) 短縮版 (FA, 合計)	-.22 *	.21 *	.24 *	.26 *	.42 *	.52 *	.24 *
v) 短縮版 (IRT, 合計)	-.24 *	.22 *	.23 *	.15	.43 *	.54 *	.28 *

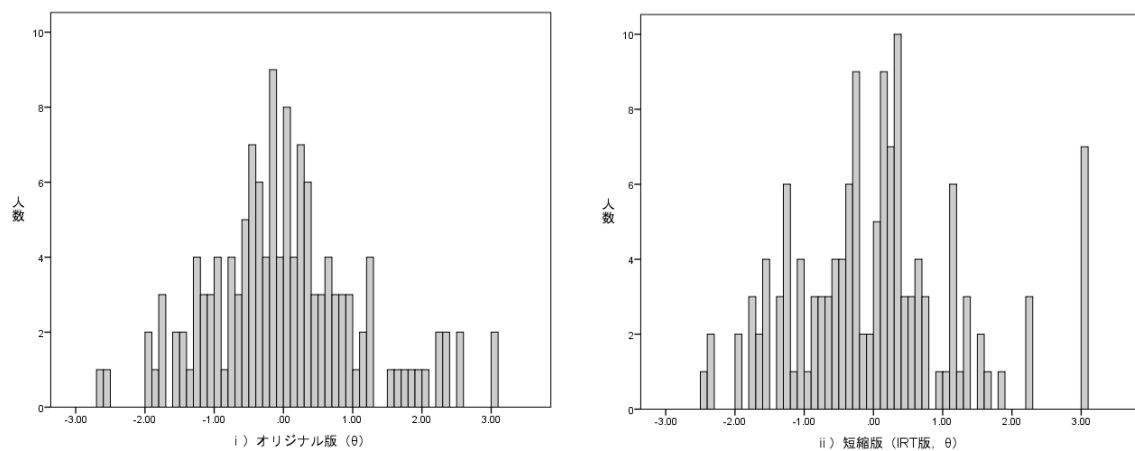
\*  $p < .05$ 

Figure9 継時的比較志向性尺度のオリジナル版と短縮版(IRT 版) 潜在特性尺度値のヒストグラム

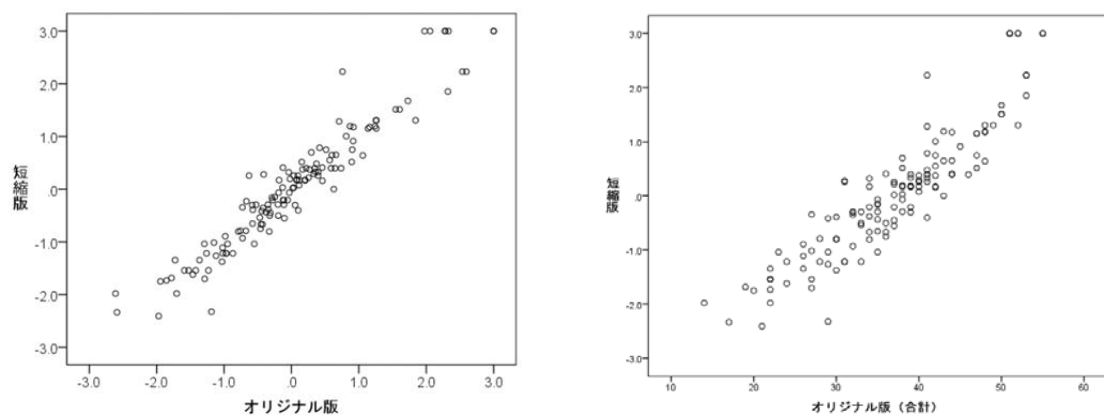


Figure10

継時的比較志向性尺度における  
オリジナル版の  $\theta$  と  
短縮版 (IRT 版) の  $\theta$  の散布図

Figure11

継時的比較志向性尺度における  
オリジナル版の合計得点と  
短縮版(IRT 版)の  $\theta$  の散布図

本研究では $\theta$ と合計得点での差異を検討してきたが、両者の相関も高く、他尺度との相関も同程度に得られていた。そのため、測定されている次元がオリジナル版と短縮版で異なる可能性はあるものの、合計得点を用いても一定程度の測定が可能であることが示唆された。そのため、前章までで作成された短縮版についても、従来の通り合計得点を算出する形で個人の得点を求めることも可能であることが示された。

#### 6.4. まとめ

本章では、短縮版の作成方法について、複数の尺度を用いて検討を行った。IRT版もFA版も、同一項目数で短縮版を構成した場合も、共通して選ばれる項目も多く、作成された短縮版尺度についても信頼性・妥当性に大きな違いは見られなかった。そのため、従来の

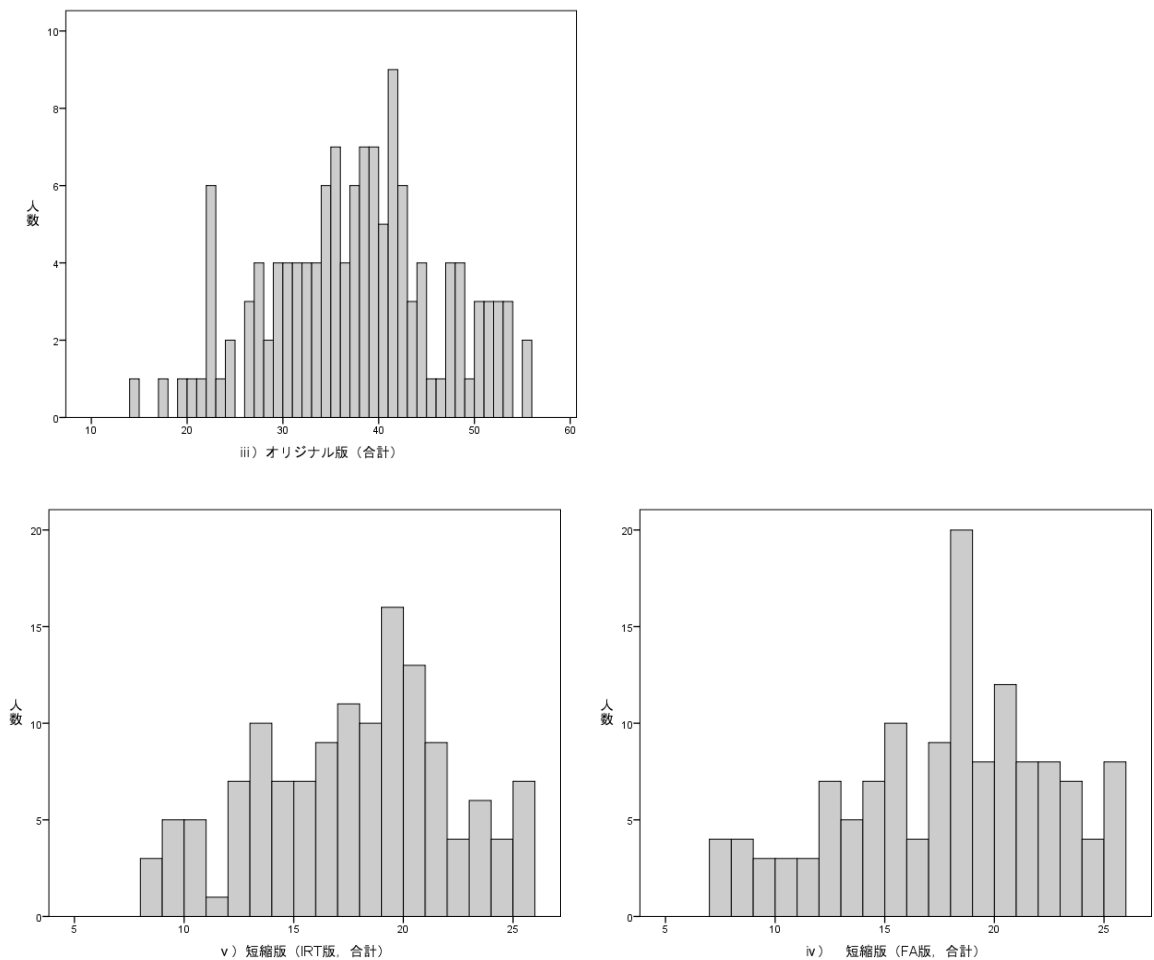


Figure12 継時的比較志向性尺度のオリジナル版と短縮版(IRT版)  
合計得点のヒストグラム



多くの研究が行っているように因子分析をもとに項目を選択した場合も、IRTをもとにして項目を選択した場合も、いずれも一定程度妥当な尺度構成が可能であったと考えられる。また、IRTについても因子分析についても、それぞれ一長一短があるため、使用目的に合った短縮版を作成することが重要であると考えられる。そのためには、いずれかの方法のみを用いるのではなく、Stanton et al. (2002) 等の提案にもあるように、複数の指標や分析結果をもとに項目選択を行うことが重要になってくるだろう。

なお、本研究では基本的にオリジナル版にできる限り忠実な短縮版の作成を意図していた。しかしながら、項目を減らしながらも特定のカットオフ値周辺の測定精度を上げる（もしくは維持する）といったことを目的とした短縮版尺度の構成も重要な課題となってくることが予想される。また、そのような一定の特徴を持った短縮版を作成した場合、6.2や6.3で検討したような信頼性や妥当性なども、今回とは異なる傾向を示す可能性もあるだろう。この点については、今後の検討課題であると言える。

## 第7章 総合考察

## 7.1. 本研究のまとめと今後の課題

本研究では、日本における心理尺度短縮版作成の現状を整理するとともに、IRT を適用した分析を行うことで、従来多くの研究で用いられてきた因子分析以外による短縮版作成方法について検討を行った。第2章で行った日本語論文のレビューの結果では、短縮版作成に因子分析結果を用いる論文が全体の85.7%と多数を占めており、英語論文を対象にした Coste et al. (1997) や Goetz et al. (2013) の結果(40%, 52%)と比較しても多いことが示された。さらに、因子分析の結果のみをもとに短縮版に用いる項目選択を行っていると考えられる論文が32.7%含まれていたことなどから、因子分析以外の多様な指標を用いて短縮版を用いる必要性が指摘された。そして第3章以降では、その短縮版作成時に用いる方法の一つとして、Goetz et al. (2013) でも近年見られるようになった IRT を用いた短縮版の作成を取り上げ、その可能性について検討を行った。第3章から第5章では、それぞれ一つの尺度を取り上げ、IRT を適用した分析をもとに短縮版の作成を試みた。その結果、作成したいずれの短縮版尺度においても、一定の信頼性・妥当性が確認され、短縮版として利用可能であることが示唆された。第6章では、従来多く用いられていた因子分析を用いた場合と、今回の IRT を用いた場合とで短縮版に選ばれる項目を比較するとともに、作成された短縮版の信頼性・妥当性を比較し、いずれの場合も両者に大きな違いが見られないことが示された。また、短縮版の得点算出方法についての比較も行い、合計得点を用いた場合も、IRT で推定された $\theta$ を用いた場合も、非常に近い結果が得られることが示された。

これらの結果から、まず IRT を用いた方法も短縮版作成に有用であることが示された。信頼性・妥当性の高い短縮版が構成された点はもちろんのこと、項目パラメタや情報量など多くの指標が得られることから、オリジナル版尺度も含めて測定上の特徴が詳細に検討できる点も有用であると言えるだろう。しかしながら、本論文で行ったものは実際の短縮版尺度作成の試みや、因子分析を用いて作成された短縮版との比較など、基礎的な検討にとどまっている。そのため、まだ十分に検討できておらず、今後の検討が必要な点もいくつか残されている。そこで、それらについてまとめて行きたい。

最初に指摘するのは、IRT のパラメタ推定のプロセスに関する点である。今回の研究では、Big Five 尺度と継時的比較志向性尺度において、選択枝のカテゴリにおいて順序の逆転が起こっていたために、複数のカテゴリへの反応を統合する形でパラメタの推定を行っていた。たとえば、Big Five 尺度では、「非常にあてはまる」「かなりあてはまる」「ややあてはまる」「どちらともいえない」「あまりあてはまらない」「ほとんどあてはまらない」「まったくあてはまらない」の7件法を用いてデータ収集を行っていたが、分析の過程で両端の二つのカテゴリをまとめることで5件法(5段階)のデータとして扱っていた。しかし、この方法が妥当だったかどうかについては、議論の余地がある。今回のように両端

の選択カテゴリを統合して得られた5件法のデータと、実際に5件法の選択枝に回答を求めて得られたデータとでは、推定されるパラメタは完全には同一のものとはならないと考えられる。そのため、実際に5件法の選択枝に回答を求めたデータについても、分析を行うことも今後必要になってくると考えられる。

また、調査協力者が尺度に回答する際の負担を軽減することが短縮版作成の目的の一つであれば、7件法でなく5件法にするなど選択カテゴリ数を減らすことも有効な方法の一つであると考えられる。特に、今回のようにIRTを用いた分析を行ったことで、オリジナル版において選択枝が十分に機能していないことが示唆されたのであれば、まずそういった方法についても検討を行う必要があると言えるだろう。また、これは単に短縮版作成時の議論にとどまらず、オリジナル版の尺度作成においても検討が必要な点であると言えるだろう。さまざまな分析を行う都合で、一定の分散を確保するため回答選択枝の数は多い方が望ましいと判断される場合もあると考えられる。しかし、その際にも各選択枝のカテゴリが適切に機能しているか否かは、回答者の負担軽減やデータの質の維持を図る上で十分に検討する必要があるだろう。

また、IRTのモデルの選択についても、議論の余地はあったと考えられる。本研究では、任意にGPCM (Muraki, 1992) を採用してパラメタ推定を行ったが、Graded Response Model (Samejima, 1969) など、他のモデルの適用を検討することも可能であったと考えられる。さらに、今回の分析では、カテゴリパラメタを項目間で一貫させるか否かについても、尺度によって異なる方法を採用していた。DSRS-Cでは、項目ごとに個別に推定したのに対し、Big Five尺度と継時的比較志向性尺度では逆転項目・非逆転項目ごとにまとめて推定していた。これらはいずれも、分析に用いるデータの数や推定するパラメタ数、各尺度の特徴などをもとに試行錯誤を経て決定されたものであったが、推定の際にどのような制約を加えるのかなどは、今後もさらなる検討が行われる必要がある。

そして、今回の研究では、作成された短縮版を用いてデータ収集を行った上で、その短縮版尺度の性能の検討を行うことが十分にできていなかったという点も、課題の一つとして挙げられる。オリジナル版に回答したデータをもとにした分析だけではなく、実際に短縮版の項目のみに回答した場合も、一定の信頼性や妥当性が担保されることは確認される必要があるだろう。また、短縮版を利用して新たにデータ収集を行うことで、実際にどの程度回答する際の負担（もしくは時間）を軽減できているのかについても検討が必要であったと考えられる。短縮化によって項目数が半分になっていたとしても、必ずしも回答に要する時間が半分になるとは限らない。たとえば、全体の教示文を理解するのにある程度時間を要するような尺度の場合は、短縮化による所要時間の変化は相対的に少なくなると考えられる。そのため、短縮版の利点をより明確にするためには、たとえば回答にかかる時間がオリジナル版に比べ、どの程度短縮されたのか等についても具体的に明示することも有効であるだろう。また、この点をさらに進めると、オリジナル版尺度作成時において

も回答にかかる凡その時間について明示するなどの工夫が可能であるように思われる。尺度に回答する際にかかる時間や負担を、調査者の経験や勘に頼ることなくある程度正確に見積もることが可能になれば、複数の尺度や課題等を組み合わせた調査を設計する際にも非常に有益である。こういった点についての議論も、今後さらに必要になってくると考えられる。

次に、短縮版の作成目的についても指摘を行いたい。本研究では、オリジナル版にできる限り近い形で、幅広い能力の測定が可能な短縮版の作成を目指していた。そのため、**Location** パラメタが偏らないように考慮し、項目の選択を行っていた。しかし、前章までの考察等でもたびたび触れてきたが、短縮版の利用目的によっては、 $\theta$  が特定の値周辺で精度の高い測定が可能になるように短縮版を構成することも有効な方法の一つである。たとえば、抑うつを測定する尺度であれば、特に抑うつが高い人たちの間で細かな違いを測定する必要がある場合もあれば、スクリーニングを目的とし、カットオフ値周辺で精度の高い測定が必要になる場合もあると考えられる。それぞれの場合で適切な項目は異なると考えられるため、**Location** パラメタ等の情報をもとにそれぞれの目的に適した項目を選択し、短縮版尺度が構成される必要が出てくる。本研究では、これらの目的に応じた尺度構成は検討できなかったが、IRT のメリットを活かす上では重要な点であると言えるだろう。ただし、この場合は、オリジナル版に十分な項目数が含まれていることが必要になってくると考えられる。今回扱った尺度はいずれのオリジナル版も 10 項目程度であり、さまざまな目的に応じて尺度を構成するには、必ずしも十分とは言えない項目数であった。たとえば、継時的比較志向性尺度を例に挙げると、オリジナル版は 11 項目あり、そこから 5 項目を選択する場合には理論上は 462 通りの組み合わせが考えられる。しかし、尺度の中

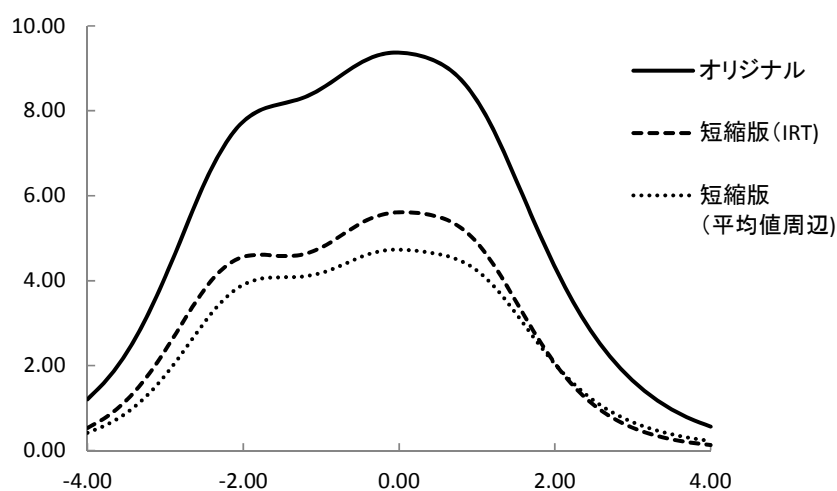


Figure1 継時的比較志向性尺度の情報量：  
オリジナル版と、短縮版(IRT 版)，**Location** パラメタが平均値周辺で  
あることを基準に選んだ場合の比較

には Slope パラメタが低い項目もいくつか含まれており、Location パラメタだけを見て項目選択を行うことなどは難しい。実際、Figure1 に第5章で推定されたパラメタをもとに、 $\theta$  が平均値周辺において精度が高い測定を目指して、単純に Location パラメタが 0 周辺（絶対値が小さい）であることのみを基準に項目を選択して短縮版を作成した場合のテスト情報量を示した。この場合、Figure1 から分かる通り、目的とした平均値周辺においても IRT 版に比べて情報量が大きく下がる形になってしまう。これは、Location パラメタが 0 付近の項目に、Slope パラメタが低い項目があるために生じている（第5章 Figure2 参照）。そのため、Slope パラメタが高いいくつかの項目は、どのような目的の尺度においても重複して選ばれる可能性が高くなり、尺度間で大きな差は生じにくくなる。言い換えれば、目的に応じた多様な短縮版を作成することを目指す場合には、より多くの項目から構成される尺度や項目プールを基にする必要があると言えるだろう。本論文の冒頭でも、尺度の乱立が指摘されていることについて触れたが、もし類似した尺度が複数存在するのであれば、それらを整理し、まとめることによって、そういった項目プールを作成することも可能であるように思われる。この点も今後の課題の一つである。

最後に、IRT 以外の方法も含めて、短縮版作成における今後の課題を指摘したい。まず、本研究では、短縮版作成のための方法として IRT を中心に検討を行っていた。そこから示唆できることは、ここまで述べてきたように、因子分析だけでなく IRT を用いた方法も取り入れて短縮版作成を行うことが有効であること、特に Location パラメタの値も参考に項目選択を行うことで、目的に応じた短縮版の作成が可能になりうること、である。しかし、IRT 以外の他の方法については、まだ十分な検討を行うことができていない。そのため、IRT を含めて多様な指標をもとに短縮版を作成すべきであるという以上の、明確な短縮版作成のガイドライン等は、現時点では十分に示すことができていないと言える。この点は非常に大きな課題である。実際、第2章では、今回の研究で扱った IRT と因子分析だけでなく、項目・合計相関や  $\alpha$  係数をもとに項目選択を行ったり、回答が特定の選択枝に集中した項目を除外したりすることによって短縮版を作成する研究も見られていた（第2章 Table1）。これら多様な方法が、それぞれどのような特徴を持った短縮版を構成することに寄与するのか、どのような場合にこういった指標が特に有効に機能するのか等については、今後検討していく必要があるだろう。

ここまで、本研究では IRT に焦点を当てた上で、短縮版作成について検討を行ってきた。しかし、IRT を適用した分析を行うためには、IRT に関する専門の知識はもちろんのこと、パラメタの推定を行うためには、因子分析等に比べて多くのデータ数が必要であったり、専用のソフトウェアが必要であったりするという面もあるように思われる。そのため、実際に利用するためのハードルは決して低くはないと言えるかもしれない。しかしながら、IRT については、近年フリーのソフトウェアである R によって分析を行う環境等も整ってきている（たとえば、加藤・山田・川端，2014）ことに加え、熊谷（2009）による

「EasyEstimation」シリーズのように、できるだけ簡単な操作で分析が行えることを目指したソフトウェアも開発されている。また、IRTについて解説した書籍も多数刊行されている（たとえば、野口・大隅，2014）。そのため、心理尺度短縮版の開発に限らず、心理学のさまざまな研究にも、今後ますます適用が広がっていくことが期待される。

## 引用文献



- Albert, S. (1977). Temporal comparison theory. *Psychological Review*, **84**, 485-503.
- Alea, N., & Bluck, S. (2007). I'll keep you in mind: The intimacy function of autobiographical memory. *Applied Cognitive Psychology*, **21**, 1091-1111.
- 有光 興記 (1999). スポーツにおける"心構え"尺度の開発 性格心理学研究, **7**, 77-87.
- Beck, A. T. (1967). *Depression : Clinical, experimental, and theoretical aspects*. New York: Harper & Row.
- Birleson, P. (1981). The validity of depression disorders in childhood and the development of a self-rating scale: A research report. *Journal of Child Psychology and Psychiatry*, **22**, 73-88.
- Buss, AH. & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, **63**, 452-459.
- Buunk, A. P., & Gibbons, F. X. (2006). Social comparison orientation: A new perspective on those who do and those who don't compare with others. In Guimond, S. (Ed.) , *Social comparison and social psychology*. NewYork: Cambridge University Press. pp.15-32.
- Buunk, B. P., Nauta, A., & Molleman, E., (2005). In search of the true group animal: The effects of affiliation orientation and social comparison orientation upon group satisfaction. *European Journal of Personality*, **19**, 69-81.
- Buunk, B. P., Oldersma, F. L., & de Dreu, C. K. W. (2001). Enhancing satisfaction through downward comparison: The role of relational discontent and individual differences in social comparison orientation. *Journal of Experimental Social Psychology*, **37**, 452-467.
- Buunk, B. P., Ybema, J. F., Gibbons, F. X., & Ipenburg, M. (2001). The affective consequences of social comparison as related to professional burnout and social comparison orientation. *European Journal of Social Psychology*, **31**, 337-351.
- Buunk, B. P., Ybema, J. F., Van der Zee, K., Schaufeli, W. B., & Gibbons, F. X.(2001). Affect generated by social comparisons among nurses high and low in burnout. *Journal of Applied Social Psychology*, **31**, 1500-1520.
- Buunk, B. P., Zurriaga, R., Gonzalez-Roma, V., & Subirats, M., (2003). Engaging in upward and downward comparisons as a determinant of relative deprivation at work: A longitudinal study. *Journal of Vocational Behavior*, **62**, 370-388.
- Buunk, B. P., Zurriaga, R., Peiro, J. M., Nauta, A., & Gosalvez, I. (2005). Social comparisons at work as related to a cooperative social climate and to individual differences in social comparison orientation. *Applied Psychology*, **54**, 61-80.
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO personality inventory (NEO-PI-R) and NEO Five Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

- Coste, J., Guillemin, F., Pouchot, J., & Fermanian, J. (1997). Approaches to shortening composite measurement scales. *Journal of clinical epidemiology*, **50**, 247-252.
- 傳田 健三・賀古 勇輝・佐々木 幸哉・伊藤 耕一・北川 信樹・小山 司 (2004). 小・中学生の抑うつ状態に関する調査 —Birleson 自己記入式抑うつ評価尺度 (DSRS-C) を用いて—. 児童精神医学とその近接領域, **45**, 424-436.
- Edwards, P., Roberts, I., Clarke, M., DiGuseppi, C., Pratap, S., Wentz, R., & Kwan, I. (2002). Increasing response rates to postal questionnaires: systematic review. *BMJ (Clinical research ed.)*, **324**, 1183-1185.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologist*. Lawrence Erlbaum associates, New Jersey
- 榎本 博明 (2006). 人格心理学領域における研究動向 教育心理学年報, **45**, 61-71.
- 榎本 博明 (2013). 人格心理学領域における研究動向—その現状と課題— 教育心理学年報, **52**, 34-45.
- Fayers, P., & Machin, D. (2007). *Quality of life: The assessment, analysis and interpretation of patient-reported outcomes (2nd ed.)*. Wiley, West Sussex, England
- Festinger, L.(1954). A theory of social comparison processes. *Human Relations*, **7**, 117- 140.
- FFPQ 研究会 (1998). FFPQ (5 因子性格検査) 北大路書房
- Franke, G. R., Rapp, A., & Andzulis, J. “Mick”. (2013). Using shortened scales in sales research: Risks, benefits, and strategies. *Journal of Personal Selling and Sales Management*, **33**, 319–328.
- 藤島 寛・山田 尚子・辻 平治郎 (2005). 5 因子性格検査短縮版 (FFPQ-50) の作成 パーソナリティ研究, **13**, 231-241.
- 福田 一彦・小林 重雄 (1973). 自己評価式抑うつ性尺度の研究 精神神経学雑誌, **75**, 673-679.
- 福井 里江・原谷 隆史・外島 裕・島 悟・高橋 正也・中田 光紀・深澤 健二・大庭 さよ・佐藤 恵美・廣田 靖子 (2004). 職場の組織風土の測定: 組織風土尺度 12 項目版 (OCS-12) の信頼性と妥当性 産業衛生学雑誌, **46**, 213-222.
- Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, **73**, 349–360.
- Gibbons, F. X., & Buunk, B. P. (1999). Individual differences in social comparison: Development and validation of a measure of social comparison orientation. *Journal of Personality and Social Psychology*, **76**, 129-142.
- Gibbons, F. X., & Gerrard, M., (1995). Predicting young adult's health risk behavior. *Journal of Personality and Social Psychology*, **69**, 505-517.

- Gibbons, F. X., Lane, D. J., Gerrard, M., Pomery, E., & Lastrup, C. L. (2002). Drinking and driving: A prospective assessment of the relation between risk cognitions and risk behavior. *Risk Decision and Policy*, **7**, 267-283.
- Goetz, C., Coste, J., Lemetayer, F., Rat, A.-C., Montel, S., Recchia, S., Debouverie, M., Pouchot, J., Spitz, E., & Guillemin, F. (2013). Item reduction based on rigorous methodological guidelines is necessary to maintain validity when shortening composite measurement scales. *Journal of Clinical Epidemiology*, **66**, 710-718.
- Herzog, A. R., & Bachman, J. G. (1981). Effects of questionnaire length on response quality. *Public Opinion Quarterly*, **45**, 549-559.
- 堀 洋道 (監修)・松井 豊・宮本 聡介 (編) (2011). 心理測定尺度集VI 現実社会とかわるく集団・組織・適応> サイエンス社
- 堀 洋道 (監修)・山本 真理子 (編) (2001). 心理測定尺度集 I 人間の内面を探る<自己・個人内過程> サイエンス社
- 稲垣 宏樹・井藤 佳恵・佐久間 尚子・杉山 美香・岡村 毅・栗田 圭一 (2013). WHO-5 精神健康状態表簡易版 (S-WHO-5-J) の作成およびその信頼性・妥当性の検討 日本公衆衛生雑誌, **60**, 294-301.
- 石井 香織・井上 茂・大谷 由美子・小田切 優子・高宮 朋子・下光 輝一 (2009). 簡易版運動習慣の促進要因・阻害要因尺度の開発 体力科学, **58**, 507-516.
- 伊藤 拓・上里 一郎 (2001). ネガティブな反すう尺度の作成およびうつ状態との関連性の検討 カウンセリング研究, **34**, 31-42.
- 岩城 直子 (2008). 「看護における社会的スキル」尺度短縮版の作成の試み 富山大学看護学会誌, **8**, 213-31.
- 上地 雄一郎・宮下 一博 (2009a). 自己愛的脆弱性尺度の妥当性の検討——友人関係への影響の検討を通して—— 岡山大学大学院教育学研究科研究集録, **140**, 1-6.
- 上地 雄一郎・宮下 一博 (2009b). 対人恐怖傾向の要因としての自己愛的脆弱性, 自己不一致, 自尊感情の関連性 パーソナリティ研究, **17**, 280-291.
- 加藤 健太郎・山田 剛史・川端 一光 (2014). Rによる項目反応理論 オーム社
- 数見 隆生・土井 豊・伊藤 常久 (2009). 宮城教育大学学生のジェンダー意識の現状と課題——一般大学生との比較調査から—— 宮城教育大学紀要, **44**, 109-123.
- 工藤 恵理子・遠藤 由美 (2007). 過去の自己は現在の自己評価維持のためにつくられるのか 日本社会心理学会第48回大会大会発表論文集, 88-89.
- 熊谷 龍一 (2009). 初学者向けの項目反応理論分析プログラム EasyEstimation シリーズの開発 日本テスト学会誌, **5**, 107-118.

- 熊野 宏昭・織井 優貴子・山内 祐一・瀬戸 正弘・上里 一郎・坂野 雄二・宗像 正徳・吉永 馨・佐々木 直・久保木 富房 (2000). Short Interpersonal Reactions Inventory 日本語短縮版作成の試み (第2報) ——33項目版への改訂—— 心身医学, **40**, 447-454.
- 黒田 祐二・桜井 茂男 (2001). 子どもの抑うつ研究の概観 筑波大学心理学研究, **23**, 129-138.
- Lennox, R. D., & Wolfe, R. N. (1984). Revision of the Self-Monitoring Scale. *Journal of Personality and Social Psychology*, **46**, 1349-1364.
- 森 慶輔・三浦 香苗 (2007). 職場における短縮版ソーシャルサポート尺度の開発と信頼性・妥当性の検討 ——公立中学校教員への調査を基に—— 昭和女子大学生活心理研究所紀要, **9**, 74-88.
- 村上 宣寛・村上 千恵子 (1999). 主要5因子性格検査の世代別標準化性格心理学研究, **8**, 32-42.
- Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, **16**, 159-176.
- Muraki, E., & Bock, R. D. (2003). PARSCALE 4: IRT item analysis and test scoring for rating-scale data [Computer program]. Lincolnwood, IL: Scientific Software.
- 村田 豊久 (1993). 小児期のうつ病 臨床精神医学, **22**, 557-563.
- 村田 豊久・清水 重紀・森 陽二郎・大島 祥子 (1996). 学校における子どものうつ病 —Birlesonの小児期うつ病スケールからの検討— 最新精神医学, **1**, 131-138.
- 村山 航 (2012). 妥当性 概念の歴史的変遷と心理測定学的観点からの考察 教育心理学年報, **51**, 118-130.
- 内藤 まゆみ・鈴木 佳苗・坂元 章 (2004). 情報処理スタイル (合理性—直観性) 尺度の作成 パーソナリティ研究, **13**, 67-78.
- 並川 努 (2010). 継時的比較志向性尺度短縮版の作成——Item Response Theoryを用いた検討—— 名古屋大学大学院教育発達科学研究科紀要 (心理発達科学), **57**, 71-76.
- 並川 努 (2011). 継時的比較の個人差——継時的比較志向性尺度の作成と検討—— 心理学研究, **81**, 593-601.
- 並川 努 (投稿中). 心理尺度短縮版の作成における現状と課題
- 並川 努・谷 伊織・脇田 貴文・熊谷 龍一・中根 愛・野口 裕之 (2012). Big Five 尺度短縮版の開発と信頼性と妥当性の検討 心理学研究, **83**, 91-99.
- 並川 努・谷 伊織・脇田 貴文・熊谷 龍一・中根 愛・野口 裕之・辻井 正次 (2011). Birleson 自己記入式抑うつ評価尺度 (DSRS-C) 短縮版の作成 精神医学, **53**, 489-496.
- 並川 努・脇田 貴文・野口 裕之 (2010). 日本における自尊感情尺度使用の状況と課題 教育心理学フォーラムレポート, FR-2010-01.
- 野口 裕之・大隅 敦子 (2014). テスティングの基礎理論 研究社

- 野村 信威・橋本 宰(2001). 老年期における回想の質と適応との関連 発達心理学研究, **12**, 75-86.
- 野村 豊子 (1998). 回想法とライフレビュー——その理論と技法—— 中央法規出版
- 野崎 秀正 (2006). 生徒の社会的比較志向性と学業的援助要請の関連 宮崎女子短期大学紀要, **32**, 113-122.
- 大野木 裕明 (2004). 主要 5 因子性格検査 3 種間の相関的資料 パーソナリティ研究, **12**, 82-89.
- 奥村 泰之・亀山 晶子・勝谷 紀子・坂本 真士 (2008). 1990 年から 2006 年の日本における抑うつ研究の方法に関する検討 パーソナリティ研究, **16**, 238-246.
- 長田 由紀子・長田 久雄 (1994). 高齢者の回想と適応に関する研究 発達心理学研究, **5**, 1-10.
- 押見 輝男・渡辺 浪二・石川 直弘(1985) . 自己意識尺度の検討 立教大学心理学科研究年報, **28**, 1-15.
- Radloff, L. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, **1**, 385-401.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton: Princeton University Press.
- Ross, M., & Wilson, A. E. (2002). It feels like yesterday: Self-esteem, valence of personal past experiences, and judgments of subjective distance. *Journal of Personality and Social Psychology*, **82**, 792-803.
- Roznowski, M. (1989). Examination of the measurement properties of the Job Descriptive Index with experimental items. *Journal of Applied Psychology*, **74**, 805-814.
- 坂井 明子・山崎 勝之・曾我 祥子・大芦 治・島井 哲志・大竹 恵子 (2000). 小学生用攻撃性質問紙の作成と信頼性, 妥当性の検討 学校保健研究, **42**, 423-433.
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometric Monograph*, **17**.
- 笹川 智子・金井 嘉宏・村中 泰子・鈴木 伸一・島田 洋徳・坂野 雄二 (2004). 他者からの否定的評価に対する社会的不安測定尺度 (FNE) 短縮版作成の試み—項目反応理論による検討— 行動療法研究, **30**, 87-98.
- 佐々木 淳・星野 崇宏・丹野 義彦 (2002). 精神病理の症状と性格 5 因子モデルとの関係 教育心理学研究, **50**, 65-72 .
- 佐藤 徳 (2008). 想起された出来事の時間的距離判断ならびに所属判断に影響を及ぼす要因の検討 パーソナリティ研究, **16**, 416-425.
- 佐藤 寛・新井 邦二郎 (2002). 子ども用抑うつ自己評価尺度 (DSRS) の因子構造の検討と標準データの構築 発達臨床心理学研究, **14**, 85-91.

- 佐藤 寛・下津 咲絵・石川 信一 (2008). 一般中学生におけるうつ病の有病率 半構造化面接を用いた実態調査 精神医学, **50**, 439-448.
- Saxe, R., & Weitz, B. A. (1982). The SOCO Scale: A Measure of the Customer Orientation of Salespeople. *Journal of Marketing Research*, **19**, 343-351.
- 清水 健司・川邊 浩史・海塚 敏郎 (2006). 対人恐怖心性-自己愛傾向 2 次元モデル尺度における短縮版作成の試み パーソナリティ研究, **15**, 67-70.
- 下仲 順子・中里 克治・権藤 恭之・高山 緑 (1999). NEO-PI-R, NEO-FFI 共通マニュアル 東京心理
- 白澤 早苗・石田 多由美・箱田 裕司・原口 雅浩 (1999). 記憶検索に及ぼすエネルギー覚醒の効果 基礎心理学研究, **17**, 93-99.
- Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*, **30**, 526-537.
- 曾我 祥子 (1983). 日本版 STAIC 標準化の研究 心理学研究, **54**, 215-221.
- Stanton, J. M., Sinar, E. F., Balzer, W. K., & Smith, P. C. (2002). Issues and strategies for reducing the length of self-report scales. *Personnel psychology*, **55**, 167-194.
- Suls, J., & Mullen, B. (1982). From the cradle to the grave: Comparison and self-evaluation across the life-span. In J. Suls. (Ed), *Psychological perspectives on the self*. Hillsdale. NJ: Lawrence Erlbaum Associates. pp. 97-125.
- Suls, J., & Wheeler, L. (Eds.) (2000). *Handbook of social comparison: Theory and research*. New York: Kluwer Academic/Plenum.
- 鈴木 淳子 (1994). 平等主義的性役割態度スケール短縮版 (SESRA-S) の作成 心理学研究, **65**, 34-41.
- 鈴木 平・依田 麻子・越川 房子・杉若 弘子・嶋田 洋徳・瀬戸 正弘・上里 一郎 (1997). Life Style Index の日本語短縮版の作成および標準化の試み 健康心理学研究, **10**, 31-43.
- 高田 利武 (1992). 他者と比べる自分 サイエンス社
- 高田 利武 (1993). 青年の自己概念形成と社会的比較——日本人大学生にみられる特徴—— 教育心理学研究, **41**, 339-348.
- 谷 伊織・吉橋 由香・神谷 美里・宮地 泰士・野村 香代・伊藤 大幸・辻井 正次 (2010). 抑うつと特性不安から見た小中学生の精神的健康の構造的検討 精神医学, **52**, 265-273.
- Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L.Berkowitz(Ed.), *Advances in experimental social psychology* Vol.21. NewYork: Academic Press. pp.181-227.
- 友寄 令子・佐藤 公代 (2006). 社会的比較志向性に関する研究——その規定因と大学生生活適応への影響—— 愛媛大学教育実践総合センター紀要, **24**, 97-106.

- 外山 美樹 (2002). 社会的比較志向性と心理的特性との関連——社会的比較志向性尺度を作成して—— 筑波大学心理学研究, **24**, 237-244.
- 内田 照久 (2005). 音声中の抑揚の大きさと変化パターンが話者の性格印象に与える影響 心理学研究, **76**, 382-390.
- 内田 照久 (2012). 教育評価・心理測定で用いる測度の妥当性検証の機運と社会的役割を担う試験をめぐる課題解決への取り組み 教育心理学年報, **51**, 63-72.
- 和田 さゆり (1996). 性格特性用語を用いた Big Five 尺度の作成 心理学研究, **67**, 61-67.
- 脇田 貴文 (2004). 評定尺度法におけるカテゴリ間の間隔について——項目反応モデルを用いた評価方法—— 心理学研究, **75**, 331-338.
- 脇田 貴文・小塩 真司・願興寺 礼子・桐山 雅子 (2007). University Personality Inventory 短縮版の開発 人文学部研究論集, **17**, 123-128.
- Wheeler, L. (2000). Individual differences in social comparison. In J. Suls., & L. Wheeler. (Eds.), *Handbook of social comparison: Theory and research*. New York: Kluwer Academic/Plenum. pp. 141-158.
- Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin*, **90**, 245-271.
- Wilson, A. E., & Ross, M. (2000). The frequency of temporal-self and social comparisons in people's personal appraisals. *Journal of Personality and Social Psychology*, **78**, 928-942.
- Wilson, A. E., & Ross, M. (2001). From chump to champ: People's appraisals of their earlier and present selves. *Journal of Personality and Social Psychology*, **80**, 572-584.
- 山口 陽弘 (2011). 心理尺度に着目したこの一年の概観——パーソナリティ研究とは「心理尺度づくり」なのだろうか？—— 教育心理学年報, **50**, 97-107.
- 山本 真理子・松井 豊・山成 由紀子 (1982). 認知された自己の諸側面の構造 教育心理学研究, **30**, 64-68.
- 吉田 寿夫・石井 秀宗・南風原 朝和 (2012). 尺度の作成・使用と妥当性の検討 教育心理学年報, **51**, 213-217.
- Zung, W. W. K. (1965). A self-rating depression scale. *Archives of General Psychiatry*, **12**, 63-70.

謝辭



本研究をまとめるまでの間に、多くの皆様からご指導やご支援をいただきました。特に指導教官である名古屋大学大学院教育発達科学研究科心理発達科学専攻心理社会行動科学講座教授野口裕之先生からは大変御多忙な中、長期間に渡って丁寧な御指導を賜りました。深く感謝申し上げます。

また、研究室の先輩である東北大学大学院教育学研究科准教授熊谷龍一先生、関西大学社会学部准教授脇田貴文先生、東海学園大学人文学部准教授谷伊織先生をはじめ、多くの方々からも、研究を進めるさまざまな場面でご指導ご鞭撻を賜りました。一人ひとりのお名前を記すことができないのは大変残念ですが、お世話になった皆さまにあらためて深く御礼申し上げます。

## 付録

## Appendix 1

## 第2章 分析対象文献リスト

- 阿部 恒之・鈴木 ゆかり・平田 祐子 (1989). エステティックフェーシャルマッサージの心理生理学的研究——リラクゼーション効果の検証 日本化粧品技術者会誌, **22**, 236-244.
- 荒井 由美子・田宮 菜奈子・矢野 栄二 (2003). Zarit 介護負担尺度日本語版の短縮版 (J-ZBI\_8)の作成: その信頼性と妥当性に関する検討 日本老年医学会雑誌, **40**, 497-503.
- 有光 興記 (1999). スポーツにおける"心構え"尺度の開発 性格心理学研究, **7**, 77-87.
- 江口 圭一・戸梶 亜紀彦 (2009). 労働価値観測定尺度(短縮版)の開発 実験社会心理学研究, **49**, 84-92.
- 藤島 寛・山田 尚子・辻 平治郎 (2005). 5 因子性格検査短縮版 (FFPQ-50) の作成 パーソナリティ研究, **13**, 231-241.
- 福井 里江・原谷 隆史・外島 裕・島 悟・高橋 正也・中田 光紀・深澤 健二・大庭 さよ・佐藤 恵美・廣田 靖子 (2004). 職場の組織風土の測定: 組織風土尺度 12 項目版 (OCS-12) の信頼性と妥当性 産業衛生学雑誌, **46**, 213-222.
- 橋本 剛 (1997). 対人関係が精神的健康に及ぼす影響——対人ストレス生起過程因果モデルの観点から 実験社会心理学研究, **37**, 50-64.
- 原田 謙・杉澤 秀博・杉原 陽子・山田 嘉子・柴田 博 (2004). 日本語版 Fraboni エイジズム尺度 (FSA) 短縮版の作成: 都市部の若年男性におけるエイジズムの測定 老年社会科学, **26**, 308-319.
- 広沢 俊宗 (2009). 学習技術、および学習特性に関する尺度化の研究 (1) —大学生用簡易版尺度の作成— 教育総合研究叢書, **2**, 71-82.
- 今津 芳恵・村上 正人・小林 恵・松野 俊夫・椎原 康史・石原 慶子・城 佳子・児玉 昌久 (2006). Public Health Research Foundation ストレスチェックリスト・ショートフォームの作成: 信頼性・妥当性の検討 心身医学, **46**, 301-308.
- 石井 香織・井上 茂・大谷 由美子・小田切 優子・高宮 朋子・下光 輝一 (2009). 簡易版運動習慣の促進要因・阻害要因尺度の開発 体力科学, **58**, 507-516.
- 伊藤 美奈子 (1994). 性格特性の一面性と個人志向性・社会志向性との関連について 心理学研究, **65**, 18-24.
- 岩城 直子 (2008). 「看護における社会的スキル」尺度短縮版の作成の試み 富山大学看護学会誌, **8**, 213-31.
- 岩田 香織・岡田 節子・朴 千萬・中嶋 和夫 (2000). 短縮版 UCLA Loneliness Scale の開発 静岡県立大学研究紀要, **14**, 225-233.

- 加賀美 常美代・大淵 憲一 (2006). 教育価値観に関する異文化間比較: 短縮版尺度開発と包括次元の探索 文化, **69**, 305-290.
- 上地 玲子 (2001). 心理的分離尺度短縮版の作成 岡山短期大学紀要, **24**, 69-72.
- 上地 雄一郎・宮下 一博 (2009). 対人恐怖傾向の要因としての自己愛的脆弱性, 自己不一致, 自尊感情の関連性 パーソナリティ研究, **17**, 280-291.
- 蒲原 龍・岡田 栄作・志渡 晃一 (2009). うつ尺度 CES-D 簡易版作成の試み 北海道医療大学看護福祉学部学会誌, **5**, 87-92.
- 金子 真理子・眞嶋 朋子・小泉 晋一・岡本 高宏・佐藤 紀子 (2006). Suffering 調査票の開発 日本看護科学会誌, **26**, 3-12.
- 数見 隆生・土井 豊・伊藤 常久 (2009). 宮城教育大学学生のジェンダー意識の現状と課題——一般大学生との比較調査から—— 宮城教育大学紀要, **44**, 109-123.
- 金城 光・井出 訓・森 伸幸 (2008). 日本版成人メタ記憶尺度 MIA (The Questionnaire for Metamemory in Adulthood) 短縮版開発のための検討 大妻女子大学紀要—社会情報系—社会情報学研究, **17**, 163-173.
- 小林 宏 (2000). 中学生用勤労観短縮版作成の試み 発達心理臨床研究, **7**, 33-37.
- 熊野 宏昭・織井 優貴子・鈴鴨 よしみ・山内 祐一・宗像 正徳・吉永 馨・瀬戸 正弘・坂野 雄二・上里 一郎・久保木 富房 (1999). Short Interpersonal Reactions Inventory 日本語短縮版作成の試み——タイプ C パーソナリティ測定を中心として—— 心身医学, **39**, 335-341.
- 熊野 宏昭・織井 優貴子・山内 祐一・瀬戸 正弘・上里 一郎・坂野 雄二・宗像 正徳・吉永 馨・佐々木 直・久保木 富房 (2000). Short Interpersonal Reactions Inventory 日本語短縮版作成の試み (第 2 報) ——33 項目版への改訂—— 心身医学, **40**, 447-454.
- 三川 俊樹・井上 知子・芳田 茂樹 (1991). 青年期における人格形成と精神的健康に関する研究(V)——性役割および自我同一性地位と役割受容・充実感の関連—— 追手門学院大学文学部紀要, **25**, 51-67.
- 三好 昭子・大野 久・内島 香絵・若原 まどか・大野 千里 (2003). Ochse&Plug の Erikson and Social-Desirability Scale の日本語短縮版 (S-ESDS) 作成の試み 立教大学心理学研究, **45**, 65-76.
- 内藤 まゆみ・鈴木 佳苗・坂元 章 (2004). 情報処理スタイル (合理性—直観性) 尺度の作成 パーソナリティ研究, **13**, 67-78.
- 中川 敦子・木村 由佳・鋤柄 増根 (2009). 乳児の行動のチェックリスト (IBQ-R) 短縮版の作成 人間文化研究, **12**, 15-25.
- 中井 裕子・宮下 光令・笹原 朋代・小山 友里江・清水 陽一・河 正子 (2006). Frommelt のターミナルケア態度尺度 日本語版 (FATCOD-B-J) の因子構造と信頼性の検討——尺度翻訳から一般病院での看護師調査, 短縮版の作成まで がん看護, **11**, 723-729.

- 西野 泰広 (1990). 幼児の自己制御機能と母親のしつけタイプ 発達心理学研究, **1**, 49-58.
- 岡安 孝弘・高山 巖 (1999). 中学生用メンタルヘルス・チェックリスト(簡易版)の作成 宮崎大学教育学部附属教育実践研究指導センター研究紀要, **6**, 73-84.
- 岡安 孝弘・由地 多恵子・高山 巖 (1998). 児童用メンタルヘルス・チェックリスト(簡易版)の作成とその実践的利用 宮崎大学教育学部附属教育実践研究指導センター研究紀要, **5**, 27-41.
- 大竹 恵子・島井 哲志・曾我 祥子 (2001). 小学生のコーピング尺度短縮版の作成 ヒューマンサイエンス, **4**, 1-5.
- 笹川 智子・金井 嘉宏・村中 泰子・鈴木 伸一・島田 洋徳・坂野 雄二 (2004). 他者からの否定的評価に対する社会的不安測定尺度 (FNE) 短縮版作成の試み—項目反応理論による検討— 行動療法研究, **30**, 87-98.
- 佐々木 万丈 (2002). 中学生用体育学習心理的ストレスレベル測定尺度の短縮版の開発と標準化 体育学研究, **47**, 383-394.
- 佐々木 淳・星野 崇宏・丹野 義彦 (2002). 精神病理の症状と性格 5 因子モデルとの関係 教育心理学研究, **50**, 65-72.
- 瀬戸 正弘 (2003). 喫煙動機評価尺度短縮版作成の試み 安田女子大学心理教育相談研究, **2**, 33-40.
- 清水 健司・川邊 浩史・海塚 敏郎 (2006). 対人恐怖心性-自己愛傾向 2 次元モデル尺度における短縮版作成の試み パーソナリティ研究, **15**, 67-70.
- 白澤 早苗・石田 多由美・箱田 裕司・原口 雅浩 (1999). 記憶検索に及ぼすエネルギー覚醒の効果 基礎心理学研究, **17**, 93-99.
- 鈴木 淳子 (1994). 平等主義的性役割態度スケール短縮版 (SESRA-S) の作成 心理学研究, **65**, 34-41.
- 鈴木 平・依田 麻子・越川 房子・杉若 弘子・嶋田 洋徳・瀬戸 正弘・上里 一郎 (1997). Life Style Index の日本語短縮版の作成および標準化の試み 健康心理学研究, **10**, 31-43.
- 鈴木公啓 (2004). 日本語短縮版 Padua Inventory (J-PI32) の作成 精神科診断学, **15**, 15-24.
- 高比良 美詠子 (1998). 対人・達成領域別ライフイベント尺度(大学生用)の作成と妥当性の検討 社会心理学研究, **14**, 12-24.
- 武田 知樹・波多野 義郎 (2008). 在宅脳卒中患者を対象とした短縮版心理的 QOL 質問票作成の試み 日本保健医療行動科学会年報, **23**, 163-171.
- 豊田 弘司・岸田 麻里 (2006). 教育用簡易版恋愛感情尺度の作成 教育実践総合センター研究紀要, **15**, 1-5.
- 植村 善太郎 (2001). あいまいさへの耐性と集団同一性が新入成員への寛容的反応に及ぼす効果 性格心理学研究, **10**, 27-34.

- 若林 明雄・サイモン バロン-コーエン・サリー ウィールライト(2006). Empathizing-Systemizing モデルによる性差の検討——Empathizing 指数 (EQ) と Systemizing 指数 (SQ) による個人差の測定 心理学研究, **77**, 271-277.
- 脇田 貴文・小塩 真司・願興寺 礼子・桐山 雅子 (2007). University Personality Inventory 短縮版の開発 人文学部研究論集, **17**, 123-128.
- 山口 桂子・服部 淳子・中村 菜穂・水野 貴子・小林 督子 (2001). 看護婦用ストレス反応尺度の作成：既成尺度の看護婦への適用と短縮版作成の試み 愛知県立看護大学紀要, **7**, 1-11.

## Appendix 2

## 第2章 補足

第2章では1980年から2009年までの論文を対象に報告を行った。その後、2013年までの結果についても集計を行ったため、その概要を報告する。まず、Table1に示したように、対象となった論文数は、全体で73本であり、2010年から2013年までの4年間で24本となっていた。2010年から2013年は、2009年以前と比べても数は多くなっており、近年も短縮版のニーズは高いことが改めて示唆された。

また、短縮版作成方法について、90年代以前と2000年から2009年の00年代、2010年代に分けて集計を行った結果をTable2に示した。全体的な傾向は、第2章で報告されたものと概ね一致していると考えられる。まず、因子分析が最も多く用いられている傾向は、2010年代も共通して見られていた。また、2010年以前と比べて2010年以降に多く見られた方法としては、 $\alpha$ 係数を用いた方法と、IRTを用いた方法が挙げられる。ただし、IRTについては本研究の第3章から第5章で行われた研究論文が含まれているため、IRTについてはまだ多く一般に普及している方法とは言えない部分がある。

Table1 対象論文の発行時期

時期	論文数
1980～1984年	0
1985～1989年	1
1990～1994年	4
1995～1999年	8
2000～2004年	17
2005～2009年	19
2010～2013年	24
計	73

Table2 主な短縮版作成方法と論文数

	2000～ ～1999年	2009年	2010年～ 2013年	計	%※
質的分析					
	5	17	12	34	45.9%
量的分析					
因子分析	10	32	17	59	79.7%
項目-合計相関	1	9	3	13	17.6%
回答の分布の偏り・集中	2	7	1	10	13.5%
$\alpha$ 係数	1	4	5	10	13.5%
欠測値、未経験等の回答率	1	4	2	7	9.5%
IRT	0	1	4	5	6.8%
G-P分析	0	3	1	4	5.4%
項目間の相関係数	1	1	0	2	2.7%

※ %は、対象となった74本のうち、当該分析が用いられていた割合を示す