

日本版NEO-FFIの信頼性の検討

岩手大学生のデータに項目反応理論を適用

阿久津 洋 巳*

(2017年10月31日 受付, 2018年1月17日 受理)

1. はじめに

人には性格とよばれる比較的安定して持続する行動と思考・感情の特性がある。そのため個人の行動を予測したり、その問題を理解する上で、性格の知識は実際的な価値がある。人の性格を画然と異なるタイプによって分ける方法は、直感に訴えわかりやすい反面、その個人の全体を十分にとらえることはできない。多くの人に共通な特性を扱いやすい数に限定して列挙し、その各共通特性を個人がどの程度もっているかを記述するほうがより正確に個性を把握することができる。性格の共通特性を数え上げたりその量を測定したり、さらに共通特性を何らかの方法で限定する試みを性格心理学の歴史にみることができる(John, Srivastava, 1999)。共通特性を数の点で比較的扱いやすい5つの特性にまとめた性格の5因子(通称Big Five)が、1990年代から今日まで理論と実践の両面で多くの研究者から支持され発展してきた。人の性格について語るときに使える共通言語として Big Five がもつ意義は大きい。

1. Big Five の発見

性格の5因子は、人の性格特性を表現する言葉の研究に起源をもつ(John&Srivastava, 1999)。Allport と Odbert は、大辞典の見出し語で人物を記述できる単語を数え、それが1万8千語以上あることを見出した(Allport&Odbert, 1936)。Cattell (1943, 1945a, 1945b) は Allport&Odbert のリストから4500語を選び、そのリストを35に縮減した。これは主に当時利用可能であった因子分析法では大量の変数を扱うことに制限があったためである。Cattellはこの35語に因子分析を適用し、12の性格因子を見出し、最終的にはかの要因を勘案して16性格因子の質問紙を作成した(Cattell, Eber, & Tatsuoka, 1970)。Cattellの変数のうち22を使って分析を進め、Fiske (1949) は5因子に類似した因子に達した。これらの因子をさらに明確にするために Tupes&Christal (1961) は様々な調査対象集団を用いデータ分析(相関行列の再分析)を進め、5つの因子を確認した。後に他の研究者たちが Cattell の35の変数リストから同様な5因子を見出し(Norman, 1967 ;

* 岩手大学教育学部

Borgatta, 1964 ; Digman&Takemoto-Chock, 1981), 今日の Big Five が成立した。

2. Eysenck の2つの特性

性格特性の考えを一般に広め、特性が一般人口と病的傾向をもつ個人の理解に有効であることを実証した研究者が Eysenck である (Eysenck, 1970)。Eysenck は性格に関連する多くの質問項目に対する回答を多数集め、その回答に因子分析の手法を適用して、最終的に2つの因子を見出した。これが, introversion-extraversion (内向-外向) と neuroticism (神経症的傾向) の2大特性である。個人差は、これらの特性を多くもっているか少なく持っているかの違いとして表すことができる。この2つの特性間に相関はないので、これらを水平と垂直の直交する2次元座標軸として座標平面を作ると全ての人をその平面のどこかに位置づけることができる。Eysenck は、後に psychoticism (精神病傾向) を付け加えて、3因子とした。Eysenck はまた、自分が見つけた特性の背後に生物学的基盤を想定していた。内向-外向に関しては、大脳皮質の覚醒度を生物学的基礎と考えた。この考えを支持する証拠は多数ある (Eysenck, 1990)。神経症的傾向との生物学的基礎として自律神経の機能を想定したが、支持する証拠は多くはない。精神病的傾向に関しては、攻撃性の程度はテストステロンの濃度に影響され、男性では高いことが示唆された (Eysenck, 1990)。Eysenck の2因子については、それ以降の研究は繰返し2因子の存在を確認しているが、2因子だけでは性格特性を記述するのに十分ではないという議論は続いた。Big Five には Eysenck の2因子が含まれており、その意味で Big Five は Eysenck の理論の延長上にあるという見方もできる。Big Five の理論家は、Eysenck と同じく性格特性の生物学的基盤を想定しているが実証にはまだ多くの道のりが予想される。

3. Big Five の特性

性格の5因子で仮定される特性は、しばしば OCEAN と略記される。これは (1) openness 開放性 (2) conscientiousness 誠実性 (3) extraversion 外向性 (4) agreeableness 調和性 (5) neuroticism 神経症傾向の5つである。英語の特性名は、統一されているが、日本語の訳語は統一されていない。本論文では、日本語版 NEO-PI-R と NEO-FFI で使われている訳語を使う (下仲・中里・権藤・高山, 1999)。これらの5つの特性のうち外向性、神経症傾向は Eysenck が主張した2つの特性と同じものである。開放性は、多くの調査からみるとやや不安定な特性である。しかしながら、これら5つの特性は様々な国と文化で繰返し見つかっている特性であり、特定の文化に限定されない人類に共通の特性といえる。

4. Big Five の質問紙

Big Five の特性は、特性の測定とともに発展してきた。Cattell たちの16因子の検討から始めた Costa&McCrae は、神経症傾向、外向性、開放性の3特性を確認し、それらを測定するために NEO Personality Inventory (Costa&McCrae, 1985) を開発した。次いで、この NEO Personality Inventory に調和性と誠実性をふくめて、Costa&McCrae (1992) は240項目からなる NEO 人格目録修正版 (NEO Personality Inventory, Revised, NEO PI-R) を発表した。この質問紙は、5因子について各6つの側面 (facets) を測定できる。NEO PI-R は

中年とより高齢の成人を標本として、因子分析と複数の妥当性検証を経て作成された。内的一貫性、時間的安定性、輻輳妥当性と弁別妥当性をもつ。さらに広い範囲の言語と文化において5因子の30側面の尺度が再現された (McCrae & Costa, 1997)。NEO PI-R の240項目は、項目数が多いため実施に時間がかかるという実用上の欠点があるため、それを補う目的で短縮版のNEO-FFIが開発された (Costa & McCrae, 1992)。NEO-FFIは各因子に12項目を配置しているが、各因子に6つの側面という構成はもっていない。この12項目はNEO PI-Rの240項目のうち各因子で高い因子負荷を得た項目を集めている。尺度の信頼性は、 α 係数の平均0.78であり適切な範囲である。

Big Fiveの質問紙としては、NEO PI-Rのほかには、GoldbergのTrait Descriptive Adjectives (TDA) (Goldberg, 1990, 1992) と John, Donahue, & KentleのBig Five Inventory (BFI, 1991)がある。TDAは100の単極の特性形容詞のリストからなる。高い内的一貫性と頑健な因子構造を持つ。形容詞のみを提示するため、解釈にあいまいさがあるため一貫した反応を得にくいという指摘がある (John & Srivastava, 1999)。この欠点を修正し、さらに短時間で効果的にBig Fiveを測定する目的で開発されたのが44項目の検査項目からなるBFIである。形容詞のかわりに特性形容詞に基づいた短い語句を使う。例えば、誠実性の"persevering"のかわりに"perseveres until the task is finished"のような具合である。 α 係数からみた信頼性は平均0.8と適切な範囲である。輻輳の妥当性も確認されている。

5. NEO-FFI の検討

本研究は、日本版の短縮版NEO-FFIに注目した。英語版のNEO PI-Rは240項目という多数の項目を含むために使いにくい、優れた尺度である。その日本語版である日本版NEO-PI-Rにもとづいて作成された日本版NEO-FFIは60項目からなり、短時間で実施できるため使いやすい。

使用マニュアルにもとづいて日本版NEO-PI-Rの作成過程を要約すると、18～87歳の男女920名を対象としてデータを得ている (下仲・中里・権藤・高山, 1999)。そのうち大学生は18～22歳であり362名が調査に参加した。 α 係数から見た信頼性は、5つの特性で0.86～0.92と高い。検査-再検査の信頼性も5つ特性で0.84～0.91と高い。5因子内の6つの側面別に因子負荷量を日本語版NEO-PI-Rの結果とオリジナルのNEO PI-Rの結果と比較したところ、おおむね似通った結果であった。これらの結果から、日本版NEO-PI-Rが適切に性格の5因子を測定できる尺度であることが理解できる。

日本版NEO-PI-Rの短縮版である日本版NEO-FFIは、英語版のNEO-FFIと同様の手続きを踏んで作成された。具体的には、日本版のNEO-PI-Rから各次元について12項目を選択し60項目としている。その信頼性と妥当性は、日本版NEO-PI-Rの240項目を実施して得られたデータから日本版NEO-FFIに相当する項目のデータを選びだして検討された。成人群659名のデータが使われた。またそのうち大学生356名は大学生群としてまとめられている。大学生のデータに限ると356名は標準化のために十分なサンプルサイズとはいえないが、一応の目安を与えるには十分なサンプルサイズと考えられる。使用マニュアルには、パーセントイル換算表が含まれており、利用者は素点から標準得点を計算することができる。日本版NEO-FFI独自の調査を十分に大きいサンプルサイズに対して実施すれば、

実用上より有用な尺度にできる。また必要によっては項目数を縮減し、より簡便な尺度にできる可能性もある。

本研究は、日本版NEO-FFIについて岩手大学の学生を母集団と想定して、多数の岩手大学生を調査参加者として、因子構造と質問項目を検討した。加えて、項目反応理論 (Item Response Theory, IRT) を適用して尺度の信頼性を検討した。尺度作成における項目反応理論の有用性については、多くの文献があるのでここでは省略する (例えば、豊田, 2002)。

2. 方法

1. 調査

調査1 調査期間 2012年11月から2013年4月の間にデータを収集した。

調査方法 岩手大学の農, 工, 人文社会, 教育学部の学科から比較的受講者が多いと期待できる授業科目をランダムに選択し, 授業時間を使って調査を実施した。NEO-FFIの市販テストの質問項目を実施者が読み上げ, 参加者は筆者が準備した回答用紙に回答を記入した。NEO-FFIの質問項目はNEO-FFIの出版元の東京心理の許可を得て使用した。

調査参加者 岩手大学学生553人が調査に参加した。未回答項目を含む参加者を除いた511人のデータを分析対象とした。学部別の参加人数を調査2とあわせて表1に示した。

調査2 調査期間 2015年10月から2016年5月の間にデータを収集した。

調査方法 授業時間を使って調査を実施した。NEO-FFIの市販テストを配布して, 参加者自身で質問項目を黙読して答える標準的方法で実施した。岩手大学学生422人が調査に参加した。未回答項目を含む参加者を除いた389人のデータを分析対象とした。学部別の参加人数を調査1とあわせてTable 1に示した。

調査1と調査2を合計すると900人が参加した。質問紙に答えた男女の割合 (男子510人, 女子390人), および学部・学科の学生比を大学が公表している学生数と比較して, 適切な比率であることを確認した。2012年当時岩手大学学生数は全学で5164であった。学部学生の実際の比率は, 工学部0.38, 農学部0.20, 教育学部0.23, 人文社会学部0.19であったので, 参加者はその比率とほぼ同じであった。調査1, 2の参加者数, 参加者の学部・男女別内訳, 参加者の学部別人数と割合をTable 1, 2, 3に掲載した。

Table 1 調査1と調査2の参加者数

	工学部	農学部	教育学部	人文社会学部
調査1	305	98	75	33
調査2	3	76	129	181

Table 2 参加者の内訳

	工学部	農学部	教育学部	人文社会学部
男子	280	75	87	68
女子	28	99	117	146
平均年齢	19.8	19.4	18.9	19.0

Table 3 参加者の学部別人数と割合

総人数	工学部	農学部	教育学部	人文社会学部
900	308	174	204	214
割合	0.342	0.193	0.227	0.238

2. データの分析方法

調査1と調査2の参加者の回答を統合して、通常の尺度構成の手順に従ってデータの分析を進めた。具体的には、まず(1)因子分析により因子数の確認と因子に含まれる質問項目を確認した。次いで、(2)因子分析結果に基づいて質問項目を5つの因子に分類したのちに、因子別に回答データに項目反応理論を適用した。データの分析にはRを使用した。項目反応理論を適用するためには、Rの ltm package (Rizopoulos, 2007) を使用した。

3. 結果と考察

1. 因子分析による質問項目の確認

スクリープロットと第1から第6までの主成分の数値を参考にして、5因子に決定した。第1因子から第6因子の固有値は、順に2.836, 2.301, 2.050, 1.897, 1.711, 1.367であった。

因子数を5因子として、最尤法を用い promax 回転を行う因子分析を実行した。因子負荷量の基準を0.2として、この基準以下の項目を選び削除した。因子負荷量の基準を0.2と通常より低くした理由は、因子に分類する段階では多くの項目拾い上げて、その後の項目反応理論を使った分析で不適切な項目を除外するためである。以下質問項目に言及する際は、Q1, Q2のように、Qを質問項目、その後の数字を項目番号とする。1回目の因子分析結果によって削除された項目は、Q8, Q18, Q28, Q38であった。これら削除された4項目はすべて開放性の項目であった。続けて、これらの質問項目を除いたデータに対して再度因子分析を行ったところ、上述の基準から削除対象の項目は無く、56項目が残った。この56項目を分析対象とした。第1因子(神経症傾向)は13項目、第2因子(外向性)は15項目、第3因子(開放性)は7項目、第4因子(調和性)は10項目、第5因子(誠実性)は11項目であった。因子名と質問項目番号を日本版NEO-FFI使用マニュアルに掲載されている質問項目の分類とあわせてTable4に示した*。Table5には56質問項目の因子負荷量と日本版NEO-FFIの分類特性を示した。因子分析結果を理解しやすくするため、5因子別に、一番因子負荷量が大きかった項目と日本版NEO-FFIの分類とは異なる項目の内容をTable6に掲載した。56質問項目の全データの分散に対する5因子による説明率は、0.30であった。また、因子間の相関は、0.01～-0.29程度であり(Table7)、内的整合性の目安となるCronbackの α 係数とMcDonaldの ω 係数を計算したところ、実質的に同じであったため、 α 係数のみ報告すると、神経症傾向0.83, 外向性0.82, 開放性0.69, 調和性0.63, 誠実性0.77であった。開放性と調和性の信頼性がやや低かった。

*質問項目を全て明示すると結果を理解しやすいが、著作権を侵害するので、本論文では一部のみ明示する。

Table 4 5因子の質問項目：日本版マニュアルの質問項目分類と本研究の結果

N (神経症傾向)		E (外向性)		O (開放性)		A (調和性)		C (誠実性)	
NEO-FFI	本調査	NEO-FFI	本調査	NEO-FFI	本調査	NEO-FFI	本調査	NEO-FFI	本調査
1	1	2	2	3	3	4	9	5	5
6	6	7	4	8	13	9	14	10	10
11	11	12	7	13	23	14	19	15	15
16	16	17	12	18	43	19	24	20	20
21	21	22	17	23	48	24	29	25	25
26	26	27	22	28	53	29	36	30	30
31	31	32	27	33	58	34	39	35	35
36	33	37	32	38		39	49	40	40
41	41	42	34	43		44	54	45	45
46	42	47	37	48		49	59	50	50
51	46	52	44	53		54		55	55
56	51	57	47	58		59		60	
	56		52						
			57						
			60						

数字は質問項目番号である。

NEO-FFI の欄は使用マニュアルに示された質問項目の分類である。

本調査欄は筆者の因子分析結果である。太字の数字はマニュアルと一致しない項目である。

Table 5 質問項目の因子負荷行列

質問項目	神経症傾向	外向性	開放性	調和性	誠実性	日本版
Q46	-0.77	0.02	0.06	0.13	-0.09	神経症傾向
Q31	-0.71	0.02	0.07	0.07	-0.07	神経症傾向
Q16	-0.70	-0.01	0.06	0.08	-0.05	神経症傾向
Q6	0.60	-0.09	0.01	-0.06	-0.14	神経症傾向
Q1	-0.59	0.08	0.01	0.05	-0.11	神経症傾向
Q21	0.57	-0.10	0.00	0.05	0.10	神経症傾向
Q11	0.54	0.03	-0.04	0.22	0.12	神経症傾向
Q26	0.49	-0.14	0.05	0.05	-0.13	神経症傾向
Q56	0.45	-0.02	-0.12	0.06	-0.03	神経症傾向
Q41	0.43	-0.06	0.08	0.11	-0.16	神経症傾向
Q42	0.42	-0.36	0.21	0.00	0.20	外向性
Q51	0.31	0.19	0.17	-0.02	-0.29	神経症傾向
Q33	-0.20	-0.10	0.08	0.02	-0.04	開放性
Q37	-0.14	0.79	0.07	0.17	0.06	外向性
Q17	-0.05	0.71	0.1	-0.06	-0.04	外向性
Q52	-0.12	0.69	0.04	0.32	0.21	外向性
Q2	-0.06	0.67	0.2	-0.05	-0.15	外向性
Q22	-0.04	0.65	0.06	0.01	-0.03	外向性
Q32	-0.06	0.62	-0.01	0.36	-0.08	外向性
Q7	0.07	0.51	-0.06	-0.07	-0.08	外向性
Q12	0.13	-0.37	0.12	0.19	-0.05	外向性

日本版 NEO-FFI の信頼性の検討

Q27	0	-0.35	-0.19	0.25	0.14	外向性
Q47	-0.03	0.35	-0.03	0.19	0.21	外向性
Q34	-0.26	0.35	0.03	-0.05	0.13	調和性
Q4	0.1	0.31	-0.09	-0.23	0.09	調和性
Q60	0.08	0.29	-0.08	0.03	0.5	誠実性
Q44	-0.18	-0.23	0.22	0.2	0.17	調和性
Q57	-0.08	-0.23	-0.2	0.18	-0.08	外向性
Q43	0.05	0	-0.67	0.01	-0.01	開放性
Q13	-0.03	-0.07	-0.63	0	-0.03	開放性
Q23	-0.12	0.02	0.57	0.21	-0.03	開放性
Q48	-0.01	0.07	0.43	-0.02	-0.04	開放性
Q58	-0.01	-0.08	-0.4	0.16	-0.09	開放性
Q53	-0.04	0.19	-0.38	0.18	0.17	開放性
Q3	-0.09	0.08	0.36	0.01	0.13	開放性
Q59	-0.11	0.02	-0.03	0.49	-0.05	調和性
Q54	-0.1	0.11	0.03	0.48	0.01	調和性
Q36	0.24	0.08	0.07	0.43	0.05	神経症傾向
Q24	0.1	-0.13	0.01	0.43	-0.12	調和性
Q14	0.2	0.05	-0.02	0.34	-0.12	調和性
Q19	0.12	0.17	0.12	-0.32	-0.05	調和性
Q9	0.11	0.2	0.04	0.31	-0.03	調和性
Q49	0.21	0.27	-0.06	-0.28	0.27	調和性
Q29	0.24	0	0.07	0.26	-0.07	調和性
Q39	0.15	-0.12	0	0.21	0.14	調和性
Q50	0.05	0.21	0.03	0.07	0.6	誠実性
Q5	0.03	-0.05	0.09	-0.01	0.6	誠実性
Q20	0.19	-0.02	-0.03	-0.17	0.57	誠実性
Q10	-0.06	-0.04	0.08	0.08	0.51	誠実性
Q15	-0.09	0.06	-0.05	-0.06	-0.5	誠実性
Q35	0.02	0.2	-0.15	-0.03	0.49	誠実性
Q55	-0.01	0.1	-0.01	0.24	-0.49	誠実性
Q25	-0.05	0.23	-0.12	0.15	0.45	誠実性
Q30	0.07	0.1	-0.02	0.04	-0.45	誠実性
Q45	0.04	0.05	-0.07	0.2	-0.37	誠実性
Q40	0.03	-0.08	0.2	0.05	0.23	誠実性

日本版の列は日本版NEO-FFIの項目特性を示す。本研究の結果と一致しない項目を太字で示した。因子負荷量の前についている（-）の符号は、逆転項目を示す。

Table 6 5因子の質問項目中一番因子負荷が大きい項目と因子分析によって分類を変更された項目

項目番号	項目内容
神経症傾向	
Q46	悲しくなったり、落ち込んだりすることはほとんどない。
Q33	場所が変わると、気分や気持ちが変わるものだが、私は気分が変わったとは感じない。

Q42	楽道家ではない。
外向性	
Q37	元気で、はつらつとした人間だ。
Q34	私はほとんどの人から好かれている。
Q4	私は、誰にでも好意を持って接しようとする。
Q60	やることすべてにおいて、志を高く持ってがんばる。
Q44	私は現実的で、情では動かない。
開放性	
Q43	詩を読んだり芸術作品をみていると、ぞくぞくしたり感情の高まりを感じる。
調和性	
Q59	自分の望むものを手に入れるためなら、人を操ることもためらわない。
Q36	人の仕打ちによく腹をたてる。
誠実性	
Q50	バリバリと仕事をやって、それをやり遂げる。

Table 7 因子間相関

	外向性	開放性	調和性	誠実性
神経症傾向	-0.09	0.01	0.14	-0.17
外向性		-0.25	-0.29	0.26
開放性			-0.07	-0.02
調和性				-0.07

2. 項目反応理論による質問項目の選択と尺度の信頼性

分析には、Irm に実装されている Samejima (1969) の段階反応モデルを使用した。以下の分析では、しばしば項目の識別力に言及するが、識別力は appendix に掲載した因子別項目母数表 (Table11) に記載されている。項目反応理論を適用すると、任意のテスト得点に対してテスト情報量を計算することができる。情報量の平方根の逆数 ($1/\sqrt{x}$, ここで x を情報量とする) は、特性得点の標準誤差に相当するため、情報量から計算した標準誤差を特性の得点 (z 得点) の関数としてグラフに描くと、尺度の信頼性を特性得点の関数として検討することができる。

神経症傾向 因子分析の結果から 13 項目を神経症傾向の質問項目と判断した (Table4)。この 13 項目に対して項目母数と情報量を計算し、1.0 以上という基準で識別力が高い項目を選んだ。10 項目が選ばれた (Table8)。標準誤差を日本版 NEO-FFI の 12 項目と、因子分析後の 13 項目および IRT 分析による選択後の 10 項目について図に示した (Fig.1)。標準得点でみると 30 ~ 70 の特性得点の間で、標準誤差が 4 ~ 5 程度で特性得点を推定できることがわかり、高い精度で特性値を推定できると判断できた。

外向性 因子分析の結果から 15 項目を外向性の質問項目と判断した (Table4)。この 15 項目のデータに項目反応理論を適用し、項目母数と情報量を計算した。次に、項目の識別力に注目して、1.0 以上という基準で識別力が高い 8 項目を選んだ。標準誤差を、最初の 12 項目と因子分析後の 15 項目および IRT 分析後の 8 項目について図に示した (Fig.1)。図から 8 項目の質問項目を使うと 30 ~ 70 の特性得点の間で、標準誤差が 4 程度で特性得点を推定できることがわかる。高い精度で特性値を推定できると判断できた。

開放性 因子分析の結果から7項目を誠実性の質問項目と判断した (Table4)。この7項目に対して項目母数と情報量を計算し、1.0以上という基準で識別力が高い項目を選択したところ、3項目が選択された (Table8)。標準誤差を、最初の12項目と因子分析後の7項目とIRT分析後の3項目について図に示した (Fig.1)。標準誤差を見ると、3項目の質問で、30～70の特性得点の間で標準誤差が4～8程度であるから、推定の精度は高くないが、35～65の範囲ならば精度が高いと判断できた。この範囲では、3項目の短縮版のほうが12項目のオリジナル版より、項目数が少ないにもかかわらずやや精度が高いことは注目すべきである。

調和性 因子分析の結果から10項目を誠実性の質問項目と判断した (Table4)。この10項目に対して項目母数と情報量を計算し、1.0以上という基準で識別力が高い項目を選択したところ、3項目のみが選択された (Table8)。標準誤差を、最初の12項目と因子分析後の8項目とIRT分析後の3項目について図に示した (Fig.1)。標準誤差を見ると、30～70の特性得点の間で標準誤差が7～8程度であるから、推定の精度は高くなかった。12項目のオリジナル版と3項目の短縮版の精度に大きな違いはなく、オリジナル版の問題点が明らかとなっている。また、オリジナル版では神経症傾向の項目であるQ36 (人の仕打ちによく腹をたてる) が調和性の3項目に含まれたことも注意すべきであろう。Q36は調和性と神経症傾向の両方に因子負荷が高い項目であった。

誠実性 因子分析の結果から11項目を誠実性の質問項目と判断した (Table4)。この11項目に対して項目母数と情報量を計算し、1.0以上という基準で識別力が高い8項目を選んだ (Table8)。標準誤差を、最初の12項目と因子分析後の11項目およびIRT分析後の8項目について図に示した (Fig.1)。8項目の質問で、30～70の特性得点の間で、標準誤差が5程度で特性得点を推定できた。比較的信頼できる精度で特性値を推定できると判断できた。

Table 8 識別力から選択された項目

神経症傾向	外向性	開放性	調和性	誠実性
Q1	Q2	Q13	Q24	Q5
Q6	Q7	Q23	Q36	Q10
Q11	Q12	Q43	Q59	Q15
Q16	Q17			Q20
Q21	Q22			Q25
Q26	Q32			Q35
Q31	Q37			Q50
Q41	Q52			Q55
Q46				
Q56				

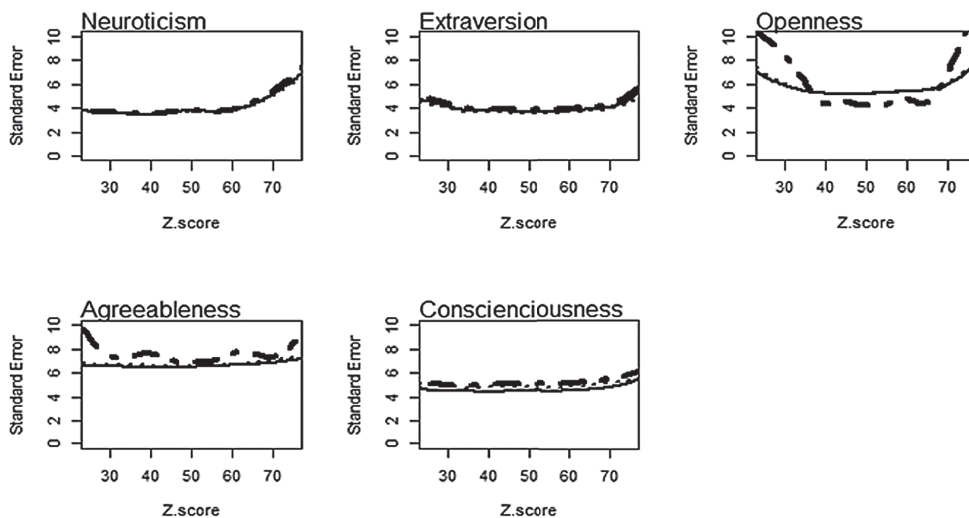


Fig.1 5つの特性別に情報量にもとづいて標準誤差を標準得点に対してプロットした。実線はオリジナル12項目，点線は因子分析後の項目，1点破線は最少項目をあらわす。

ほとんどの特性で実践と点線は重なっている。

開放性では，3項目の短縮版の方が実用的範囲で精度が高い。

3. 特性得点の比較

本研究では，(1) 因子分析の結果から特性の項目を選択する尺度と，(2) 因子分析で選択された項目に項目反応理論を適用して得られた識別力から選択した尺度の2種類が作成された。各々を因子分析版と短縮版とよぶことにする。因子分析版の各尺度の項目番号を Table4の本調査の欄に，また短縮版の各尺度の項目番号を Table8に掲載した。

実際に NEO-FFI のような質問紙を使用して性格特性の量を測定する場合を考えると，通常は項目母数を使って性格特性の得点を計算することはまれであろう。むしろ，NEO-FFI の場合であれば5件法の回答[0, 1, 2, 3, 4]を特性に該当する項目に渡って合計し(素点を作り)，素点を標準化に使われた準拠集団のパーセンタイルと素点の対応表(パーセンタイル換算表)に照らし合わせて，各回答者の特性を標準得点に変換するという手順をとる。そこで，この手順で得られた尺度値を用いて，日本版NEO-FFIのオリジナルの特性12項目版，因子分析版，短縮版の特性尺度の性能を比較した。なお，比較するための基準尺度として因子分析結果にもとづいて選択した質問項目に，因子別に項目反応理論を適用して得られた尺度を使った。なお尺度を比較する際に，得点が1パーセンタイルより低い値(< 26.7)と99パーセンタイルより高い値(> 73.3)を越える値の参加者は除外した。

基準尺度との相関を NEO-FFI オリジナル版，因子分析版，短縮版について表にした(Table9)。NEO-FFI オリジナル版の開放性と調和性は，他の特性に比べるとやや相関が低かった。興味深いことに，短縮版の開放性はオリジナル版より基準尺度と相関が高く，よ

り精度が高い尺度である可能性が見出された。短縮版の開放性は3項目からなる。オリジナル版の開放性特性には不適切な質問項目が多数含まれていて、尺度の精度を低下させているのであろう。当然のことであるが、全体的にみると因子分析版が基準尺度と相関が高く、より精度が高い尺度といえる。

Table 9 基準得点との相関

	神経症傾向	外向性	開放性	調和性	誠実性
オリジナル版	0.95	0.93	0.84	0.87	0.97
因子分析版	0.97	0.96	0.95	0.97	0.97
短縮版	0.97	0.97	0.94	0.82	0.97

IRTにより求められた得点を基準得点とする。これは Table 4 の本調査の項目に対して IRT を適用して得られた得点である。

Table 10 因子間相関

	神経症傾向	外向性	開放性	調和性	誠実性
神経症傾向	1	-0.11	0.06	-0.12	-0.11
外向性	-0.17	1	0.03	0.33	0.29
開放性	0.00	0.11	1	0.05	-0.04
調和性	-0.17	0.22	-0.04	1	0.16
誠実性	-0.09	0.23	-0.04	0.14	1

右下がりの対角要素（1）の上にオリジナル版の因子間相関、下に短縮版の因子間相関を示した。

次に重要な発見としては、短縮版の相関係数は、調和性の0.82を除いて因子分析版の相関係数とほとんど変わらないことである。今後、短縮版の調和性を修正できるならば、素点を計算してそこから標準得点を求めるという手順を使う場合、短縮版でも十分な精度で性格特性の値を推定できるであろう。いずれにせよ、オリジナルの各12項目を用いた性格特性の値は、開放性と調和性で精度が劣ると考えられる。形容詞を用いた Big Five 尺度の短縮版を検討した研究は、並川らの短縮版と日本版NEO-FFIを比較しているが、開放性と調和性で、2つの尺度の相関が低い結果を見出した（並川・谷・脇田・熊谷・中根・野口、2012）。開放性は0.21、調和性は0.46の相関であった。他にも、NEO-PI-Rの開放性と主要5因子検査（村上・村上、1999）の知性との相関が低い（0.29）という結果が報告されている（大野木、2004）。日本版NEO-FFIにおけるこれらの特性尺度の信頼性が低いことがその原因の1つと推測できる。

最後に、特性尺度間の相関を、オリジナル版と因子分析版について調べた（Table10）。オリジナル版では、外向性と調和性の間と、外向性と誠実性の間に相関があった（0.33, 0.29）が、因子分析版では外向性と調和性、外向性と誠実性の相関はそれほど高くなかった（0.22, 0.23）。

4. 残された問題

本研究は、Big Fiveの5つの性格特性を測定する日本版NEO-FFIの尺度について、項目反応理論を適用して得られた情報量の観点から信頼性を検討した。その結果、神経症傾向、外向性、誠実性の3つの特性は信頼性が高いことが見出された。標準得点は30～70の区間で5点の測定誤差で尺度値を得ることは集団テストでは満足すべき精度といえる。これに対して、開放性と調和性の尺度の信頼性は高くなかった。識別力がある項目は、12項目中で開放性が3項目、調和性が3項目であった。開放性と調和性では役に立たない項目が多いことが原因である。開放性と調和性におけるこのような問題点は、日本版NEO-FFIがもつ測定用具としての機能の一部の欠陥であり、日本版NEO-FFIがもつ実用的な価値を大きく損なうとはいえないかもしれないが、近い将来日本版NEO-FFIの開放性と調和性の項目が修正され、より優れた日本版NEO-FFIができあがることを期待する。

5. まとめ

岩手大学生の性格検査に使いやすい尺度を作るために、性格の5因子を測定する日本版NEO-FFIの信頼性を検討した。学部で学生比率と男女の比率を考慮に入れて岩手大学生900人からNEO-FFIに対する回答を得た。質問項目を5因子に分類するために因子分析を行ったところ、NEO-FFIの各因子12項目とは異なる分類結果を得た。さらに、5因子間に弱い相関があった。この新たに分類された項目を使って因子別に項目反応理論(IRT)を適用し、識別力にもとづいて項目を精選したところ、最多10項目(神経症傾向)、最少3項目(開放性と調和性)が選ばれた。次に、もとの各因子12項目を使う尺度、因子分析結果による項目を使う尺度、IRT分析結果による最少項目を使う尺度の3種類の尺度を比較した。情報量からみると、3種類のどの尺度においても開放性と調和性の信頼性は低かった。

謝辞

本研究のデータ収集に当たっては、岩手大学教育学部の多くの学生にお手伝いをしていただきました。特に教育学部学生の小島早織さん、小石裕理香さん、八重樫大周さんには調査1の大部分のデータの収集、計算表への入力、データのチェックの仕事をさせていただきました。ここに記して感謝の意を表します。また調査に参加していただいた岩手大学の多くの学生に感謝します。

引用文献

- Allport, G.W. & Odbert, H.S. (1936) Trait-names: P psych-lexical study. *Psychological Monograph*, 47, No.211.
- Borgatta, E.F. (1964) The structure of personality characteristics. *Behavioral Science*, 9, 8-17.
- Cattell, R.B. (1943) The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology*, 38, 476-506.

日本版 NEO-FFI の信頼性の検討

- Cattell, R.B. (1945a) The description of personality: Principles and findings in a factor analysis. *American Journal of Psychology*, 58,69-90.
- Cattell, R.B., Eber, H.W. & Tatsuoka, M.M. (1970) *Handbook for the Sixteen Personality Factor Questionnaire (16PF)*. Champaign, IL: IPAT.
- Cattell, R.B. (1945b) The principal trait clusters for describing personality. *Psychological Bulletin*, 42, 129-161.
- Costa, P.T. & McCrae, R.R. (1985) *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P.T. & McCrae, R.R. (1992) *NEO-PI-R Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Digman, J.M. & Takemoto-Chock, N.K. (1981) Factors in the natural language of personality: Reanalysis and comparison of six major studies. *Multivariate Behavioral Research*, 16, 149-170,
- Eysenck, H.J. (1990) Biological dimensions of personality. In L.A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 244-276). New York: Guilford Press.
- Eysenck, H.J. (1970) *The structure of personality*. (3rd edition.) London: Methuen
- Fiske, D.W. (1949) Consistency of the factorial structures of personality traits from different sources. *Journal of Abnormal and Social Psychology*, 44, 329-344.
- Goldberg, L.R. (1990) An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology*, 59, 1216-1229,
- Goldberg, L.R. (1992) The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4,26-42.
- John, O.P. & Srivastava, S. (1999) The big five taxonomy: history, measurement and theoretical perspectives. In *Handbook of Personality*, (Pervin & John ed.)
- John, O.P., Donahue, E.M. & Kentle, R.L. (1991) *The Big Five Inventory — Versions 4a and 54*. Berkeley, CA: University of California, Berkeley, Institute of of Personality and Social Research.
- McCrae, R.R. & Costa, P.T. (1997) Personality trait structure as a human universal. *American Psychologist*, 52, 509-516.
- 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.
- 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, 2, 91-99.
- Norman, W.T. (1967) *2800 personality trait descriptors: Normative operating characteristics for a university population*. Department of Psychology, University of Michigan, Ann Arbor, MI.
- 大野木裕明 (2004) 主要5 因子性格検査3 種類間の相関的資料 パーソナリティ研究 12, 82-89.
- Rizopoulos, D. (2006) Irm: An R package for latent variable modeling and item response theory analyses. *Journal of Statistical Software*, 17, 5.
- Samejima, F (1969) Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement*, 34.
- 下仲順子, 中里克治, 権藤恭之, 高山緑 (1998) 日本版NEO-PI-R の作成とその因子的妥当性の検討 性格心理学研究 6, 2, 138-147.
- 下仲順子, 中里克治, 権藤恭之, 高山緑 (1999) *日本版NEO-PI-R, NEO-FFI 使用マニュアル* 東京心理株式会

阿久津 洋 巳

社

豊田秀樹（編）（2002） 項目反応理論 [事例編] – 新しい心理テストの構成法– 朝倉書店

Tupes, E.C. & Christal, R.C. (1961) *Recurrent personality factors based on trait ratings*. (Tech.Rep.) Lackland Air Force Base, TX:USAF.

Appendix

Table 11 項目母数と識別力

神経症傾向					
	Extrmt1	Extrmt2	Extrmt3	Extrmt4	Dscrmn
Q1	-2.92	-1.469	-0.904	1.097	1.466
Q6	-2.777	-1.505	-0.673	0.928	1.561
Q11	-1.666	-0.347	0.283	1.894	1.229
Q16	-2.22	-1.09	-0.619	0.894	2.009
Q21	-2.774	-1.226	-0.417	1.159	1.434
Q26	-1.994	-0.562	0.076	2.029	1.251
Q31	-2.427	-1.233	-0.661	0.959	2.116
Q33	-11.682	-6.208	-3.272	3.032	0.39
Q41	-3.423	-1.273	-0.472	1.938	1.161
Q42	-2.695	-0.559	0.54	2.693	0.951
Q46	-2.57	-1.446	-1.026	0.532	2.533
Q51	-4.656	-2.025	-0.768	2.317	0.718
Q56	-2.826	-1.106	-0.276	1.651	1.041
外向性					
	Extrmt1	Extrmt2	Extrmt3	Extrmt4	Dscrmn
Q2	-2.569	-0.78	0.397	1.868	1.484
Q4	-3.624	-1.31	0.203	2.928	0.834
Q7	-4.296	-2.22	-0.816	1.153	1.13
Q12	-2.916	-0.788	0.555	3.1	1.115
Q17	-2.828	-1.616	-0.493	1.022	2.047
Q22	-2.527	-1.016	0.111	1.533	1.651
Q27	-3.102	-0.383	2.075	5.567	0.623
Q32	-1.801	0.154	1.005	2.578	1.205
Q34	-2.067	-0.024	2.286	4.721	0.982
Q37	-1.379	-0.174	0.755	1.964	2.582
Q44	-7.431	-3.92	-1.204	3.643	0.506
Q47	-4.32	-0.828	0.472	2.904	0.758
Q52	-1.512	-0.117	0.856	1.899	1.917
Q57	-3.405	0.454	3.473	8.486	0.457
Q60	-3.654	-1.181	0.35	2.353	0.917
開放性					
	Extrmt1	Extrmt2	Extrmt3	Extrmt4	Dscrmn
Q3	-4.576	-2.268	-0.792	2.407	0.709
Q13	-1.953	-0.855	-0.172	1.187	1.734
Q23	-2.097	-1.061	-0.047	1.681	1.399
Q43	-1.156	-0.184	0.464	1.615	2.186
Q48	-2.821	-1.416	-0.474	1.667	0.893
Q53	-5.121	-2.155	-0.246	2.245	0.843
Q58	-3.494	-1.411	0.262	2.947	0.818
調和性					
	Extrmt1	Extrmt2	Extrmt3	Extrmt4	Dscrmn
Q9	-6.534	-3.068	-1.173	2.071	0.552

阿久津 洋 巳

Q14	- 3.634	- 0.896	1.376	3.64	0.877
Q19	- 5.891	- 3.314	- 1.095	1.978	0.633
Q24	- 2.529	- 0.547	0.61	2.758	1.238
Q29	- 5.46	- 3.125	- 0.489	2.342	0.759
Q36	- 2.787	- 0.548	0.797	3.065	1.037
Q39	- 5.381	- 2.114	0.346	2.933	0.701
Q49	- 5.783	- 2.918	- 0.755	2.636	0.693
Q54	- 4.334	- 2.223	- 0.842	1.192	0.884
Q59	- 2.985	- 1.381	- 0.334	1.267	1.098
誠実性					
	Extrmt1	Extrmt2	Extrmt3	Extrmt4	Dscrmn
Q5	- 2.073	- 0.379	0.48	2.129	1.496
Q10	- 1.938	- 0.268	0.589	2.422	1.224
Q15	- 2.22	- 0.212	0.754	3.177	1.049
Q20	- 3.884	- 2.235	- 1.47	0.919	1.545
Q25	- 2.155	- 0.253	1.051	3.038	1.041
Q30	- 1.051	1.427	2.406	4.981	0.892
Q35	- 3.744	- 2.516	- 1.2	1.353	1.292
Q40	- 5.535	- 0.052	2.498	7.92	0.395
Q45	- 3.68	- 0.296	0.871	3.533	0.805
Q50	- 2.559	- 0.871	0.47	2.25	1.625
Q55	- 2.444	- 0.485	0.94	3.316	1.122

右端の Dscrmn は識別力である