

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 24 年 5 月 24 日現在

機関番号：23903

研究種目：基盤研究(C)

研究期間：2009～2011

課題番号：21530734

研究課題名（和文）項目反応理論にもとづく妥当でない回答の検出指標が臨床群を識別できるか

研究課題名（英文）Can a person-fit method based on an item response theory differentiate clinical persons from a normal sample?

研究代表者

鋤柄 増根 (SUKIGARA MASUNE)

名古屋市立大学・人間文化研究科・教授

研究者番号：80148155

研究成果の概要（和文）：

項目反応理論の person-fit を利用することで、MMPI の尺度 8 において統合失調症群を健常群と区別できたが、尺度 2 でうつ病群の区別はできなかった。このことは、統合失調症群は健常群と不連続であるが、うつ病群は連続性のあることが示唆された。また、社会的望ましさはテスト得点に加算的に影響するのではなく、社会的望ましさに敏感な個人とそうでない個人とではその影響が異なるという交互作用があることが明らかになった。

研究成果の概要（英文）：

A person-fit method based on the item response theory in the clinical scale 8 on the MMPI was able to differentiate patients with schizophrenia from the normal group, but it on the scale 2 was not to differentiate patients with depression from normal. The results implies that the schizophrenia individuals are discontinuous from the normal group but the depress ones are continuous. Moreover, a way of influences of social desirability on test scores were dependent on the individual's sensitive to social desirability.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2009 年度	1,300,000	390,000	1,690,000
2010 年度	1,000,000	300,000	1,300,000
2011 年度	700,000	210,000	910,000
総計	3,000,000	900,000	3,900,000

研究分野：社会科学

科研費の分科・細目：心理学・臨床心理学

キーワード：性格検査・項目反応理論・person-fit・臨床群・MMPI・社会的望ましさ・多値の項目反応モデル・短縮版

1. 研究開始当初の背景

パーソナリティ検査には、カットオフ得点以上(あるいは以下)の個人は臨床群であると分類する目的に利用されるものがある。その代表的なものが、MMPI(Minnesota Multiphasic Personality Inventory)である。このような分類をするときに、健常群の個人がカットオフ得点を超えてしまうエラー(擬陽性)、逆に、臨床群の個人がカットオフ得点未満にしか

らないエラー(見落とし)がある。健常群と臨床群とではその回答パターンに違いがあると考えることで、これらのエラーを検出できる可能性がある。臨床群の回答パターンは、健常群の回答パターンから見たとき、妥当ではないパターンととらえることになる。このような妥当でない回答の検出について、MMPI は、妥当性尺度 (L,F,K 尺度) や F 尺度の粗点から K 尺度の粗点を減算した F-K 指標、Dr 尺度な

どを使用してきた。今回は、能力検査において、項目反応理論(IRT: Item Response Theory)に基づいて開発された person-fit 指標を、パーソナリティ検査での妥当でない回答パタンの検出に有効かどうかを検討する。

この person-fit 指標は、以下のような原理に基づくものである。ある特性値を持つ個人が、ある項目パラメタ(困難度と識別力)を持つ項目に正答あるいは採点キー側に回答する確率が推測される。さらに、複数の項目に対しても生起する確率の高い回答パターンも推定できるので、このパターンからの逸脱する生起確率の小さい回答パターンは、妥当でないあるいは逸脱した回答パターンであるといえる。このような逸脱した回答パターンを持つ個人の person-fit は低くなる。ここで、臨床群が各項目に健常群とは異なる解釈をすると仮定すると、健常群から推定された項目パラメタに基づいて推定された臨床群の回答パターンは健常群のそれと異なり、臨床群の person-fit は低くなるといえる。つまり、カットオフ得点を超えかつ person-fit の低い個人は臨床群であるが、カットオフ得点を超えていても person-fit の高い個人は健常群であると判断できることになる。

回答パターンに違いがあるという考え方の背後には、臨床群の個人は健常者とは検査項目に異なる解釈をしている、さらにいえば、臨床群は健常者とは質的に異なる集団であるという考えがあることになる。しかし、近年、アナログ研究では、健常群と臨床群の間の連続性を仮定している。今回の研究は、このような考え方の検討にもなる。

しかし、項目反応理論が能力検査を対象として発展してきたものであることから、そこで考えられてきた指標がそのままパーソナリティ検査に適用できるかを確認がまず必要であった。妥当でない回答の発生機序として、能力検査では、パーソナリティ検査ではよく問題になる回答の偽装はほとんど関係していない。この偽装的な受検態度を person-fit が検出可能であるかを、まず検討しなければならない。

さらに、パーソナリティ検査の場合、項目の社会的望ましさに反応して回答が歪められる場合がある。仮に、回答パターンに違いがあったときに、その違いが個人ごとが考える社会的望ましきの違いが反映されており、項目が本来測定しようとしている特性についての反応の違いではないかもしれない。ここで、社会的望ましきと個人の特性との交互作用が存在する可能性も検討しておかなければならないことになる。

2. 研究の目的

研究 1

項目反応理論に基づいた person-fit 指標は、ある特性値の個人が回答する確率の高い妥当な回答パターンからの逸脱として定義され、その逸脱が大きくなれば負、小さければ正の値をとる。臨床群は健常群の平均からの逸脱、つまり連続した逸脱と考えるなら、臨床群の得点は健常群の平均より逸脱して高いが、その回答パターンは健常群と同様に妥当なものであり逸脱していない。一方、この連続性が仮定できないなら、臨床群の回答パターンは妥当なものから逸脱している可能性があり、person-fit 指標による識別可能性がある。さらに、回答をよりよく見せる(faking-good)あるいはより悪く見せる(faking-bad)という偽装的な態度で受検した時の person-fit との比較もする。

研究 2

自己報告式のパーソナリティ検査では、項目特性としての社会的望ましきと、社会的に望ましい方向への反応傾向という個人特性との2つがその回答に影響している。しかし、従来の研究ではこれらを必ずしも明確に分けていないし、これらの交互作用も検討していない。そこで本研究ではこれらの点を階層的重回帰分析によって検討する。

さらに、項目の社会的望ましきの尺度化が必要である。この点について、段階反応(graded response)に関して、Raschを基礎とする Generalized Partial Credit Model(GPCM)と Thurstoneを基礎とする系列範疇法のそれぞれの推定値を比較することで、GPCMにおける項目(i)の位置パラメタ(b_i)が、項目の社会的望ましきの推定値として利用可能かを検討する。また、カテゴリーをデータ収集後にまとめたときの推定値の変化も検討する。

3. 研究の方法

研究 1

健常研究協力者 短・四大学の 1, 2 年生 377 人(男 145, 女 232)。

臨床データ 198 名(男 98, 女 100)。診断名の主なものは、統合失調症(46 名)、アルコール依存(43 名)、うつ病(18 名)、不安障害(15 名)、神経症(13 名)、摂食障害(7 名)、気分障害(7 名)、人格障害(6 名)であり、他の診断名に該当する人数は 5 名以下であり 1 名しかいない診断名も多くあった。

faking 群 faking-good 群は 58 人(男 23, 女 35), faking-bad 群が 62 人(男 34, 女 28)で、いずれも大学生。

MMPI MMPI 新日本版(1993)

手続 健常者群と faking 群は冊子式を集団実施した。臨床群は、複数の臨床心理士がカード式を実施し、彼らを通して主治医から承諾を得た上で、データの提供を受けた。

研究 2

調査協力者 503 名(女: 313, 男: 189, 不明: 1)の大学生

調査内容 パーソナリティ検査として和田(1996)のBigFive(E: Extraversion, N: Neuroticism, O: Openness to Experience, C: Conscientiousness, A: Agreeableness; 7段階評定), その各項目の社会的望ましさ(項目特性; 9段階評定), 社会的望ましさへの反応傾向(個人特性)を測定するためにMMPIの項目から選択された39項目からなるEdwardsの社会的望ましさ尺度(2件法)。

手続 まず, BigFive と Edwards の社会的望ましさ尺度を冊子にしたものを配布・実施。BigFive の 60 項目の社会的望ましさの測定が影響を及ぼさないように, これらが終了後, 項目の社会的望ましさを測定する調査票を配付・実施した。

4. 研究成果

研究 1

特性値の推定に使用したパラメタは, 鋤柄(1998)が MMPI 新日本版の標準化データ(男 500, 女 522)について, 2-パラメロジスチックモデルによって推定したものを使用した。

各尺度ごとの person-fit 値, 推定特性値 \square , z 得点を求めた。z 得点は MMPI の手引きの基準値を使って換算したものである。person-fit 値の平均を臨床群と健常群で比較した結果, いずれの尺度においても, 臨床群の平均 person-fit 値は健常群より有意(1%水準, $df = 573$, t 検定による)に低く, より逸脱回答をしていることを示す。最も差が大きかった尺度は尺度 3(Hy)であり, 逆に最も差が小さかった尺度は尺度 8(Sc)であった。

次に, person-fit, 推定特性値 \square , z 得点の間の相互相関を求めた。 \square と z 得点の間の相関のほとんどが.8 前後であり, \square は z 得点のよい推定値になっていることを示している。person-fit と \square の相関が 5%水準で有意になったものは, 臨床群で-.401(Hy)から-.141(Si)の間であり, 健常群では-.397(Hy)から.195(Hs: 唯一正の相関)であった。また, person-fit と z 得点との有意な相関は臨床群では Mf(-.146)のみであり, 健常群では Mf(-.214), Hs(.134), Sc(-.111)であった。推定特性値あるいは z 得

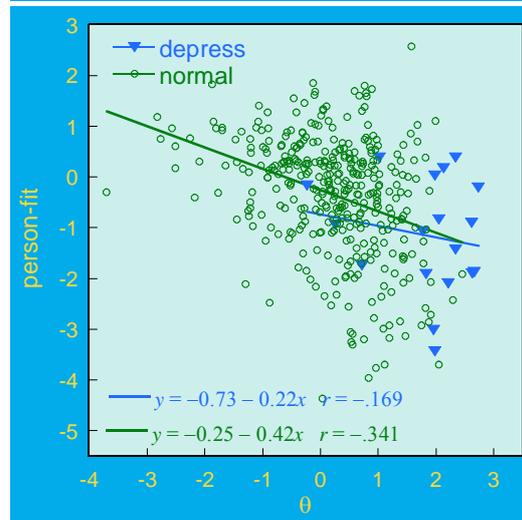
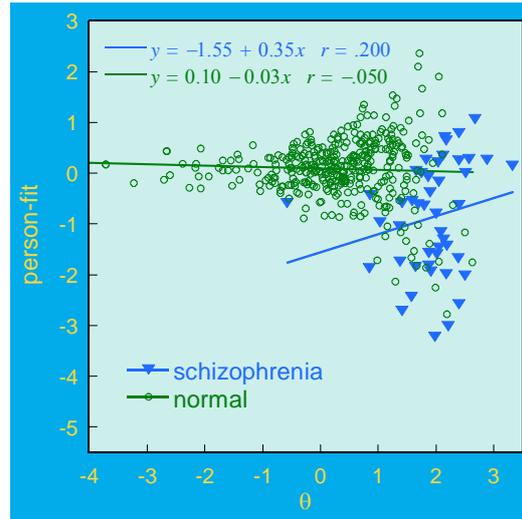


図 1 健常群と臨床群の person-fit と推定特性値 θ の散布図と回帰直線(上) 統合失調症群と健常群 (下) 抑うつ群と健常群

点いずれにおいても, person-fit との相関は負であり, 特性値が高くなると逸脱回答である可能性が高くなることを示す。

臨床群としてまとめた人数のある統合失調症と抑うつをとりあげ, それぞれに関連の強い尺度[統合失調症では尺度 8(Sc), 抑うつでは尺度 2(D)]における推定特性値 \square と person-fit との関連を図 1 に示した。統合失調症群では, 両変数の間に正の傾きを持つ回帰式が見られた(図 1 上)。しかし, 抑うつ群(図 1 下)では, 健常者・臨床群ともに両変数間には負の傾きの関係があった。

次に 尺度ごとの person-fit 値を尺度についてまとめた平均 person-fit を各個人ごとに求めた。その平均 person-fit の分布を, 偽装的な受検態度で受検した標本と比較したのが図 2 である。臨床群の分布

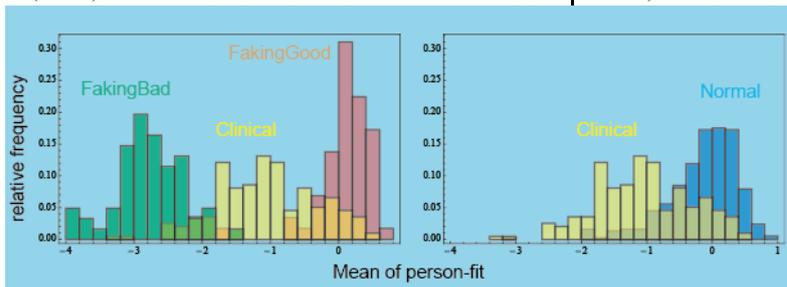


図 2 臨床群, 健常群, Faking-Good, Faking-Bad 群の person-fit 平均値のヒストグラム

は、普通の受検態度の健常者群と faking-bad 群の中間に位置している。person-fit は faking-bad と臨床群の区別を可能にする指標の一つとなることを示唆する。なお、faking-good 群は健常群とほぼ重なりあっており、faking-good を検出する指標に person-fit は使えないことを示す。

また、MMPIは項目数が多いので、項目反応理論のパラメタである識別力を利用して、短縮版の作成も検討した。短縮版と完全版との等価性の検証方法は、一度実施した完全版から完全版の得点と短縮版の項目による得点を求め、それらの間の相関を利用するものが多い。しかし、項目が重複している得点による相関であるので、過大推定になっていると考えられる。この点を改善するために、完全版の得点でマッチングをとった等価と考えられる2群を作成し、片方を完全版実施群、残りの半数を短縮版実施群とし、両群の得点間の相関を取るという方法で検討した。その結果良好な短縮版が得られることが確認された。

研究 2

項目特性としての社会的望ましさとそれに対する反応傾向との交互作用を検討するために、階層的重回帰分析を性格因子ごとに、因子の尺度得点を被説明変数として実施した。説明変数である項目特性としての社会的望ましさと Edwards の主効果のみのモデル 1 を検討し、次にこの両者の交互作用を投入したモデル 2 を検討した。なお説明変数はそれぞれの平均を減算している。また、各得点に性差のあるものがあつたが、性の主効果を入れた重回帰分析ではそれらの偏重回帰係数が有意になることはなかったため、上の分析では男女をまとめて分析した。

モデル 1 から 2 への R^2 の増分が有意すなわち交互作用が有意になった性格因子は N と A であつた(順に、 $\Delta R^2 = .024, .015, F_s(1, 486) = 18.650, 8.983, p < .005$)。図 3 に性格因子ごとに、Edwards の尺度得点の $\pm 1SD$ と平均での、性格因子の社会的望ましきから性格因子の尺度得点への回帰直線を示す。N については $+1SD$ と平均での傾きは、有意にゼロより大きい(順に、 $t_s = 3.84, 2.10, df = 487, p < .05$)が、 $-1SD$ における傾きは有意ではなかった。A については $+1SD$ での傾きは、有意にゼロより大きい($t = 2.93, df = 487, p < .05$)が、平均と $-1SD$ における傾きは有意ではなかった。交互作用が有意でなかった性格因子では、社会的

望ましきへの反応傾向とは関係なく、尺度の望ましきを高く判断した個人の性格因子の尺度値は高くなっている。

図 3 を見てわかるように、5 因子の E, O, C 尺度においては、その尺度を社会的に望ましいと判断した個人ほど尺度得点も高くなっている。また、社会的望ましきへの反応傾向の高低とは関係なくこの傾向は同じであつた。しかし、N と A に関しては、社会的望ましきへの反応傾向の高い個人についてのみ E, O, C と同様の傾向が見られたが、低い個人では、その尺度に関する社会的に望ましきの判断と尺度得点は関係ないことが見て取れる。このことは、社会的望ましきへの反応傾向の低い個人は、尺度や項目の社会的望ましきに無関心あるいは反応できていないと考えられる。これらの個人が正直に回答していると解釈することも可能であるが、検査場面や状況によっては要求される自分をよく見せようとする防衛的な反応ができないことを表しているとも解釈できるものである。

GPCM による社会的望ましきの推定は、IRTPRO 2.1 を使用し、系列範疇法による推定は Guilford (1954) の計算方法によって求めた。さらに、9 段階のカテゴリーの 1 と 2, 3 と 4, 6 と 7, 8 と 9 をそれぞれカテゴリー 1, 2, 4, 5 とまとめ、5 はそのまま 3 として 5 段階カテゴリーにまとめたものについても同様に分析した。

GPCM における項目の位置パラメタは、多

表 1 GPCM の項目位置パラメタ(b)および系列範疇法による尺度値(median)それぞれ上位 5 項目と下位 5 項目(項目の後のアルファベットは 5 因子の性格次元、ゴシックは反転項目)

GPCM (top5)		GPCM (bottom5)			
b	item	b	item		
-2.28	温和な	A	-0.84	地味な	E
-2.20	話し好き	E	-0.80	飽きっぽい	C
-2.07	人嫌い	E	-0.68	成り行きまかせ	C
-2.03	自己中心的	A	-0.68	緊張しやすい	N
-1.90	無節操	C	-0.50	心配性	N
successive (top5)		successive (bottom5)			
median	item	median	item		
3.19	親切な	A	-0.78	悩みがち	N
3.16	短気	A	-0.75	不安になりやすい	N
3.14	社交的	E	-0.67	弱気になる	N
3.13	良心的な	A	-0.59	憂鬱な	N
3.12	怒りっぽい	A	-0.54	悲観的な	N

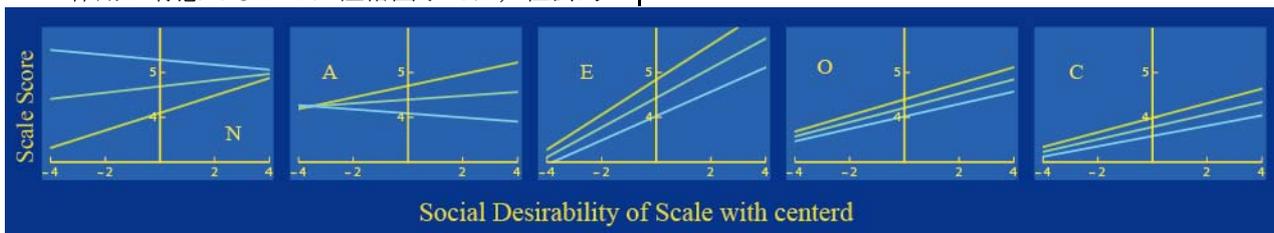


図3 Edwardsの $\pm 1SD$ と平均での社会的望ましきから尺度得点への回帰直線
N, A: 交互作用のあつた尺度 E, O, C: 交互作用のなかつた尺度 黄: $+1SD$, 緑: Mean, 青: $-1SD$

くの個人が社会的に望ましいと答える項目ほど小さな値になる。一方、系列範疇法で得られる項目の尺度値は、社会的望ましい項目ほど大きな値となる。表1に、9段階カテゴリーの分析について、それぞれの推定方法による、パラメタと項目内容を上位5項目と下位5項目について示す。表を見て分かるように上位には望ましいと考えられる項目が、下位には望ましくない項目がきているので、推定値はいずれの方法も適切なものであるといえる(ゴシックの反転項目は内容を逆に解釈)。両者の相関は-.671であり、先に述べたことから予想されるように負の関連を示している。ただし、N(神経症的傾向)次元の項目だけを取り出した場合の相関は.666と正の関連を示している。N次元は社会的に望ましくない項目で構成されており、N次元そのものが社会的に望ましくないものといえる。推定値はいずれも社会的望ましくないことを表してはいるが、正の関連を示す理由は明確でない。

9カテゴリーを5カテゴリーにまとめた場合の推定値と9カテゴリーのままの推定値の関連は、GPCMによる項目位置の推定値間の相関は.865であった。また、系列範疇法による推定の場合、推定値間には.997の相関があった。いずれの推定法も、それぞれの方法の中では、ほぼ同じ推定値を示しているといえる。ところが、5カテゴリーにおける2つの推定方法による推定値間の相関は-.467となり、カテゴリーをまとめると推定方法間で、この推定値が一致しなくなるといえる。特に、E次元の「話好き」のGPCMの項目位置パラメタがカテゴリーをまとめたとき極端に社会的に望ましい方向に変化している(-2.20から-4.16)。しかし、この項目のカテゴリー頻度は他の項目と大きく異なっていない。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計4件)

1. 中川敦子・木村由佳・鋤柄増根・水野里恵 幼児の行動のチェックリスト (ECBQ) 短縮版の作成, 名古屋市立大学院人間文化研究科紀要人間文化研究, 16, 2011, 1-16.
2. 中川敦子・木村由佳・鋤柄増根 乳児の行動のチェックリスト (IBQ-R) 短縮版の作成, 名古屋市立大学大学院人間文化研究科人間文化研究 2009, 12, 15-25
3. Akechi, T., Ietsugu, T., Sukigara, M., Okamura, H., Nakano, T., Akizuki, N., Okamura, M., Shimizu, K., Okuyama, T., Furukawa, T. A., & Uchitomi, Y. Symptom indicator of severity of depression in cancer patients: a comparison of the DSM-IV criteria with

alternative diagnostic criteria. *General Hospital Psychiatry*, 2009, 31, 225-232.

[学会発表] (計6件)

1. 鋤柄増根 シミュレーションデータによる心理検査の短縮版と完全版の等価性の検証方法, 日本心理学会第75回大会(於日本大学), 2011年9月
2. 鋤柄増根 MMPIにおける臨床群の識別へのperson-fitの応用可能性, 日本パーソナリティ心理学会第19回大会(於慶応大学), 2010年10月
3. 鋤柄増根・酒井恵子 パーソナリティ検査における項目の社会的望ましさと社会的望ましさへの反応傾向との交互作用効果, 日本心理学会第73回大会(於立命館大学), 2009年10月

[図書] (計0件)

[産業財産権]

○出願状況 (計0件)

なし

○取得状況 (計0件)

なし

[その他]

なし

6. 研究組織

(1)研究代表者

鋤柄 増根 (SUKIGARA MASUNE)

名古屋市立大学・人間文化研究科・教授

研究者番号: 80148155

(2)研究分担者

なし

(3)連携研究者

なし