

令和 6 年 5 月 29 日現在

機関番号：14501

研究種目：研究活動スタート支援

研究期間：2022～2023

課題番号：22K20316

研究課題名（和文）回答時間を利用した適応型の多肢強制選択心理測定の開発

研究課題名（英文）Development of an adaptive multiple forced-choice psychological measurement using response time

研究代表者

分寺 杏介（Bunji, Kyosuke）

神戸大学・経営学研究科・准教授

研究者番号：40962957

交付決定額（研究期間全体）：（直接経費） 2,000,000円

研究成果の概要（和文）：本研究は、多肢強制選択型の心理尺度を回答時間データを利用し適応型で実施する方法論開発に向けて、実践的な課題点についての検討を行った。具体的には、文脈効果・順序効果が Thurstonian IRTモデルで仮定される「選択肢の効用の不変性」に及ぼす影響および回答方法の違いがモデルパラメータに及ぼす影響を、既存データおよび新規に収集したデータを用いて分析した。ベイズ統計モデリングによる評価の結果、文脈効果と順序効果はいずれも完全にはゼロでないが、実用上の影響は限定的であり、適応型測定の実施には僅かな注意が必要とされた。また異なる回答方法が推定結果に及ぼす影響も小さい可能性が示された。

研究成果の学術的意義や社会的意義

心理尺度によるパーソナリティの測定は、社会科学の学術研究に限らず、社会の様々な場面で用いられており、より短時間で・より精度良く個人の性格特性を評価する方法のニーズは高い。本研究の成果は、心理測定の効率性・有効性を高めるための方法論開発へと繋がる重要な知見を提供するものである。

研究成果の概要（英文）：This study examines practical issues in the development of an adaptive version of a multiple-choice psychological measurement using response time data. Specifically, the study analyzes the impact of context effects and order effects on the assumed "invariance of the utility of choice options" in the Thurstonian IRT model, as well as the influence of different response methods on model parameters, using both existing and newly collected data. The evaluation, conducted through Bayesian statistical modeling, reveals that while context and order effects are not completely zero, their practical impact is limited, requiring only slight attention when implementing adaptive measurements. Additionally, the results suggest that the influence of different response methods on the estimation outcomes is likely minimal.

研究分野：心理統計学

キーワード：強制選択 文脈効果 順序効果 ベイズ統計モデリング

1. 研究開始当初の背景

心理学や社会学を始めとする多くの研究分野では、心理尺度によるパーソナリティの測定が頻繁に行われる。心理尺度による調査に関しては、以下に挙げたような方略による改善の試みが行われてきた(図1)。

A) 多肢強制選択型への変更 心理尺度の大多数はリッカート型で実施されている。リッカート型測定では、様々な文(例:「わたしは社会的である」)が連続で呈示され、それぞれの文にどの程度当てはまるかを複数の選択肢の中から選ぶ。特に重要度の高い採用試験などの場面において、リッカート型測定は意図的に回答を歪めて自分をよく見せようとする行為(フェイキング)などの測定バイアスに弱いことが知られている。対応策の一つとして、複数の文を同時に呈示し最も当てはまるものを選択させる多肢強制選択型への変更がある。

B) 回答時間データの利用 従来、心理測定は紙面上で実施されていたが、近年ではコンピュータを用いた測定(**Computer Based Testing; CBT**)が増加している。この流れは特にここ数年のコロナ禍で急激に進展しており、今後も様々な測定が紙から**CBT**に移行すると考えられる。**CBT**では回答情報に加えて、回答時間データを容易に入手できる。回答時間データの利用方法は数多く検討されており、実際に測定の信頼性・妥当性向上に寄与していることが知られている。

C) 適応型の項目呈示 また**CBT**では比較的容易に、個人ごとに異なる項目を呈示可能である。適応型測定などと呼ばれるこの方法は、全回答者に同じ項目セットを呈示する方法(項目固定型)と比べて、より短い時間で個人の特性値を精度良く推定できる。

以上の3点を統合することができれば、各改善法が持っている利点を兼ね備えた、従来のものよりも妥当かつ効率的な測定法となることが期待される。

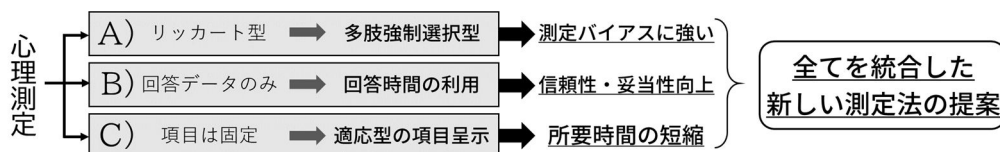


図1 心理測定の改善方略と研究のゴール

2. 研究の目的

本研究では、多肢強制選択型の心理尺度を、回答時間データを利用しながら適応型で実施するための方法論開発に向けて、実践的な課題点について検討することを目的とした。具体的には、多肢強制選択型心理測定データに対して広く用いられている**Thurstonian IRT**モデルにおいて仮定されている「選択肢の効用の不変性」を検証することで、適応型測定の開発に向けての示唆を提供することを目指した。

多肢強制選択型心理測定では、複数の選択肢が同時に提示される。**Thurstonian IRT**モデルでは、回答者は各選択肢に対して潜在的な「効用」をもち、効用が最も高い選択肢を選ぶと考える。このとき、各選択肢の効用は回答者自身の心理特性の強弱に基づいて独立に決まると仮定されており、他の選択肢の効用や選択肢の提示される順番等の本来無関係な要素によって変化することはないことが期待されている。多肢強制選択型の適応型測定の開発に際しては、まず同時に提示する選択肢の組み合わせをどのように決めるかが問題となる。仮に効用の不変性が満たされている場合には、選択肢単位で最適なものを選ぶことができるため、より「適応的」に実施することができる。一方満たされていない場合には、あらかじめ選択肢を組み合わせたブロックを作成する必要がある。また実施の際にも効用が変化しないように様々な注意を払う必要がある。加えて、**Thurstonian IRT**モデルの仮定が満たされないことになるため、モデル自体の改善等を行う必要が生じると考えられる。

3. 研究の方法

本研究期間には、大きく分けて以下の3つの研究を行った。

(1) 選択肢の文脈効果に関する二次分析

文脈効果とは、ある選択肢の効用が、同時に提示される別の選択肢の影響を受けることを指す。仮に多肢強制選択型心理測定において文脈効果が認められる場合、例えば文Aの効用が、「文A,Bが同時に提示されたとき」と「文A,Cが同時に提示されたとき」で変わるようになるため、事前に文を組み合わせたブロックを作成する必要があるといえる。この文脈効果の有無を検証するため、**Bunji & Okada (2021, Multivariate Behavioral Research)**で収集されたデータを利用して、ベイズ統計モデリングによる文脈効果の有無の検証を行った。このデータでは、回答者は4つの文(**Big Five**を測定するための文)について、

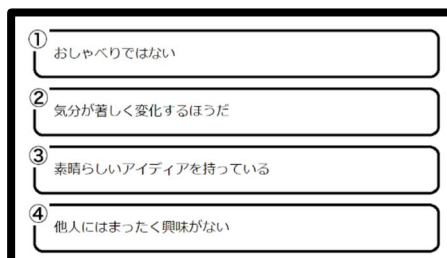


図2 提示された画面のイメージ(1)

「自分に当てはまる」と思う順に回答することを求められた(図 2)。また本データでは、合計 24 種類の文を、それぞれ 3 回ずつ用いて、18 ブロックの項目を作成している。したがって、ある文(例:「おしゃべりではない」)は、3 つの異なるブロックにおいて、それぞれ 3 つの異なる文と同時に提示されている。

(2) 選択肢の順序効果に関する検証

順序効果とは、あるブロック内の選択肢について、どの選択肢が先に提示されるかによって、その効用が変化することを指す。仮に順序効果が確認される場合には、適応型測定実施時に、選択肢の表示順を変えるべきではないことになり、ひいては事前に文を組み合わせたブロックを作成する必要性にも繋がる。順序効果の有無を検証するため、クラウドソーシングにて 500 名の日本人を対象にデータを収集した。このデータでは、画面に 2 つの文(Big Five を測定するための文)が左右に提示され(図 3)、回答者は「自分に当てはまる」と思う方を回答することを求められた。その際、文の組み合わせは全回答者で共通であったが、どちらの文が左に表示されるかを回答者ごとにランダムに決定した。

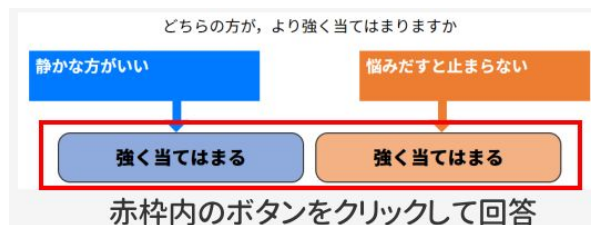


図 3 提示された画面のイメージ

(3) 回答方法の違いがおよぼす影響についての検証

CBT を実用化する場合、回答者は様々なデバイスから回答を行う可能性があり、またデバイスに応じて異なる方法での回答を許容することが望ましい。特に回答時間を考慮した適応型測定実施のためには、回答方法によって異なりうる回答時間が選択肢の効用等に影響しないことが期待される。回答方法の違いが回答時間や文の特性におよぼす影響を見るため、同じ文・ブロックを用いて収集された 2 種類のデータを用いた二次分析を行った。1 つ目のデータは Bunji & Okada (2020, Behavior Research Methods) で使用されたものであり、このデータでは回答者は提示された 2 つの文のうち「自分に当てはまる」と思う方を、キーボードのボタンを押して回答することが求められた。2 つ目のデータは本研究(2)のために収集したデータであり、このデータでは同じ文の中からマウスクリックで回答することが求められた。

4. 研究成果

(1) 選択肢の文脈効果に関する二次分析

異なるブロックで提示されている同一の文に関して、項目特性を表すパラメータが不変であると仮定したモデル(以下、不変モデル)および、同じ文であっても異なるブロックで提示されている場合は異なる文であるとみなして異なる項目パラメータを設定するモデル(以下、変動モデル)の 2 種類を考え、モデル比較を行った。また、異なるブロックにおける同一の文の項目パラメータや効用を直接比較するため「ある文以外のすべての文が文脈効果に対して不変である」ことを仮定したモデル(以下、部分変動モデル)を追加し、検証を行った。

得られた結果は次のとおりである。心理測定の最も重要な目的である「個人の心理特性の推定」に関しては、不変・変動モデル間でほぼ同じ値が得られた($r > .999$)。情報量規準(WAIC)を比較したところ、わずかに変動モデルが良い値を示したが、標準誤差と比べても小さな差異であった。全モデルのベイズファクターを比較したところ、不変モデルが明らかに良いことが示された。部分変動モデルにおいて、同じ文から異なるブロック間で示される効用の 3 つの事後分布の重なり率を見たところ、平均で 83.6%の重なりが見られた(図 4)。

以上の結果をまとめると、本研究で使用したデータにおいては、文脈効果は完全にゼロではないが、実用上は結果に大きな影響を与えるほどのものではないと言えるだろう。文脈効果を考慮して変動モデルを採用するよりも、僅かな差異を気にせず不変モデルを採用したほうが、推定するパラメータ数の減少のメリットが大きいと考えられる。

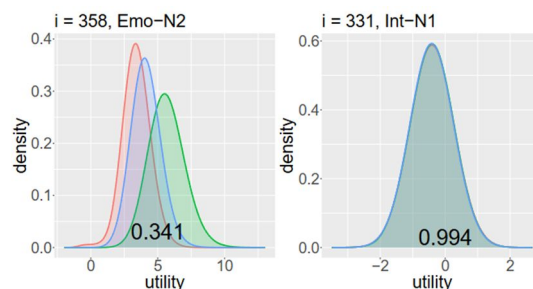


図 4 効用の事後分布の重なり率。

左: 最も重なり率が低かったケース (34.1%)
右: 最も重なり率が高かったケース (99.4%)

(2) 選択肢の順序効果に関する検証

はじめに記述統計量として、選択肢の提示位置による選択率の違いを確認した(図 5)。その結果、基本的に選択率の差は 10%未満にとどまった。以下では、同じ文であれば、左右どちらに

提示されても項目パラメータが不変であると仮定したモデル(以下,不変モデル)および,同じ文であっても提示位置が異なっている場合は異なる文であるとみなして異なる項目パラメータを設定するモデル(以下,変動モデル)の2種類を考え,モデル比較を行った。

得られた結果は次のとおりである。「個人の心理特性の推定」に関しては,不変・変動モデル間でほぼ同じ値が得られた($r > .993$)。情報量規準(WAIC)を比較したところ,わずかに変動モデルが良い値を示したが,標準誤差と比べても小さな差異であった。変動モデルにおいて,「ある項目パラメータの差が0である」という(ヌル)モデルとのベイズファクターを **Savage-Dickey** 法で近似的に項目レベルで算出したところ,すべての項目についてヌルモデルが優勢となった。変動モデルにおける同一項目の項目パラメータの差異に関して **95%確信区間 (HDI)** を求めたところ, **9** 割以上の項目・ブロックにおいて **HDI** は **0** を含む区間であった。項目パラメータの事後分布の重なり率を算出したところ,一部の項目について重なり率が **50%未満**であった。

以上の結果をまとめると,個人特性の推定に関して言えば順序効果は無視できるほど小さいと考えることができる。一方で,事後分布の重なり率が低いことなどを踏まえると,やはり順序効果は完全にゼロではないと言える。適応型測定を実施するためには項目パラメータを事前に推定する必要があるため,その際には注意が必要である。

(3) 回答方法の違いがおよぼす影響についての検証

2種類のデータについて,始めに記述統計の比較を行った。項目選択率のオッズのデータ間相関(図6)は $r = .916$, また平均回答時間のデータ間相関(図7)は $r = 0.92$ といずれも高い値であった。また,両データを同一のモデル(Thurstonian Diffusion IRT; Bunji & Okada, 2020, Behavior Research Methods)で分析し,得られた項目・ブロックパラメータを比較したところ,次のような結果が得られた。モデル間での推定値の相関は **0.875-0.926** であった。因子負荷に相当するパラメータについて,データ2(マウスクリック)のほうが僅かに高い傾向が見られた。項目の所要時間に相当するパラメータに関して,データ2のほうが僅かに速い傾向が見られた。項目情報量については,概ね標本誤差のレベルに収まった。

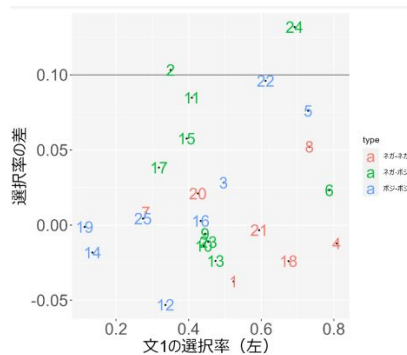


図5 提示位置による選択率の差

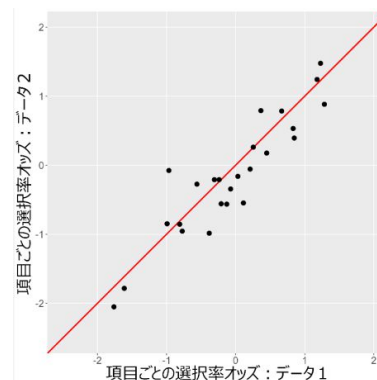


図6 項目選択率のオッズ

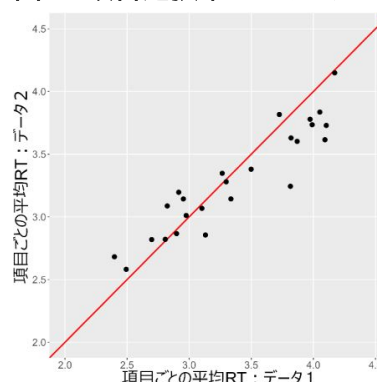


図7 平均回答時間

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 1件）

1. 著者名 分寺 杏介	4. 巻 19
2. 論文標題 コンピュータを用いたアセスメントに関する研究トピックの整理と最新の動向	5. 発行年 2023年
3. 雑誌名 日本テスト学会誌	6. 最初と最後の頁 191 ~ 225
掲載論文のDOI（デジタルオブジェクト識別子） 10.24690/jart.19.1_191	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計2件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 分寺 杏介, 岡田 謙介
2. 発表標題 異なる回答方法における一対比較型心理測定の安定性の検討
3. 学会等名 日本計算機統計学会第37回大会
4. 発表年 2023年

1. 発表者名 分寺 杏介
2. 発表標題 一対比較型心理測定における順序効果に関する検討
3. 学会等名 日本行動計量学会第51回大会
4. 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------