

尺度開発

橋本佐由理*

Scale development

Hashimoto Sayuri, Ph.D

キーワード

尺度 scale

測定 measurement

妥当性 validity

信頼性 reliability

I 健康行動科学に基づいた研究の必要性

行動科学が知られるようになり、人間行動の研究が各分野で進んだが、本学会で最も関心が寄せられているのは、健康をめぐる人間行動に関する研究であろう。それは健康行動科学と呼ばれ、だれもが高いQOL(生活の質)を保ちながら生きていけるよう、心身の健康を維持・増進・回復するための行動変容を成功させ、セルフケア行動がとれるようになることに重大な関心がある。人は健康になるために生きているのではなく、個人のライフスタイルや価値観があるので、自分の大切な人生の目標を犠牲にしてまで行動を変えられないし、本人が必要を感じ、変えたいと思わないかぎり行動変容は不可能である。最終

*日本保健医療行動科学会奨励研究員

的に本人が自己決定したことでなければ、せっかく変容した行動も続かない。

現在の疾患は、生活習慣病といわれるような本人の長年の生活習慣、生活環境、すなわち生き方そのものが疾患をつくりだしており、その多くがストレス関連疾患である。また、事故やけがではなく、病氣らしいといって病院にやってくる患者は、明確な原因や疾患が見つからないことも多い。こうした疾患や病気の予防や治療の成功には、生活環境の改善、生活習慣や生き方、人生目標の変容までをも求められる。

たとえば、患者の保健行動変容の支援のために様々な指導法や教育法、相談法を行ってきたが、これまでの方法だけでは患者の行動変容や病気の予防や治療の成功していない現状がある。したがって、行動科学に基づいた技術や技法が不可欠であり、行動科学的な方法を取り入れた研究が必要である。保健医療分野だけではなく、教育分野やその他の分野でも人間行動について取り上げるかぎり、行動科学的な視点が必要であることはもちろんである。ここでは行動科学的な視点のなかでも焦点を絞り、人の態度や意見などを測定するための研究のツールである尺度の開発方法や手順を述べていく。

II 調査研究のための尺度開発方法

1. 尺度開発の前に

当然のことではあるが、まず調査研究のテーマを決定し、その目的を明らかにする。研究方法は、大きくは質問紙調査と決定しているが、調査時期や調査対象者、調査票の内容などの方法を検討して計画書を作成する。もちろん取り上げるテーマや対象などに関する文献を検討する。そして、調査対象者の背景を知る必要がある。これらを検討するとき、机上のデータだけではなく、現場での臨床経験からつかんでいるものが役立つことはいうまでもない。先行研究に照らし合わせ、自分の経験や専門分野、現場での実践や臨床から集積した対象者からの生きた声を十分に活用し、オリジナリティを発揮することが大切

である。

2. 尺度とは

尺度とは大まかにいえば、測定における規則のことである。それは表1に示した、名義尺度・順位尺度・間隔尺度・比率尺度の4つの測定水準である。

さて社会測定には、尺度を用いる第1種測定法と主観的な第2種測定法とに分けられる。第1種測定法には、内的一貫性尺度、外的な基準のある予測尺度、入学試験のように外的基準のない選抜尺度・指数・指標などがあげられる。第2種測定法には、投票法・評定法・序列法・対比較法などがある。このように社会測定では、間接的に数値に直すことで測定をするのである。

保健医療や教育分野で作成する尺度も、人々の態度や意見などの主観的なも

表1 尺度の4つの種類

- 1) 名義尺度：間隔の概念も大小や後先の概念もない尺度であり、数値で表されるものではない。性質や属性、具体的な分類名(カテゴリー)を値とする質的なデータである。性別であれば男か女か、右側・中央・左側、内向性・外向性などの対象者における対立する反応や特徴、参加・不参加、所有・未所有などの対象者の特定の反応や特徴のあり・なし、賛成・反対、快・不快などの複数の選択肢の質問に対する回答結果。
- 2) 順位尺度：A・B・Cや優・良・可、社員・管理職・経営者といった値の大小や後先(順序)、ランクを保証するものである。たとえば、運動会の徒競走における1位、2位、3位というのは順位尺度であり、鼻の差であっても大差であっても1位の次は2位、その次は3位と決まる。したがって、個々の値の間に等間隔性は保証されない。間隔を問題としていないので、小数はない。これも記号で表す質的なデータである。
- 3) 間隔尺度：個々の値の間には、等間隔が保証されているものである。温度というのはこれにあたる。温度の0度というのは、温度がなくなるわけではないけれども、0度と1度、1度と2度との間隔は必ず等間隔である。必ず数値で表される量的データである。
- 4) 比率尺度：等間隔であるだけでなく、ゼロを基点とすることができる尺度のことをいう。たとえば、時間のゼロというのは、時間がないことをいうが、1秒と2秒、2秒と3秒の間隔は等しい。距離も同様である。これも必ず数値で表される量的データである。

のを測定しようとしているので、その作成法には、評定法や序列法などが用いられる。たとえば、評定法により作成された尺度を評定尺度と呼び、評定尺度得点が得られる。それは、形容詞対（広い—狭い）や程度副詞（非常に、かなり、まあ、あまり）を横に並べて物差しの役割をするものを作成し、そのなかから選択してもらったり、丸をつけてもらう方法で回答を得る。それらを集計する際に得点化していくが、ここに1つ問題が生じる。広いから狭いまでの間を5段階設けたのであれば、1～5あるいは0～4の値を得られる。得られた値そのものは、程度や強さなどの大小を示しているので、順位尺度と判定するのであれば問題ではないが、間隔尺度といえるかどうかである。間隔尺度としたほうが、統計的に処理をするにも処理方法の数も質もまさっているが、非常に誤った結果を招くことも起こりうる。尺度の判定ミスは、研究の結論の誤りとなる。評定尺度法で得られたものを、順位尺度とするか、間隔尺度と判定するかは、実は研究者の考え次第である。結果が妥当な仮説を支持するものであったり、有益な知見を得ているのであれば間隔尺度と判定することは適切であろう。研究者の考え次第なので、同じ専門分野の熟練した研究者に参考意見やアドバイスを得ることは有効である。ここでは詳細は省くが、このような態度測定のための尺度作成には、サーストンの等現間隔法・リッカート法などの尺度構成法があげられる。

3. 調査票作成の手順

尺度を用いた調査票の作成手順を簡単に述べる。

- 1) 測定目的・測定領域の決定。
- 2) 項目収集（関連文献・既成調査票）・自由記述。

3) 尺度項目作成：項目の表現（意見項目の作成基準）は、まず、①簡潔な文であること、②1つの項目のなかに2つの意見を含まない、そして、③単なる事実の表現ではなく、肯定的否定的な態度の表れでなくてはならない、④表現はなるべく現在形を用いる、⑤特定の人からだけ肯定されたり、特定の人だけにあてはまる記述を避ける、⑥意見は測定される態度に関して意味のあるも

のでなければならない、以上のような作成基準を参考にする。また、たとえば、予備調査の尺度項目の選択肢には、大変そう思う・そう思う・普通・あまりそう思わない・まったくそう思わないの他に、どちらともいえないといった選択肢を加えておくことと尺度項目の選定や修正の際に参考になる。

4) 予備調査。

5) 項目の修正(仮尺度の作成)・項目分析：予備調査のデータから、以下の観点から検討し、項目の修正や削除をする。①各項目の性質を各項目における個人差の大きさ(分布に極端な偏りがあるものを削除、標準偏差が小さいものを削除)について統計的に検討する。このときに削除対象となる項目や先に述べた「どちらともいえない」の選択が多かった項目は、項目の表現に問題がないか検討する。たとえば、筆者の尺度で運動効果に対する信念尺度の作成過程で「運動は健康や美容に効果がある」という項目表現には「どちらともいえない」が多かったが、「運動は健康や美容にも効果がある」という表現に変更すると、選択肢から選択された。②各項目の識別力(弁別力・因子分析・上位一下位分析)の検討。③項目全体の等質性(内的整合性 α 係数)の検討。

6) 尺度の決定。

7) 尺度の反応歪曲の補正・尺度の信頼性の推定。

8) 調査票の編集(実施時間の決定・教示・項目の配列)：調査票により、なるべく正確な態度や意見を反映したデータを収集したい。そのためには、不適切なデータを排除するために、①編集の際に、逆転項目を入れておく(すべて1を選ぶなど項目を読まずに選択しているケースなどを削除できる)、②調査票の尺度の並べ方、尺度項目の並べ方は特に意図がないかぎり、時間的なあるいは思考的な流れに沿うようにする、あるいは、初めから難しい尺度や答えにくい尺度、非常にプライバシーにかかわる尺度などをおかないほうがよい。また、③ただ、調査に協力を願うというだけでなく、調査票に回答することで対象者にとっても、気づきや学びにつながるように工夫するとよい。

9) 標準化のための調査。

10) 妥当性の検証。

11) 調査票の実施：調査をしたら、その結果によっては、再度標準化の修正や検証をする必要が出てくることもある。

12) 調査票の手引作成。

13) 実用化。

III 尺度の信頼性と妥当性について

調査研究において、質問紙調査の結果から考察や結論を述べるためには、調査票を構成している尺度が、信頼性・妥当性があるかどうかが問題になる。具体的な統計手法は省くが、信頼性、妥当性の考え方を本稿の最後に述べておく。

1. 信頼性とは

信頼性とは、たとえば、測定器具でいうのであれば、測定器の精度にあたるもので、その尺度が、目的とする対象をどの程度正確に、一貫して測定できるかを示す指標である。たとえば、信頼性の高いカメラは、いつでもどこでもだ

表2 信頼性係数の算出法

- 1) 再検査法：同一尺度を一定期間において同一集団に2回測定し、2度の尺度の結果の間の相関係数によって、再検査信頼性係数を推定する。記憶と学習が再検査法による信頼性に影響を及ぼすことがある。性格検査や健康調査票などに有効な方法であるといわれている。
- 2) 平行テスト法：測定内容は同一であるが、設問の表現様式が異なる2種のテストを実施して、その相関係数で信頼性係数を推定する方法である。たとえば、新しい尺度を作成した場合には、ほぼ同じ内容を測定しているとみられる既成の尺度との相関を求める。
- 3) 折半法：2種のテストをするかわりに、1回のテストの結果を2つの部分に折半する方法で、尺度項目をランダムに2等分して、分割されたそれぞれのグループに含まれる項目の合計得点の相関を求めるものである。
- 4) 内の一貫性による方法：尺度の各項目間の反応の一貫性の程度を信頼性係数の推定値とする方法であり、各項目間の得点の等質性が検証できる。クロンバッチの α 係数などがよく算出される。その係数は、項目数が多いと上がる傾向がある。

れが撮っても同じように撮れ、状況が変化しても同じ機能を発揮するもので、この特性は繰り返し得られるはずである。すなわち、同一条件と考えられる枠のなかで何回測定あるいは調査をしても対象の属性・指標のデータに変動が少なく、一貫している場合を信頼性が高いといい、時間や期間をおいても結果に差がないのである。また、調査のたびにデータがばらばらになる場合は信頼性が低いという。何度測定や調査をしても結果はほぼ同じであることを計算で示した値を信頼性係数という。それは、0 から 1 の間の数値をとるが、信頼性係数の算出法には表 2 のようなものがある。

どのくらいの信頼性が得られればよしとするかは、その調査の分野や性格によるのでいちがいにはいえないが、一般的に社会調査における尺度の信頼性係数は、0.70以上であれば信頼性があると判断される。予備調査の段階で算出したときなどデータ数が少ないと、多少0.70を下回ってしまうことがあるが、データを増やせば0.70以上になることが多い。また、最も用いられる方法は、内的一貫性を示すクロンバッハの α 係数の算出によるものであろう。信頼性のない尺度で測定した結果から、いくら考察をしたとしても意味のないものとなるので注意が必要である。

信頼性が低い場合の対処法は、それが既成の尺度の場合には原著をあたって、自分の研究対象に対して適切な尺度であるか、測定方法や時期が適切であるかを再確認する。独自開発した尺度であれば、尺度を構成する項目のなかに不適切な項目が含まれている可能性があるため、それぞれの項目を削除したときの α 係数の算出をして検討するとよい。

2. 妥当性とは

妥当性とは、尺度が測定しようとしているものを本当に測定しているかどうかを調べるための概念である。たとえば、子どもにテキストから文章を写しながら、意味のないリズムを覚えさせるテストをすると、信頼性は高くてもこのことが文章を理解しているかの読解力を知る指標としては適切ではなく、妥当性を欠いているということになる。すなわち、妥当性とは測定対象そのものに

表3 妥当性の検討法

- 1) 内容妥当性：これは、調査項目のサンプリングが適切であるかどうかを検討するものであり、その適否については、主として専門家の個人的かつ主観的判断を通して評価される。スポーツのことであればスポーツをやっている人に、うつなことなら精神科医に訊くなどで表現妥当性の検討ができる。また、論理的に一致しているかどうかを検討することで論理妥当性の検討ができる。
- 2) 基準関連妥当性：妥当性のなかでも最も中心的なものである。尺度が測定したいものを測定していることを確かめるために、検討したい尺度と基準尺度(外的基準)との相関関係によって検討したい尺度の効力を示す。ここには、予測妥当性、併存妥当性も含まれる。一般には、相関係数を妥当性係数とする。サンプル数によって、どの程度の妥当性をもってよしとするかが異なるが、目安としては、100サンプルであれば、0.4~0.5、1000サンプルであれば、0.2~0.3程度の妥当性係数が得られていけばよいと思われる。
- 3) 構成概念妥当性：その尺度によって測定される諸概念(学力、不安、抑うつなど)が理論的に導かれる仮説とどの程度一致するかを示す指標である。因子分析によって尺度の構成概念(因子)を抽出し、尺度がどの程度まで構成概念によって説明されるかを検討するものである。

対する概念であり、それは測定対象と対比させることではじめて意味をもつ。尺度が多目的に構成されている場合、その目的に応じていくつもの妥当性が考えられる。妥当性を測定する主な方法を表3に示した。

参考文献

- 1) ボーンシュテット&ノーキ、海野道郎・中村隆監訳(1990)：社会統計学、ハーベスト社。
- 2) 海保博之編著(1995)：心理・教育データの解析法10講[2]応用編、福村出版。
- 3) 田中敏・山際勇一郎(1992)：教育・心理統計と実験計画法、教育出版。
- 4) 橋本佐由理・岩崎義正・宗像恒次・江澤郁子(1996)：運動行動をめぐる心理社会的要因に関する尺度の検討、日本保健医療行動科学会年報、11：215-232。
- 5) 橋本佐由理・岩崎義正・宗像恒次・江澤郁子(1997)：運動・食生活・健康管理の自信感に関する調査研究、日本保健医療行動科学会年報、12：205-222。