

質問紙調査入門

— 医療安全研究に役立つ Tips 集 —

Introduction to Questionnaire Survey — Tips for Research on Patient Safety —

芳賀 繁 HAGA Shigeru

社会安全研究所・立教大学 名誉教授
Research Institute for Social Safety/ Professor Emeritus, Rikkyo University

要約

はじめに、質問紙調査（本稿では紙に印刷された調査票を使うものだけでなく、電子媒体によるものを含む）について素直な回答を得るための工夫、倫理的配慮、調査票の構成、質問項目の順序や言葉遣いに関する注意事項をまとめた。

次に尺度とは何か、どのような種類があるのかを述べ、心理尺度開発のプロセス、および尺度の信頼性、妥当性について初学者向けに解説した。

キーワード：質問紙調査、心理尺度、信頼性、妥当性

I. はじめに

人を対象とする調査には行動を直接または間接的に観察する「観察調査」と、質問を投げかけて回答を求める「質問調査」がある。質問調査は日本では一般にアンケートと呼ばれているが、英語では questionnaire survey という。アンケートはフランス語の “enquête” から来ていて、フランス語の意味は「調査、探索」、英語の “survey” に当たる。

質問調査における質問の投げ方は、電話、戸別訪問や街頭でのインタビューもあるが、研究目的では紙の調査票を配って調査対象者に記入して貰う形が多かった。これを「質問紙調査」と呼ぶ。しかし、近年ではインターネット（web）調査やイントラネット調査、メールでの調査など、電子媒体を用いる調査が増えている。英語

の questionnaire survey には口頭での質問もインターネットによる調査も含まれるが、日本の心理学界では questionnaire survey を伝統的に「質問紙調査」と呼び習わしてきた。「紙」の一字が邪魔になるが、「質問調査」という言葉はあまり定着していないので、本稿では「質問紙調査」で通すことにする。以後、質問紙調査は紙に印刷された調査票を使うものだけでなく、電子媒体を含む文字を使った質問に回答を求める形の調査を指すものと考えて読み進めて欲しい。

なお、実験や行動観察ではなく、質問紙調査を用いる研究方法を「質問紙法」という。人の意見、感情、信条、過去の経験、将来の予測、未来への展望などは、実験や行動観察で探ることが難しいので、質問紙法は心理学の重要な研究手法の一つとなっている。

本稿では、質問紙調査で素直な回答を得る工夫、倫理的配慮、調査票の構成、質問項目の順序や言葉遣いに関

する注意事項をまとめた後、心理尺度について初学者向けに解説する。

II. 質問紙調査の基本

1. 素直な回答を得るために

自分自身の安全に関する行動、勤務する組織や職場、上司に対する意見などを素直に答えてもらうのは難しい。自分が書いた答えを上司が見たり、勤め先に調べられたりしたら、怒られないか、人間関係が悪くならないか、仕事がやりにくくならないか、あとで不利益が生じないかと不安になるのは当然である。

回答が無記名であっても、職場で調査票を回収したり、組織のシステムに入力したり、電子メール等で回答を送信する形だと、回答者が特定されてしまう心配がある。したがって、そのような不安を取り除くための工夫が必要である。

筆者が企業に頼まれて、あるいは企業に頼んで安全意識や安全文化調査をする場合、まず、その企業と文書による覚書を交わして、個人の回答は決して会社に伝えないことを取り決める。その上で、質問紙と回答用紙を配るときに回答用紙を入れて提出するための封筒も一緒に配り、回答を入れて封をした封筒を提出して貰う。企業側は集めた封筒を開封せずに筆者の研究室に送る。封筒代と開封する作業が余分にかかるが仕方ない。調査対象者には回答の秘密が守られることを覚書によって保証されていることを伝えて、安心して正直な意見や気持ちを答えるよう依頼する。

外部機関と提携して、社員意識調査を匿名化して実施できるシステムを持っている会社もある。これだとコーディングの手間がかからず、データになった回答を貰ってすぐに分析できるのでありがたい。

病院内の組織（安全管理室など）が自病院の職員を対象に調査をする場合は、回答内容を安全管理室のスタッフ等に見られてしまうことに不安を感じる可能性がある。安全管理室は病院経営組織に回答を決して開示しないことを約束して、そのことを調査対象者に周知するくらいしかないが、それを信じて貰うには日頃の信頼関係が重要であることは言うまでもない。

我々のような心理学者やその所属機関の大学・研究機関、あるいは調査会社などの第三者に入って貰うのも一案である。調査結果を研究データとして使って貰う代わりに、データの回収先になって貰うのである。この場合には筆者が企業と取り交わしているような覚書を交換す

るとよい。

回答は統計的に処理されること、回収された調査票は鍵のかかるロッカー等に保管すること、データはノートパソコンやUSBメモリに入れて持ち歩かず、安全な場所に保存することなどを約束する

また、後述するフェイスシートで性別、年齢、職種、所属、職位などを尋ねると、匿名であっても回答者が1人または数人に絞られてしまう場合がある。たとえば看護師対象の調査で性別を聞くと、ある病棟の男性看護師は一人しかいないとか、年齢段階を「1. 20代、2. 30代、3. 40代、4. 50歳以上」から選ばせると職場では50歳以上が一人しかいないとかのケースである。このような場合、そもそも性別での比較が必要なのかどうか考えて性別を聞くのをやめる、40代と50歳以上を合併して「40歳以上」とするなどの工夫が必要である。

2. 倫理的配慮

今は多くの病院に倫理審査委員会があるから、その審査を受ければよいが、いくつかの基本的事項を書いておこう。

まず、回答を強制してはいけない。調査を実施する側は対象者全員に答えてもらいたいのだが、現在、大学等の研究機関で常識になっているのは、「回答したくない人は回答しなくてもよい」という原則である。そして、「たとえ回答しなくてもあなたの不利益にはならない」と伝えなければならない。

そして、回答者を困惑させるような質問をしてはいけない。年収、配偶者の有無、性的な事柄に関する質問など、もし尋ねたいなら調査目的に必須であることを倫理委員会に示す必要がある。

性別は、男女だけでなく、「その他」や「答えない」という選択肢をもうけよう。

3. 調査票の構成

最初に依頼文、次に回答方法の説明、フェイスシート、質問項目、お礼の言葉が続く。

依頼文には、調査の目的、趣旨、依頼者（病院長、医療安全推進室長など）、問い合わせ先を記載する。

フェイスシートは性別、年齢、所属、職種、職位など、回答者の属性に関する質問群である。これらの値を「デモグラフィック変数」という。回答のコーディングや後の統計処理に便利なよう、選択肢には番号を付けておく。たとえば「性別：1. 男 2. 女 3. その他 4. 答えない」のように、である。フェイスシートは質問文の前に

おかれることが多いが、筆者は「最後にあなた自身のことについてお伺いします」という前置きをして、質問の最後にデモグラフィック変数を尋ねる。その方が、回答にバイアスがかからないと思うからである。一方、自分の立場、地位などを踏まえて回答して貰いたい場合は最初に置くがよい。学術集会でのパネルディスカッションでは、藤本先生はフェイスシートを先に置くべきだと主張されていた。どちらに置くのがよいかは調査票を作る際に考えて欲しい。

お礼の言葉は一言「ご協力ありがとうございました」でよい。

4. 質問項目の並びとワーディング

ワーディングとは言葉の選び方のことである。ルールを変えることへの質問で「改正」と書いたり「改悪」と書いたりすると回答に影響を与える。「改定」や「変更」といった価値中立の言葉を使うべきである。「お年寄り」「高齢者」「若者」「子ども」などは回答者によってイメージする年代が違う可能性がある。ずっと前のエピソードだが、研究室の若者が大阪のことを「地方都市」と呼んだのを聞いた関西育ちの私が「おいおい大阪は地方都市じゃないだろう」と言ったら、東京出身のその若者に「えっ？東京以外は全部地方じゃないですか」と反論されショックを受けた。いろいろな言葉は人に寄って意味合いが異なるので、質問の文案は調査を開始する前に、できるだけ世代や出身地や性別の異なる様々な人に見てもらい必要がある。同じ病院内でも診療科や病棟や職種によっても器具の呼び名や言葉の使い方が異なるので、なるべく多くの人に頼んで試しに回答してもらおうとよい。

判断に影響を与えるような文章を質問文に含めてはいけない。「○○ですが、○○についてあなたはどのように思いますか？」というような質問の「○○ですが」の部分である。たとえば、「小まめな手指消毒は院内感染予防に大きな効果があることが知られていますが、あなたは1

つの作業毎に手指消毒をしていますか？」と尋ねると「していない」とは答えにくい。

ときどき質問文に二つの質問内容が含まれているケースを見かける。たとえば「マスクやフェイスシールドは新型コロナウイルス感染予防に役立つと思いますか？」という質問は、「マスクは役立つか」と「フェイスシールドは役立つか」の2問に分けなければならない。マスクは役立つけどフェイスシールドは役立たないと思っている人もいだろうから。

質問の順番にも注意が必要だ。前の質問が後の質問への回答に影響を与えることがある。これを「キャリオーバー効果」という。患者からの暴言やハラスメントについての質問を重ねた後に、「患者の問題行動に対してはこれまで以上の強い対応が必要だと思いますか？」と尋ねたら肯定的な回答が増えるだろう。また、インシデントの経験の有無やエラーの頻度を尋ねた後に、自分自身の性格やストレスについて問えばネガティブな側に回答が傾くだろう。

自身の安全行動や組織の安全施策についていくつかの質問を並べるときに、質問の向きを時々変える方がよい。つまり、全部「非常にそうだ」と答えれば肯定的（安全側）の回答になると、しっかり質問文を読まずに回答する傾向を生む。ときどき「声出し確認は意味がない」というような項目を混ぜておくのがコツである。このような質問項目を「逆転項目」と呼ぶ。

Ⅲ. 心理尺度の基本

1. 尺度とは

尺度 (scale) とはモノサシである。ちょっと難しく定義すると「対象に数値のある規則に従って割り当てたもの」である。

尺度は表1にまとめたように4種類ある。

名義尺度とはカテゴリに数字を割り当てただけのもの

表1 尺度の種類と特徴

種類	特徴	可能な統計処理
名義尺度	カテゴリーの区別 例：男女、血液型、職種、所属	度数分布、回答の割合
順序尺度	大小関係に意味がある。値間の差は一定ではない 例：順位、時代、先輩後輩の関係	中央値、順位相関係数
間隔尺度	値の差が一定。差に意味がある。原点と単位は任意 例：温度（摂氏／華氏）、西暦、心理尺度得点	平均値、標準偏差、積率相関係数
比例尺度	比率に意味がある。「無」の状態が原点（ゼロ） 例：長さ、質量、反応時間、年齢、入院期間	平均値、標準偏差、積率相関係数、比率

である。たとえば「看護師=1. 医師=2」と決めてもいいし、「医師=1, 看護師=2」と決めてもいい。ただのラベルだからどちらでもいいのだ。したがって、このタイプの尺度で平均値を計算するのはナンセンスである。ところが筆者が査読依頼を受けた投稿論文にそのような計算をしているものがあったりして仰天したことがある。

順序尺度とは順位が何番目かという情報が数値で与えられる。1番と2番は僅差で、2番と3番の間に大きな差がある場合もあるので、平均値を計算することはできない。

一方、間隔尺度は1と2の間、2と3の間、…100と101の間…が等間隔のモノサシである。しかし原点（値がゼロの場所）は任意に決められたものなので比率に意味がない。気温を測る単位の摂氏も間隔尺度である。英米で使われている華氏とは原点が違うので同じ気温でも摂氏と華氏とでは値が異なることは読者もご存じだろう。2022年の夏はヨーロッパを猛暑が襲ったが、それを伝える日本の某テレビが「英国ヒースロー空港の最高気温が40.2℃を記録した。平年の最高気温の“1.5倍”」と報じた。平年の最高気温が27度なので40を27で割って約1.5と計算したのだろうが、摂氏40度は華氏104度、摂氏27度は華氏80.6度なので、華氏で考えると約1.3倍となる。もちろんどちらの計算も間違いである。30度は20度の1.5倍ではないし、10度は20度の半分ではない。ただし、30度と20度の差は10、10度と20度の差も10であり、温度差は等しいと言える。10度と20度の平均が15度というのも正しい。だから平均気温を計算できるのである。

比例尺度は原点ゼロが「無」の状態、そこからの距離を同じ単位（間隔）で測るモノサシである。巻尺や定規で測る長さ、体重計で測る重さ、年齢、患者数、ベッド数などがこれに当たる。重さの単位がキログラムでもポンドでも、60キログラム（132ポンド）の体重は40キログラム（88ポンド）の1.5倍である。これは重さゼロの物体はグラムで測ってもポンドで測ってもゼロで、そこから別の単位であっても等間隔に測られるからである。比例尺度は加減乗除などの統計計算ができる。

2. リッカート尺度

さて、質問紙調査では5段階や7段階、まれには4段階の評価尺度が頻繁に使われる。これをリッカート尺度という。リッカート尺度で得られた1～5、あるいは1～7の値は順序尺度だろうか間隔尺度だろうか。「あま

例1

	全くそうではない	ややそうではない	どちらとも言えない	ややそうだ	非常にそうだ
問1. あなたは○○ですか？	1	2	3	4	5
問2. あなたは△△ですか？	1	2	3	4	5

例2

	全くそうではない	問1. あなたは○○ですか？			非常にそうだ
	1	2	3	4	5
	全くそうではない	問2. あなたは△△ですか？			非常にそうだ
	1	2	3	4	5

図1 評価を間隔尺度として扱うための工夫

り～ない」は「全く～ない」より肯定的で、「どちらとも言えない」は「あまり～ない」よりさらに肯定的だと言えるが、「全く～ない」と「あまり～ない」の差と、「あまり～ない」と「どちらとも言えない」の差は同じだろうか。

質問紙法を使った調査や研究では、やや強引に間隔尺度であるとみなして、平均や標準偏差の計算を行っている。5つ（あるいは7つ、あるいは4つ）の選択肢の心理的感覚が等しいと仮定するのだが、そのためには、図1のような等間隔の目盛りで回答してもらうことが望ましい。最近ではWeb調査や電子メールを使った調査が多いため、選択肢をリストにしてクリックさせたり、プルダウンメニューから選ばせたりするが、その場合でも、せめて選択肢に番号を振って、「1. あまり～ない」「2. やや～ない」などと表記して欲しいと筆者は考えている。

程度や頻度の副詞表現に関する論文も複数あるが、行動や経験の頻度を問う質問は難しい。「めったに～ない」「たまに」「ときどき」「しばしば」という表現は順序尺度ではあるが、相互の間隔が等しいと仮定するのは無理があると筆者は思う。うっかりミスなどの頻度について「1. 年に数回程度 2. 月に数回程度 3. ほぼ毎週…」というような聞き方をしている調査を見かけるが、この回答の平均値を計算するのは奇妙である。むしろ、「年に何回くらい○○しますか？」と尋ねて数値で答えてもらうのが望ましい。

3. 心理尺度の作り方

店や従業員のサービスに対する意見、購入した商品の満足度、支持する政党、政府の政策についての賛否などに関する調査は、一つひとつの質問に調査対象者がどの

ように答えるかを探るものが多い。一方、心理尺度は複数の質問を重ねて心理的構成概念を定量的に測るものである。

心理的構成概念とは、直接観察したり、物理的なモノサシで測ったりすることはできないが、人間の心理や行動を説明するのに役立つ概念である。私たちが日常的に使っている言葉である「やさしさ」「慎重さ」「ずるさ」「うぬぼれの強さ」「頑固さ」なども心理的構成概念と言えなくもないが、通常は心理学できちんとした操作的定義を与えられて、測定法あるいは評価方法が提案されているものを差す。たとえば、「自尊心」「リーダーシップ」「外向性/内向性」「モチベーション」「職務満足度」「安全態度」などである。心理的構成概念には個人の行動特性を説明するだけでなく、組織や集団の特性を表すものもある。「安全文化」「心理的安全性」「集団凝集性」などである。

心理的構成概念を質問紙法で定量化するために、心理尺度の開発が行われる。藤本先生は先行研究で使われている尺度を利用することを勧めているが、リサーチクエストを解明するためには既存尺度では間に合わないこともあるだろう。そんな時のために、自ら新しい心理尺度を開発する手順を大まかに示したのが図2である。単に思いつきで質問を並べただけでは心理尺度にならないことを知って欲しい。そして、実際に心理尺度を作りたいなら、心理尺度の教科書、解説書を勉強し、尺度開発を報告している研究論文を熟読して欲しい。

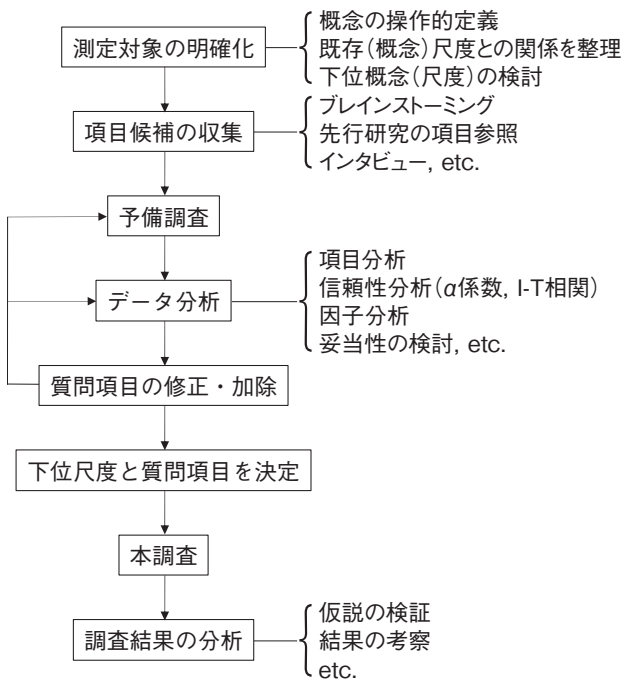


図2 心理尺度開発の大まかな手順

4. 心理尺度の信頼性と妥当性

どのような尺度にも言えることだが、心理尺度でもっとも大事なことは信頼性と妥当性である。

信頼性とは文字通りその尺度で測った値が信頼できるかどうかである。測る度に別の結果が出るようなモノサシは信頼できない。測る人が変わると結果も変わるようでも困る。心理尺度では複数の質問項目で一つの心理的構成概念を測定するが、その項目群に一貫性があるかどうかも重要視される。安全態度を測る尺度の中に「カレーライスが好きか」という質問項目があっても意味がない。安全態度尺度を構成する質問項目はどれも安全態度に関連している必要がある。「カレーライスが好きか」という質問項目は安全態度尺度の信頼性を損なうのである。このタイプの信頼性は内的整合性とも呼ばれ、クロンバックのα係数で数値化される。一般にα係数が.70ないし.80を超えていれば信頼性があると考えられている。α係数だけでなく、ω(オメガ)係数、修正済みI-T相関(その項目と、その項目を除いた他の項目の合計との相関)や、因子分析の因子負荷量等も参照して欲しい。

妥当性とはその尺度が測ろうとしているものを本当に測っているかということである。身長を知りたいのに体温計を持ってきても妥当性のある測定値は得られない。心理尺度の場合は、尺度を構成する質問項目全体で捉えたい内容をどの程度正確に捉えているかの程度を指す。

一つ一つの質問項目が測定したい心理的構成概念の内容を捉えているか(内容的妥当性)、全体として過不足がないか(構成概念妥当性)などを専門家に聞いて、その判断を参考にするとよい。何人かの専門家に項目を見てもらい、測定したい概念を調べられているかを4段階できいて、1人でも1や2をつけたら項目を排除するという手法もある。

妥当性を定量的に評価するには、外部の基準に照らして判断する方法がある。これを「基準関連妥当性」という。質問紙で測られた安全態度が、インシデントの頻度や、上司がその人の日頃の安全行動を評価した値と相関が高ければ、その安全態度尺度は妥当性があると判断できる。

あるいは、この尺度で測る心理的構成概念は、既存の別の尺度で測られる心理的構成概念と関連があるはずであると理論的に推定できるなら、両者の結果にある程度の相関がみられるはずである。これを「併存的妥当性」という。しかし、既存尺度との相関があまりにも相関が高ければ新しい尺度を作る意味がない。心理学研究でよ

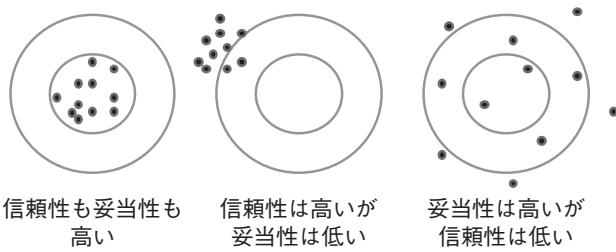


図3 尺度の信頼性と妥当性のメタファ

く行われるのは、新しく開発した尺度は既存尺度のこの部分（下位尺度）と相関が高く、別の部分（下位尺度）あるいは、別の既存尺度との間の相関は低いはずだと仮説を立て、それを検証することで新尺度の妥当性を論じることである。このほか、妥当性の評価については様々な考えや手法があるので参考書や文献に当たって欲しい。

信頼性と妥当性について説明するためによく使われるイラストが図3である。射撃の的の中央に弾痕が集まっていれば信頼性も妥当性も高く、中心から外れたところに弾痕が集まっている場合は的外れ、すなわち妥当性が低い尺度であり、弾痕が集まらずにバラバラになっているのが信頼性の低い尺度である。

IV. おわりに

本稿では、質問紙調査で素直な回答を得る工夫、倫理的配慮、調査票の構成、質問項目の順序や言葉遣いに関する注意事項をまとめた、次に心理尺度について初学者

向けに解説した。質問紙調査を立案して実施する際、質問紙調査を使った研究をする際、質問紙調査を使った論文や報告書を読む際の参考にしていただければ幸いである。

文献表

- 1) 芳賀繁・中村玲香・山出康世, 看護職のためのエラータイプチェックリストの開発, 医療の質・安全学会誌, 第1巻第1号, 15-22, 2006.
- 2) 久田満, 看護に活かす心理尺度—その選び方・使い方, ナカニシヤ出版, 2015.
- 3) 堀洋道 (監修), 心理測定尺度集 I~IV, サイエンス社, 2001-2011.
- 4) 鎌原雅彦ほか (編著), 心理学マニュアル—質問紙法, 北大路書房, 1998.
- 5) 川本利恵子ほか, 「尺度」を使った漢語研究のキホンとコツ—看護研究の精度向上・時間短縮のために, 日本看護協会出版会, 2016.
- 6) 宮本聡介, 質問紙調査と心理測定尺度—計画から実施・解析まで, サイエンス社, 2014.
- 7) 村上宣寛, 心理尺度の作り方, 北大路書房, 2006.
- 8) 種田憲一郎ほか, 安全文化を測る—患者安全文化尺度日本語版の作成, 医療の質・安全学会誌, 第4巻1号, 10-24, 2009.
- 9) 谷岡一郎, 「社会調査」のウソ—リサーチリテラシーのすすめ, 文春新書 110, 2000.