

「調査票の設計～質問項目の作り方」

1 調査票とは？

調査票とは、調査をするときに得たい情報をリストアップしておくものをいう。項目だけをあげたものを狭義の調査票といい、実際の文章として質問文を形にしたものを質問紙という。もちろん区分は明確ではない。

調査票は、面接（インタビュー）調査で使われる場合もあるし、電話調査で使われる場合もあるし、配票（留置）、郵送、一堂に集めるなどの方法で被調査者自身が記入したものを回収する場合もあるなど、さまざまな使われ方をされる。面接調査と電話調査で使われる場合は調査員が回答を記入するので他計式、それ以外の場合は被調査者自身が記入するので自記式または自計式と呼ばれる。

なお、不特定多数を対象とするインターネット調査は、母集団が不明確（代表性も回収率も不明）なので、研究には使えないと思った方がよい。対象者を限定して、ID とパスワードの入力によって本人確認をするならば、この限りではないが、記名式の調査に限られてしまうという欠点がある。

2 調査の形式ごとの長所と短所

2.1 面接調査の長所

1. 直接に被調査者に面接することにより、標本として指定された本人の応答を確実に捉えることができる
2. 調査票の回収率が高くなる
3. 他の調査法よりも調査票が長くなっても良い
4. 質問をよく被調査者に理解させることができる
5. 必要なら他の関連質問によって回答の真偽を確認することができる

2.2 面接調査の短所

1. 一般に多数の調査員が必要で（一人で調査できるサンプル数には限界がある）、交通費など費用も多くかかる
2. 追求訪問（一度で聞き取りきれなかった場合など、何度も繰り返して調査すること）が必要
3. 調査員によって応答が異なることがある
4. 調査員による不正（捏造）が生じ得る
5. 特殊な階層の人たちには面接できないこともある

2.3 配票調査の長所

1. 面接に比べると追求訪問が少なく済む
2. 面接より回収率がやや高いことが多い
3. 面接より費用が少なく済む

2.4 配票調査の短所

1. 被調査者自身が記入したかどうか確認できない
2. 本人が記入したとしても周囲の意見が影響することが多い。相談したり、他人の意見を聞いて記入される可能性もある
3. 自計式共通の欠点（調査項目が誤解されやすい、記入が不正確になりやすい、記入漏れが生じやすい）がある

2.5 集合調査の長所

1. 調査の説明や条件を被調査者の全部に対して一定にできる
2. 費用は安くて済む
3. 簡便に調査できる
4. 調査員も少数でよい
5. 無記名にできる

2.6 集合調査の短所

1. 学生を教師が集めるような特別の場合を除いて用いにくい
2. 出席率は一般にあまり良くない
3. 自計式共通の欠点がある
4. 被調査者の差を無視することになる*¹
5. 被調査者中の何人かの発言が大きな歪みを引き起こす可能性がある
6. 出席者の日当や交通費を出さねばならない場合は、かえって費用がかさむ

2.7 郵送調査の長所

1. 一般には費用は安くて済む
2. 広い地域に調査票を配布することができる
3. 無記名にできる
4. 面接と違って調査員による偏りは生じない
5. 面接しにくいような特殊階層の人にも調査票を届けられる
6. 返送率や返送速度が、被調査者の性格の一面を表す情報として利用できる

2.8 郵送調査の短所

1. 協力要請の方法に工夫が要る。手紙の書きかたによって記入し返送する意欲が左右され、切手を貼り宛名を書いた返送用の封筒も同封するなど、回収の仕方にも工夫を要する。発送の時期や催促状の書きかたも大事。
2. 調査票の回収率は低い
3. 正しく本人に渡るかどうかわからない
4. 自計式共通の欠点がある
5. 回収が難しいので催促状を何度も出さねばならないし回収に時間がかかる
6. 調査項目が長すぎると記入が忌避されるので調査票の長さが制限される（一般に面接の半分以下）

2.9 電話調査の長所

1. 簡単に迅速に調査できる
2. 面接調査より安上がり
3. 電話帳を使って無作為標本をとりやすい（携帯電話のメールではランダム・デジタル法が使えたが、一般にSPAMとして迷惑な場合があるため、最近不可能になった）
4. 面接しにくい人にも調査できる

*¹ 全員を一堂に集めて説明しても、背景も違えば能力も違うので、説明をどの程度理解できたかということには一人一人差がある。面接調査ならば被調査者の反応を見ながら一人ずつわかってもらうまで手を変え品を変え説明できるが、集合調査ではそこを無視せざるを得ない。

2.10 電話調査の短所

1. 電話を持っている人だけに対象に限られる（現代の日本ではほとんど問題にならないが、20年前には大きな問題だった）
2. 遠隔地を対象にした場合は費用がかさむ
3. 1人あたりの調査時間は短くならざるをえないので、簡単な調査しかできない

3 調査項目の選択に当たって注意すべき点

- 調査目的に関連した、あるいは作業仮説に関連した、必要十分な調査項目を含ませねばならない
- 一般には、調査目的になっている大きい主題をまずいくつかの次元に分解し、次にその各々をさらに細かいいくつかの次元に分解し、といった手続きを繰り返し、最終的に細分されたものが調査項目となり、見出し語あるいは質問の形で調査票に取り入れられる
- 何か興味があり価値がありそうに思われると何でも入れたくなるが、吟味が必要：
 1. その項目でデータが得られるか？
 2. データが得られるとしても分析に使えるか？*2
 3. 全体のバランスからみて重要度が低いのか？
 4. 被調査者に抵抗や反感や困惑を起ささないか？
- 必要最小限+聞きにくいことを聞くための導入的な質問、他の質問の応答を確認するための限られた無駄な質問

4 用語上の注意

- 簡単明瞭かつ正確に被調査者が理解できること。例：年齢でも、ただ「おいくつですか？」という問いでは、満年齢か数え年かわからないし、年まででいいのか何ヶ月まで必要なのかという精度もわからない。被調査者の最低理解水準を基準にしても「わかる」ようにする
- 単純な日常会話的用語、副詞や代名詞に注意する。例：「どんな洋服を着ますか？」「それは何故ですか？」
- 一般名詞と固有名詞に注意する。例：「新聞を何部とっていますか？」では新聞の種類がわからないし、部数を聞くときに客商売の接客用をふくめるかどうか曖昧。
- 被調査者の階層や環境の違いによるイメージの違いに注意する。例：「風呂」といっても下宿している学生などでは銭湯をイメージするかもしれない
- 難しい用語や専門用語は避ける。やむを得ず使う場合は明瞭な定義を与える。例：USの調査で、まったく架空のMetallic Metals法について尋ねた結果、70%の被調査者が「それは連邦か州で調査すべきだ」と答えたという報告がある。被調査者は「ことばがわからない」とは言いたがらない。
- ステレオタイプ的な単語は避ける。例：「左翼」というコトバでイメージするものは、被調査者によって大きく異なる

5 文章上の注意

- 過度な形容詞の使用は避ける
- 被調査者が質問について肯定する傾向にある（“yes” tendency）ことに配慮して文章を工夫する
- とくに複文で誘導的な前置き（威光暗示効果など）にならないよう注意する。例：「世間では ×といわれていますが、あなたは……」という問いは、世間の威光を借りて回答を歪める

*2 この点の確認には、ダミーテーブル（データが得られたとして、どのように集計して作表するのかをイメージするために作ってみる仮の表）が役に立つ。可能ならやるべき。

- 単位によって限定されるような聞き方は避ける。例：「1ヶ月に何冊くらい本を読みますか？」という質問では、年に2～3冊の人は0か1に無理に分類される可能性が大きい
- 文章の意味内容が2つ以上の論点を含んでいる質問（Double-barreled question, 日本語でもそのまま「ダブルバーレル^{*3}」と呼ばれる）は、各論点について1つずつの質問群に分解する。
- 過去の細かい記憶をもとにした質問はしない。
- 否定的語法の質問は曖昧なので避ける。例：「市営動物園の缶詰会社への払い下げは、阻止すべきでしょうか、それともすべきではないと思いますか？」という問いで「すべきではないと思う」という回答は、「払い下げすべきではない」のか「阻止すべきではない」のか曖昧。
- あまりにも突飛な質問はしない。例：唐突に「もし火星に住むとしたら...」というような問いは、調査全体への信頼性を失わせかねない。

6 質問の仕方のタイプ

1. 個人的質問と一般的質問。
2. 意識を聞くのか実態をきくのか。
3. 意見を聞くのか知識を問うのか。知識を問う質問を濾過質問として使うことも。
4. 平常の行動をきくのか、特定日時の行動をきくのか。
5. 単一の質問で聞くか、質問群で捉えるか。
6. 特定質問に yes または no と答えた者に対して、その判定を覆させるような第2の誘導的な質問を發して、第1の質問の yes または no の強さを測る。この第2の質問を biased question という。例：「今度の総選挙には投票に行きますか？ それとも行きませんか？」で「行く」と答えた人に対して、「投票日に雨が降っていたら/投票日に何か用事ができたら/どうしますか？」、「行かない」と答えた人に対して、「知人に誘われたらどうしますか？」と尋ねてみる。言い回しの効果が大きいので難しい。

7 回答形式のタイプ

1. 自由回答質問
2. プリコーディッド自由回答質問^{*4}
3. 回答選択式質問（賛否的 = 2項選択, 品等的 = rating, 質的多肢選択, 量的多肢選択）
4. 序列質問：選択肢を並べておいて、それらに順位をつけさせるもの。
5. 複数選択：データ化の際に注意が必要。

8 スコア化のいろいろ

項目選択と過重の与え方が研究者の恣意によっているようなスコア化によってできた尺度は、任意尺度と呼ばれる。次のような例がある。

- チェイピンの社会経済的地位 (socio-economic status) 尺度：社会経済的地位を「文化的所有, 有効所得, 物質的所有, および地域社会活動への参加の, 普通の平均的標準からみて, 個人または家庭が占める位置」として, 家庭の調度品の各々に任意の点数を与え, 有効所得をアメイン (adult male maintenance の意味) ではかり, 社会的参加度については団体別に名目的会員なら1点, 集会に参加していれば2点, 賛助会員なら3点, 委員なら4点, 役員なら5点といった任意の点数を与え, これら全部の合計を社会経済的地位尺度のスコアとしたもの。

^{*3} barrel は、普通は樽という意味だが、この場合は違う。石川淳志, 佐藤健二, 山田一成 (1998)「見えないものを見る力【社会調査という認識】」八千代出版の, p.284によると、「ちなみにダブル・バーレルとは双筒銃のことで、一度に二つの弾丸が飛び出すしかけになっている」とのこと。

^{*4} 質問形式はまったく自由回答で、被調査者は自由に答えるが、調査者側で予め予想される回答としていくつかのカテゴリを用意しておいて（このことを「プリコーディッド」という）、聞き取ったときの判断でそのカテゴリの回答ボックスのどれかにチェックするもの。

- 態度測定 of 尺度として、
 - ポイント尺度：測定しようとしているトピックについて賛成の態度を表している単語、文章、絵などと反対の態度を表しているそれを多数用意し、前者にそれぞれ+1点、後者にそれぞれ-1点を与えておく。回答者に示して、賛成・同感できるものを選ばせ、その合計点をポイントとする。
 - 序列尺度：一連の単語・文章・絵などを示して、好きな順に並べさせて、それが予め任意に与えておいた順序とどの程度一致するかで態度を採点する。
 - 評定尺度：単一のある意見に対して、大いに賛成・賛成・中立・不賛成・大いに不賛成といった、通常は3段階から7段階までの回答カテゴリーのどれかを選ばせる（5段階の場合は5件法と呼ぶ）。数直線を示して丸をつけさせるのがいいとされる。その後、各回答カテゴリーに任意の点数を与える（例えば、5件法ならば1から5までのスコアを振ることが多い）。
 - 文章尺度：賛成から反対へと順に並んだいくつかの文章からなり、各文章には研究者の主観からなるスコアが与えられていて、回答者がチェックした文章についているスコアがその回答者の得点となる。

各項目への過重の与え方が研究者の恣意によらず、もっとも適当と思われる判定者の一団（学識経験者、そのトピックに精通している集団、被調査者と同じ集団に属する人々など）の総合的判定に委ねる尺度を判定尺度という。判定には第二種尺度の評定法、序列法、一対比較法などが使われる。例えば評定法の場合、測定対象を判定者集団に示して3段階から7段階までのどれかで評定させ、判定者集団全体についての平均値をその測定対象のスコアとする。

1つの次元にのっている複数の質問項目によって1つの尺度が構成されると考え、それらの合成得点として作られる尺度が内的一貫性尺度と呼ばれる。科学的な尺度としては任意尺度や判定尺度よりもすぐれている。また、いくつかの項目の合計得点として得られる尺度に内的一貫性があることが信頼できるためには、慣例的にクロンバックの係数が少なくとも0.7以上であることが必要とされる。内的一貫性尺度の代表的なものとしては、次の3つがある。

- 項目分析によって作られる尺度：測定しようとしている量的特性との相関が高い項目のみを尺度内に採用する項目選択法。典型的には good-poor analysis を行う。手順としては、pre-test において、尺度に組み入れる候補となっている全項目に仮のスコアを与えて合計し、合計得点が第3四分位より大きい上位群と第1四分位より小さい下位群との間で、回答カテゴリーに差があった^{*5}項目を、その尺度を構成する質問項目として採用するものである。採用した項目のそれぞれについて、各回答カテゴリーに与える点数は、(1) 任意に決める、(2) 判定者集団に依頼する、(3) リッカートのシグマ法（正規分布を仮定し、 $z[i] = (y[i-1] - y[i]) / (p[i] - p[i-1])$ ）としてスコアを与える。単純な合計得点をスコアとする場合なら、各カテゴリーに1から順に整数を振るとほとんど変わらないことがリッカートによって示されている）、(4) シーウェルらのシグマ法（式が違うが、リッカートと同じく正規分布を仮定するのでスコアもほとんど同じ）、(5) ギルフォードの方法などがある。
- 尺度分析によって作られる尺度：ガットマンが考案した方法。スケイログラムを用いる。
- 因子分析によって作られる尺度：基本的には同じ因子に分類された項目を採用すると考える。

9 質問票の流れとレイアウト

- 質問の順序の原則
 1. 答えやすい質問は前
 2. 関連する事柄や似ているものは集める（システムティックにつくる）。ただし、それゆえにキャリー・オーバー効果（回答が、それまでの質問項目の影響を受けてしまうこと）が問題となる場合もある。
 3. 対象者を限定する枝分かれ質問（サブクエスチョン）で間違いにくい順番を工夫する
- タイトル：反発を起こすものは避ける
- 調査主体や連絡先の明記
- 挨拶

^{*5} 通常、回答カテゴリーをまとめて二反応形式に直し、上位群において+に反応した人の割合が下位群において+に反応した人の割合よりも統計的に有意に高ければ差があったとみなす。回答カテゴリーが多くて二反応形式に直しにくいときは平均値の差の検定とか順位和検定で有意差をみることもある。

折半法とクロンバックの係数について

質問紙によって何らかの概念の尺度を知ろうとするとき、多くの概念は直接聞き取ることができないので、複数の質問を組み合わせることによって対象者の差異をより細かく把握しようと試みることになる。例えば、自然への親近感を聞き取りたい場合に、

- (1) あなたは自然が好きですか？ 嫌いですか？
(好き, どちらかといえば好き, どちらかといえば嫌い, 嫌い)

だけでは対象者は4群にしか分かれぬ(順序尺度として数値化すると、好きを4点、嫌いを1点として1点から4点の4段階)。しかし、

- (2) 休日に海や山で過ごすのと映画館や遊園地で遊ぶのとどちらが好きですか？
(海や山, どちらかといえば海や山, どちらかといえば映画館や遊園地, 映画館や遊園地)

を加えて、これも「海や山」を4点、「映画館や遊園地」を1点とする順序尺度として扱うことにすれば、(1)と(2)の回答の合計点を計算すると、2点から8点までの7群に回答者が類別される可能性があり、より細かい把握が可能になる。さらに、

- (3) 無人のジャングルで野生生物の観察をする仕事に魅力を感じますか？ それとも感じませんか？
(感じる, どちらかといえば感じる, どちらかといえば感じない, 感じない)

の4点を加えると、3点から12点までの10段階になる。この合計得点を「自然への親近感」を表す尺度として考えてみると、3つの項目は同じ概念を構成する項目(下位概念)として聞き取られているので、互いに回答が同じ傾向になることが期待される。つまり(1)で好きと答えた人なら、(2)では海や山と答える人が多いだろうし、(3)では感じないと答えるよりも感じる人の方が多いだろうと思われる。同じ概念を構成する質問に対して同じ傾向の回答が得られれば、その合計得点によって示される尺度は、信頼性が高いと考えられる。

複数の変数(項目)の関連をみる指標の1つに、相関係数がある。変数 x と変数 y の相関係数 r_{xy} は、 i 番目の人の x に対する回答を x_i 、 y に対する回答を y_i 、 x についての回答の平均値を \bar{x} 、 y についての回答の平均値を \bar{y} 、総回答者数を n と書くことにすれば、

$$r_{xy} = \frac{s_{xy}}{s_x s_y} \quad \text{但し } s_x = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n-1}} \quad s_y = \sqrt{\frac{\sum_i (y_i - \bar{y})^2}{n-1}} \quad s_{xy} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

として定義される。相関係数は-1から1までの値をとり、まったく無関係なとき0となり、 (x_i, y_i) を xy 平面にプロットしたときに傾きがプラスの直線上に完全に乗るとき1となる。

上記3つの質問に対して一貫した答えが得られたかどうかを調べる方法の1つに折半法がある。例えば質問(1)と(2)の合計点の変数 x_{12} と質問(3)の点の変数 x_3 という具合に、同じ概念を構成する全質問を2つにわけて、 x_{12} と x_3 の相関係数を $r_{x_{12}x_3}$ とすれば、これらの質問の信頼性係数 $\alpha_{x_{12}x_3}$ は、 $\alpha_{x_{12}x_3} = \frac{2r_{x_{12}x_3}}{1+r_{x_{12}x_3}}$ となるというのがスピアマン・ブラウンの公式である(ふつうは、奇数番目の項目と偶数番目の項目に二分)。

しかし、(1)の点と(2)と(3)の合計点という分け方もあるわけで、下位概念が3つ以上ある質問だったら、これらの回答に一貫して同じ傾向があるかどうかをスピアマン・ブラウンの公式で出そうと思うと、 α の値はいくつもの(n 項目だったら n 項目を2つに分ける組み合わせの数だけ)計算される。この場合だったら、 $\alpha_{x_1x_2}$ 、 $\alpha_{x_1x_3}$ も計算しなくてはならないことになる。

それをまとめてしまおうというのがクロンバックの α で、仮に(1)(2)(3)の合計得点が「自然への親近感」を表す変数 x_t だとして、(1)(2)(3)の得点をそれぞれ変数 x_1, x_2, x_3 とすれば、クロンバックの α は、

$$\alpha = \frac{3}{3-1} \left(1 - \frac{s_{x_1}^2 + s_{x_2}^2 + s_{x_3}^2}{s_{x_t}^2} \right)$$

として計算される。クロンバックの α が0.8以上なら十分な、0.7でもまあまあの、内的一貫性(信頼性)がその項目群にはあるとみなされる。なお、クロンバックの α は、考えられるすべての組み合わせについてスピアマン・ブラウンの公式で計算される α を求め、その平均値をとった場合と同じ値を示す。

- 記入上の注意
- 調査票についての処理の記録欄：コーディングで使う
- 小見出しや説明：対象者に調査の順序をわかってもらうための説明
- 質問番号：論理的階層性が明確な方がよい
- 回答上の指示：【 】に入れるとが書体を変えるなど，質問との区別がはっきりするように。
- お礼の挨拶
- 調査員判定
- 最終レイアウトとページ数：最後のページがだいたい一杯におさまるようなレイアウトにし，通しのページ番号を振る。

10 参考文献

- 森岡清志 (1998) 「ガイドブック社会調査」(日本評論社)
- 福武直 (1984) 「社会調査 補訂版」(岩波書店)
- 池田央 (1980) 「調査と測定」(新曜社)