

問題点を踏まえた上での Web調査の利活用

京都先端科学大学人文学部 学部長・教授

佐藤嘉倫

Web調査に関する提言の発出

2020年7月に私が委員長を務めた日本学会会議社会学委員会Web調査の課題に関する検討分科会が「Web調査の有効な学術的活用を目指して」という提言¹を発出した。この分科会の設置目的は、多くの人々がその問題点を意識せずにWeb調査を用いていることに危惧を覚え、その問題点、とりわけ無作為抽出標本を用いていないことを指摘し、安易なWeb調査利用に警告を発することだった。

しかし分科会で議論を重ねるにつれ、「無作為抽出標本を用いている従来型社会調査のほうが調査会社のモニター標本を用いているWeb調査よりも優れている」とは言えないこと、また「Web調査の特性を活用すれば従来型社会調査では収集できない情報を集められる」ということが明らかになってきた。そこで分科会ではWeb調査の問題点を踏まえた上で利活用することを提言することになり、タイトルに「有効な学術的活用」という文言を入れることになった。本稿ではこの提言に基づいてWeb調査の利活用の解説をする。

「総誤差」という発想

総誤差という概念を従来型社会調査とWeb調査に適用すると、両者の優劣は簡単には決められないことが分かる。以下ではまず総誤差の概念を説

図1 総誤差の定義

$$\begin{aligned} \text{総誤差} &= \text{観察誤差} + \text{非観察誤差} \\ \text{非観察誤差} &= \text{カバレッジ誤差} + \text{標本抽出誤差} + \text{無回答誤差} \end{aligned}$$

明して、それをを用いて両者の比較を行う。

総誤差の定義を定式化すると図1のようになる²。

総誤差は観察誤差と非観察誤差からなる。前者は、ある調査項目の真の値と回答者の回答とのずれを表している。調査員効果から生じる誤差がその典型例である。

後者の非観察誤差は3種類の誤差からなる。第1のカバレッジ誤差は調査の標本抽出フレームがどれだけカバーしているかによって決まる。これは従来型調査でもWeb型調査でも生じうる。たとえば京都市住民を母集団とするならば、京都市の住民基本台帳から抽出された標本はカバレッジ誤差が少ない。しかし選挙人名簿だと有権者のみが記載されているのでカバレッジ誤差は大きい。またインターネット非利用者を母集団とするならば、調査会社の登録モニターから抽出された標本はカバレッジ誤差が大きい。

第2の標本抽出誤差は文字どおり標本抽出に伴う誤差のことであり、京都市住民を母集団とする調査で京都市の住民台帳から無作為に抽出され

た標本は標本抽出誤差が小さい。後述するように、Web調査の場合、この誤差が大きくなる傾向がある。

第3の無回答誤差は調査対象者に無回答者が存在することから生じる誤差である。無回答者の属性が偏っていると（たとえば大都市居住の男性若年層）無回答誤差は大きくなる。

従来型社会調査とWeb調査でこれらの誤差を比較すると、必ずしも前者が後者よりも優れているというわけではないことが明らかになる。

観察誤差は一般的に従来型社会調査よりもWeb調査の方が小さい。調査員効果が小さいことに加えて、分岐型質問に関して回答者の混乱を回避することができる。たとえば「いま仕事を探していますか？」という質問に対して、「探している」と回答した人には「どのような仕事を探していますか？」という質問をし、「探していない」と答えた人にはこの質問が表示されないようにして、代わりに「仕事を探していない理由は何ですか？」という質問を表示することができる。もちろん従来型の面接調査や郵送調査でも周到な準備をすれば分岐型質問で

回答者の混乱を回避することはできないが、それでも混乱する可能性をゼロにはできない。

カバレッジ誤差は従来型社会調査の場合、母集団と標本抽出フレームの関係が明確なので小さい。一方、Web調査の場合、調査会社の登録モニターからの抽出なので、上述したように調査対象となる母集団によってはカバレッジ誤差は大きくなる。

従来型社会調査の場合、標本抽出誤差は小さい。これに対して、Web調査の場合、先着順で回答者を受け付けるならば調査テーマに関心を持った人に回答者が偏ってしまうため、標本抽出誤差は大きくなる。

無回答誤差は、現在のように従来型社会調査の回収率が低い状況では大きくなる。Web調査でも、先着順の場合、早く回答した人と遅く回答した人、さらには必要回答者数に達しなかったため回答できなかった人の間の回答パターンの差が大きければ、無回答誤差は大きくなる。

このように従来型社会調査とWeb調査を比較すると、どちらが優れているかは一概には言えない。京都市住民を母集団とし、住民基本台帳から無作為に調査対象者を抽出し、経験豊富な調査員が面接訪問を行い、回収率が100%近い、という理想的な社会調査が存在するならば、従来型社会調査の方がWeb調査よりも優れていると断言できる。しかし今はそのような理想的調査は存在しないし、今後も存在しないだろう。

このように考えると、「従来型社会調査かWeb調査か」という問いは不毛であり、両者の特性を理解して使い分けることが生産的である。以下では、

このことを踏まえてWeb調査の利活用の可能性を探る。

Web調査の積極的な利活用

今までのWeb調査の多くは従来型社会調査よりも安く迅速にデータを集めることができるので用いられてきた。しかしある意味でこのような「消極的な」利用ではWeb調査の利点を十分に生かしていない。ここでは、上述の提言でも言及している利点の中から「従来型社会調査ではアクセスしにくかった対象者へのアクセス」と「センシティブな質問の可能性」について解説する。

第1の利点はWeb調査では大規模な登録モニターのリストを利用できることから生まれる。たとえば家庭内暴力の被害者や性的マイノリティの人々が直面している困難を調査しようとする場合、従来型社会調査ではそのような人々が調査対象者になることは稀である。しかし大規模Web調査ではある程度の回答数を確保することができる。

第2の利点は、Web調査の高い匿名性ゆえに、性的オリエンテーションのようなセンシティブな質問を尋ねることができることである。従来型社会調査では、そのような質問をして回答してもらうことは困難である。

まとめて言えば、Web調査はマイノリティの人々にセンシティブな質問をすることで彼ら・彼女らが直面する課題とその要因を分析することができ、彼ら・彼女らの状況を改善することに貢献する可能性を有している。従来、このテーマの調査は少数の対象者に詳細な聞き取りをする質的調査が多かったが、Web調査は大量データとその

統計分析により、新しい光を当てることができる。質的調査とWeb調査を相補的に用いることで社会的包摂を推進することが期待できる。

むすび

以上、Web調査の問題点と利点を見てきた。ありきたりな結論だが、従来型社会調査、Web調査の長所と短所をよく理解して、自分の研究テーマに適した手法を用いることが肝要である。

なお上記提言では詳しく触れられなかったが、近年注目すべき動向が2つある。第1は無作為抽出した標本を対象にWeb調査を行うという手法が用いられるようになってきたことである³。これは従来型社会調査とWeb調査の「いいとこ取り」であり、総誤差を少なくすることに貢献するだろう。

第2はビッグデータの利活用である⁴。たとえばあるWebサイトを訪問した人々をランダムに対照群と処置群に分けて、異なる質問をしたり異なる環境で回答をしてもらうという、ランダム化比較試験を行うことができる。従来の実験室実験に比べて大規模の実験参加者を確保できる。

このように、Web調査の可能性は広がっているので、積極的な利活用により実験系、調査系の研究のさらなる進展が期待できる。



さとう・よしみち

博士（文学）。専門は社会学。2020年より現職。著書に『AIはどのように社会を変えるか：ソーシャル・キャピタルと格差の視点から』（共編著、東京大学出版会）、『Sociological Foundations of Computational Social Science』（共編著、Springer）など。

1 日本学術会議社会学委員会Web調査の課題に関する検討分科会（2020）Web調査の有効な学術的活用を目指して。https://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-24-t292-3.pdf 2 Tourangeau, R. et al./大隅昇他訳（2013）ウェブ調査の科学：調査計画から分析まで。朝倉書店 3 杉野勇・平沢和司編（2024）無作為抽出ウェブ調査の挑戦。法律文芸社 4 日本学術会議社会学委員会Web調査の課題に関する検討分科会（2023）報告 社会的ビッグデータの利活用に向けて。*COI本記事に関連して開示すべき利益相反はない。