

# 混合研究方法における ウェブ調査利用の可能性と限界

田 島 恵 美

## 1. はじめに

本研究は、個人の直面する課題に関し混合研究法的観点から質的調査と量的調査を併用する際に、ウェブ調査を用いる場合の可能性、問題点、限界を探るものである。

意思決定の際に個人が認識している様々な課題について、対象となる領域を限定しても、なお物理的、社会的、心理的さまざまな側面があり、時として本人も明確に整理できていないこともある。例えば、太陽光発電システムを導入するかどうかの決定には、現在の技術の性能など技術的条件・導入コストや運用コストなどの経済的条件・固定価格買取制度などの社会的条件が関わり、さらには、それぞれの将来の見込みや背景としての環境問題への意識などが複雑に関係しており、切り分けられることなく認識され、意思決定につながっていることもある。これまでこうした多面性を損なうことなく調査を行うとすれば、それは質的調査になり、聞き取り調査などによって、本人の課題についての認識の全体像を示していた。<sup>1)</sup>

しかしながら、そうした事例研究は、一般化することは困難であり、同様の状況にある個人間に共通する課題として、析出することも困難であった。今日、混合研究方法の発展により、プラグマティックな観点から質的研究法ののち量的研究を行うなど、順次もしくは同時並行的に研究を行う方法 (Teddlie & Thshakkori, 2009) が注目されている。こうした方法を援用すれば、リサーチクエスションとして、個人の認識している課題を探索するとともに、量的研究を行い、同様の状況にある複数の個人の認識している課題

の探求へとつなげていくことが可能となってきた。

こうした状況を踏まえ、本稿では混合デザインによって、聞き取り調査など質的調査を行ったのち、もしくは同時に量的調査を行う際に、一般的な質問紙による量的調査ではなく、アクセスパネルを利用したウェブ調査を用いることを検討する。

ウェブ調査については、費用・期間の面で圧倒的な利点がある一方、代表性が問題となるためその扱いに慎重であるべきという議論もある。

ここでは、まず混合研究法のデザインとそこにおける量的調査の位置づけを確認したうえで、ウェブ調査について整理し、混合研究法の質的・量的研究のつなぎめとしてのウェブ調査の可能性と、限界を検討したい。

## 2. 混合研究法における質的・量的研究のデザイン

まずここでは混合研究法のいくつかのデザインの類型を検討しておく。Teddlie & Tashakkoriによると、混合研究法のデザインは「使われた方法論アプローチの数」「工程、または、フェーズの数」「実施プロセスのタイプ」「アプローチを統合する段階」(Teddlie & Thshakkori,2009=2017:101)という4つの基準に基づき、混合研究法単一工程デザイン、複数工程デザインに分けられ、さらに複数工程デザインにおいては、5つのファミリーがあるとされている。<sup>2)</sup> 並列型混合デザイン、順次型混合デザイン、変換型混合デザイン、マルチレベル型混合デザイン、完全統合型混合デザインの5つである。(Teddlie & Thshakkori,2009) これらは、研究に応じ、様々な組み合わせが可能であるとされている。ここでは、議論を複雑にしないため、データ自体の変換については扱わず、質的調査及び量的調査から得られる知見の接合を考察するため、おもに、並列型混合デザイン、順次型混合デザインを念頭に論を進めることとする。

### 3. 混合研究法による量的研究法の利用フェーズ

ここでは、混合研究法で量的調査を利用するフェーズについて検討する。質的研究と量的研究を同時もしくは順次に行う混合研究法において、量的調査を行うとすれば、仮説探索のフェーズと仮説検証のフェーズが考えられる。それぞれの段階を整理しておく。

#### (1) 仮説探索フェーズ

この場合は、仮説の探索を主な目的として、質的・量的調査を単独で、またはそれを併用してを行うものである。この段階では、質的調査も、量的調査もともに仮説を探索することを目的としている。両方の調査法を併用する場合は、調査対象者に関する物理的制約から、時間的に同時に進行できる場合と、どちらかが先行して行われる場合があるが、目的はともに仮説の探索にあるので、両方が終わり、それぞれの分析の結果として探索された仮説が出てくることで、次の段階に移ることになる。

#### (2) 仮説検証フェーズ

混合研究法においては、質的調査において出てきた仮説を、量的調査により検証するまたは、量的調査から導かれた仮説を質的調査によって説明するという研究デザインがある (Teddlie & Thshakkori, 2009)。順次型デザインが多くなると考えられるが、デザインの工夫によって同時に近い形になることも考えられる。つまり、仮説探索フェーズでは質的研究、量的研究のどちらか一方、もしくは両方を用いて行い、仮説が見いだせたら、次に検証段階で仮説探索で使わなかった方法を用いるというものである。<sup>3)</sup>

### 4. ウェブ調査とはどのような調査か

ここでは、混合研究法で、従来の量的調査に代わり利用を検討するウェブ調査について明確にしたい。まずウェブ調査というとき、用語として、二つ

の使用法が混在している。「明確な標本抽出枠に（例えば住民基本台帳や選挙人名簿など）にもとづく無作為抽出標本に依拠しつつ、回答方法をウェブによるものとする調査」（三輪・石田・下瀬川,2020:30）と「調査会社等の登録モニターにより構成される集団（日本ではアクセスパネルなどと呼ばれる）などから標本抽出を行う調査」（三輪・石田・下瀬川,2020:30）の2つである。

本論文で扱うウェブ調査とは後者のものである。これは、各調査会社による募集に対し、自発的に応募した人からなるモニターを大量に確保しておき、その中から、対象となる属性を、登録情報をもとに抽出して依頼する調査のことである。

こうしたアクセスパネルによるウェブ調査の利点としては、まず第一に調査費用の低さがあげられ、第二に調査にかかる時間の節約があげられる。アクセスパネルの利用の有無にかかわらずウェブ調査全体に言えることであるが、すでにデータ化されている状態で上がってくるので、データ化する費用・時間が節約できるうえ、アクセスパネルを利用すれば標本抽出にかかる時間及び費用も節約することが可能となる。

さらに昨今のコロナウィルス感染症拡大の社会状況下で広がった見ず知らずの他者との接触を忌避する風潮という点でも、自分の所有する端末（スマートフォンやパーソナルコンピュータ）を利用してインターネットを通じて回答するウェブ調査は、抵抗感が弱いと考えられる。

またアクセスパネルに登録しているモニターは、その事情はさまざまであるが<sup>4)</sup>、自発的に調査に協力してもよいという前提で登録しているので、自分の情報を他者に提供するという抵抗感がなく、調査協力者として貴重な存在である。<sup>5)</sup>

一方、こうしたアクセスパネルを利用することは、量的社会調査が重視してきた、母集団からの無作為抽出によって得られたデータから、統計学的に一定の誤差をもって母集団全体の動向を推測するということが成立せず、その代表性については疑念が付きまとい、データの質に留保をつけたうえでの公表という事態も多くなっている（三輪・石田・下瀬川,2020:30）。

## 5. 代表性の問題

次にこうしたアクセスパネルを利用したウェブ調査の問題を整理しておく。

ウェブ調査で得られるデータに影響を及ぼす誤差を分類するには、総調査誤差（TSE:Total survey error）の枠組みが役立つという（三輪・石田・下瀬川,2020:40）。総調査誤差には、観測されないことによる非観測誤差と観測誤差からなる（吉岡,2020:52）。

非観測誤差は、以下の三つに整理することができる（Tourangeau, Conrad & Couper 2013 = 2019 : 32）。

- 1) カバレッジ誤差：調査で代表される母集団（たとえばインターネットにアクセスできる成人母集団）と調査のための目標母集団（たとえば、全成人の母集団）との差違
- 2) 抽出誤差：調査母集団（たとえば、インターネット利用者）と、募集によりパネルに登録され、特定の調査に参加するよう抽出された人びととの差違
- 3) 無回答誤差：調査のために抽出された人びとと、実際に回答した人びととの差違

それぞれの誤差について、アクセスパネルを利用する場合について具体的にみていこう。1) のカバレッジ誤差であるが、アクセスパネルを利用するにせよ、しないにせよ、ウェブ調査という手法をとる以上、避けることはできない。総務省の調査によれば、2019年のインターネット利用率（個人）は、89.8%であるが、年齢別にみると13歳から69歳では9割を超えているのに対し、それ以上の年齢では減少傾向にあり、80歳以上では57.5%となっている（総務省 2020）。したがって、特に高齢者を対象とした場合は、ウェブ調査を用いることでカバレッジ誤差が大きくなると考えられる。

2) の抽出誤差は、アクセスパネルを用いる場合に大きく関係してくる。アクセスパネルの募集方法には、公募型と非公募集があり、多くのアクセスパネルは、調査会社による募集や勧誘によってモニターを集める公募型である（吉村,2020:67）。また、登録者集団から具体的にどのように回答者を

定めるかという方法でも、多くのアクセスパネルでは「目標回答数充当法」(轟・歸山,2014:47)をとっている。これは非確率抽出法である割り当て法が用いられている。たとえば、住民基本台帳人口を母集団として、地域(各自治体)、性別、年齢階級を組み合わせたセルを作り、母集団におけるサイズに比例して標本サイズを割り当て、条件に該当する回答者に回答を求めるといった方法である。さらに、条件に該当する回答者には、先着順に求める標本数に達するまで受け付けるという先着型で行われる場合も多い(轟・歸山,2014:47)。こうした場合には、母集団の推定にかかわる統計学的根拠を有していないことになる。ここがアクセスパネルを用いるときの最も問題となる点である。

最後に3)の無回答誤差であるが、回答するアクセスパネルの人数のうち、実際に回答した割合を回答率としてあらわすことができるが、抽出誤差が明確にできない以上、回答率がわかっても、それをもって、母集団とのずれをあらわす指標とはなりえない。

以上のように、アクセスパネルを利用することの問題点は、特に非確率抽出法を用いることにより明確にできない抽出誤差の存在にあるといえる。

## 6. 非確率抽出データの利用法

このように、抽出誤差が確定できない以上、使うべきではないという議論がある一方で、現実の調査で確率抽出が困難になりつつあり、「確率サンプルと非確率サンプルは多くの研究者が思っているほど違いはない」(Salganik, 2018=2019:104)と論じるものもいる。もちろん、そのための工夫なしにあまり違わないということとはできないが、いくつかの工夫をすることで、アクセスパネルのデータをもって、全体の推定に結びつけようとする試みもある。

三輪らは、Kohler(2019)の議論を引用しながら、非確率標本が利用できる条件として「一定の同質性を想定できるような限定された母集団を設定できるかに依存する」(三輪・石田・下瀬川,2020:44)としている。

一方、Salganik は、事後層化付きマルチレベル回帰を用いた分析によって調整した Wang (2015) の研究をあげ、その成果を強調している。(Salganik, 2018=2019:105-109) サンプル数が十分大きければ、事後層化に用いる集団の種類を多くすることで、同質性の維持と集団内にサンプルのないカテゴリーをなくすことを両立できる。またマルチレベル回帰を併用することで、サンプルのないカテゴリーも推定することを可能とし、結果として従来型の電話調査と同じような結果が得られたとしている (Salganik, 2018=2019:105-109)。

このように、非確率標本では、同質性を想定できるような母集団でとったデータを、目標母集団についての付加的な情報を用いて補正することで、利用することが可能となる。

## 7. 標本抽出の偏りの補正

ここでは、調査参加者に対して割り当てる重みを調整することによって偏りを補正する方法について検討する。

Tourangeau らは、こうした方法として、事後層化・サンプルマッチング・レイキング法・一般化回帰モデルによる加重平均・傾向スコア法をあげている (Tourangeau, Conrad & Couper 2013 = 2019 : 32-37)。

事後層化とは、「標本総計 (sample total) が母集団総計 (population total) にセル単位で一致するように標本加重を」調整する (Tourangeau, Conrad & Couper 2013 = 2019 : 33) ものである。事後層化では、調査に回答する確率 (参加確率) が、調査変数と無関係の場合には、偏りを除去できる (Tourangeau, Conrad & Couper 2013 = 2019 : 33)。

事後層化に類似した手法であるサンプルマッチングは、事前に母集団に近いように割り付けを行う手法である。

レイキング法も、他から得た母集団特性値に合うように標本の重みを調整するが、セル総計ではなく、周辺和に対して、調整する。

一般化回帰モデルによる加重調整は、標本推定値を 1 組の共変量との線

形関係を仮定して、調整する方法である。

この3つの方法は、標本とは別で得られた目的母集団の変数分布のデータが必要になる（Tourangeau, Conrad & Couper 2013 = 2019 : 32-39）。

これに対し、傾向スコア調整法は、カバレッジや抽出のバイアスがない参照調査が必要となるが、共変量のすべてを傾向スコアで表すことができるとしている。

これらの有効性を検討し、Tourangeauらは、次のように結論付けている。

「全体としては、ウェブ標本に固有のカバレッジの偏りと選択バイアスの調整を行うことに関して、加重調整は有用ではあるが、必ずしも正確ではない」（Tourangeau, Conrad & Couper 2013 = 2019:42-43）としており、「ここで上げたような偏りを部分的にしか解決できない」（Tourangeau, Conrad & Couper 2013 = 2019 : 43）としている。

星野は、上のレイキング法については、共変量として5~6変数しか調整できない点が問題であるとしている。一般化回帰モデルは、回帰関数を誤って設定した場合にうまくいかないのが、共変量の候補が多数存在する可能性が高いインターネット調査ではうまくいかないとしている。割付を事前に行う方法については、割付を行った方が、行わなかったものよりかえって、バイアスが大きかったという実験調査の結果を紹介している。（星野,2009:179-180）

このような検討から、星野はインターネット調査の補正方法について次のように述べている。まず補正の目的となる調査項目、さらにその調査項目に関連があると考えられる項目についてインターネット調査を行い、これと同様の調査項目の従来型調査を行う。その結果を用いて、共変量調整に必要な共変量のセットを選択し、補正によって従来型調査を十分予測できるか確認したうえで、インターネットでの主調査を行い、その得られたデータを先ほどのスコアを用いて調整するというものである（星野,2009:180-182）。

また6でSalganikの紹介をあげたが、Wangら（2015）は、非代表的な標本を用いて選挙結果の予測を行っている。これは米国のXboxユーザーを対象にしたもので、実際の有権者は47%が男性なのに対し、Xboxの標本では93%が男性というように偏ったものであった（Wang 2015:983）。この

偏ったデータを補正したのであるが、34万5858人の情報をもとに、性別2カテゴリー、人種4カテゴリー、州51カテゴリー…というように細かくカテゴリー化して、176,256のセルを作った(Wang 2015:984-5)。そして空になってしまったセルは、マルチレベル回帰を併用することで推定している(Wang 2015)。サンプル数が十分大きければ、事後層化に用いる集団の種類を多くすることで、同質性の維持と集団内にサンプルのないカテゴリーをなくすことが両立できる。

このように、事後層化でかなり同質性を維持できるレベルまでカテゴリーを細分化することに耐えられるような規模の大きな調査をインターネット上で行うこと、もしくは、傾向スコア法を用い、実験的ウェブ調査、実験的従来型調査を経て、ウェブ上で本調査を行い、補正を行うことが、非確率抽出における標本の偏りの補正として、可能な方法であると考えられる。

## 8. 仮説探索フェーズにおけるウェブ調査の利用

次に2で述べた、混合研究法の各フェーズにおいて、アクセスパネルを用いたウェブ調査が利用可能かどうか、それぞれ検討を行う。本節では、仮説探索フェーズにおけるウェブ調査の利用について検討する。

仮説探索フェーズの段階で両研究法を併用する場合は、量的研究法の扱いは、量的研究法のみを行う場合の、統計調査の探索的利用となる(森岡, 2007:69-70)。この場合、検証は、後の段階で行うため、代表性は保証されなくてもかまわないので、ウェブ調査の結果をもとに仮説を提示することは可能である。

また、アクセスパネルを利用した調査を、予備調査として利用することの可能性について轟・歸山は、変数間の関連の有無を吟味する目的を持つ調査研究を想定し、「新しく着想した質問項目に関する興味深い変数間関連を予備調査で見出したときには、それを本調査でも設定することは適切」(轟・歸山, 2014:48)だと判断できるが、そのための基準として「既知の変数間関連が予備調査で確認できるならば、その予備調査の質は高い」(轟・歸

山,2014:48-49)と判断され、未知の変数間関連について仮説として本調査に設定できるとしている。実際轟・歸山らはSSPプロジェクト(階層と社会意識研究プロジェクト)の一環として行われた全国規模の個別面接法調査と全国を対象とした調査会社の登録モニター(公募型)を用いた調査を比較した研究を行い、各意識変数の分布では2つの調査に差異がみられたが、意識変数の2変数間の関連、および属性変数と各意識変数との間の2変数間の関連では、差異がある結果は少なくなったという(轟・歸山,2014)。もちろんこのインターネット調査は、セル設定を通常の二段階無作為抽出による計画標本と外形的に等しくなる工夫をしているため、これをもって一般化できないが、アクセスパネルを利用し、目標回答充当法であっても、本調査で利用できる信頼のおける仮説の探索につながる可能性を示している。

質的調査と併用して、ウェブ調査を利用する場合であっても、ウェブ調査の質問項目に既知の変数間関連の項目を入れること、セル設定の時点で割付に工夫を行うことで、検証に耐えうる仮説を見出すことができると考えられる。

以上のように、代表性の問題はあっても、仮説を探索するフェーズにおいては、ウェブ調査は利用可能であり、工夫によって検証に耐えうる仮説を生み出すことができる。

## 9. 仮説検証フェーズにおけるウェブ調査の利用

次に仮説探索の段階で生み出された仮説について、検証の段階でウェブ調査を用いることについて検討を行う。7で検討したように、アクセスパネルを利用したウェブ調査をそのまま用いるのでは、仮説が検証されたことにならない。何らかの補正が必要であると考えられる。

アクセスパネルを利用した調査では標本数に応じて費用が発生するので、同質性維持のために新しく大量の調査を行うことは、現実的ではない。もし、既存のデータで利用可能なものがあるのであれば、事後層化法を用いて、検証のため利用できる可能性はあると考えられる。

そのように考えると、仮説検証段階でアクセスパネルを用いたウェブ調査を行うのであれば、先にあげた星野のウェブ調査→同じ項目での実験的な従来の調査→ウェブ上での本調査という手順を踏むことが、仮説検証においての方法であると考えられる。

## 10. 結論

本論文では、混合研究法のデザインとして、おもに、並列型混合デザイン、順次型混合デザインを念頭に、量的研究法の利用フェーズとして仮説探索段階と、仮説検証段階を整理した。次に本論文で検討するウェブ調査を、アクセスパネルを利用したインターネット調査であると明確にした。アクセスパネルを用いたインターネット調査にはいくつかの利点があるが、代表性についての疑問があるため利用に疑念が示されることもある。代表性の問題を整理し、主に抽出誤差によるものとした。非確率データ利用の可能性としては、同質性を想定できるような母集団でとったデータを、目標母集団の付加的な情報を用いることで、補正することがあげられた。そこで補正するための方法についての検討を行った。

そのうえで、仮説探索段階と仮説検証段階それぞれにおいて、このアクセスパネルを用いたインターネット調査の利用可能性と限界を検討した。

仮説探索段階においては、アクセスパネルを用いた場合でも割り付けに工夫を行い、既知の変数間関連の項目を入れることで、検証に耐えうる仮説を見出すことができると考えられた。

仮説検証段階については、ウェブ調査→同じ項目での実験的な従来の調査→ウェブ上での本調査という手順を踏むことで補正できると考えられた。もしアクセスパネルを用いたインターネット調査を混合研究法の仮説検証段階で行うとすれば、事前にウェブ調査と従来の調査をしておく必要がある。

以上のようなことから、アクセスパネルを用いたインターネット調査は仮説の探索には有効であり、十分な研究の設計が可能であれば仮説検証段階でも用いることができることがわかった。今後、実際のデータ補正を通じて補正方法に関する知見を蓄積していくとともに、混合研究法における詳細な利

手順や研究方法の設計について、実際の研究で用いながら、検討していきたいと考えている。

## 註

- 1) 電力システムなど科学技術のシステムの形成・変容に対して科学技術の社会学は、それを技術的要素と社会的要素に分け社会的要素のみを取り出して課題として分析するのではなく、技術的成果物と社会的成果物を区別なく扱おうと試みている（たとえば、Hughes, 1983=1996）。本稿では個人が直面する課題の認識と意思決定とくに技術的・経済的・社会的条件が複雑に絡み合う中での認識と意思決定に関心があるが、そうした状況を扱う際にも、このような観点は有効であると思われる。
- 2) ここで Teddlie & Thshakkori はファミリーという言う用語を使っている。これは「各デザインとも他のデザインの特性に基づき、あらゆる組み合わせが可能であることを示している」（Teddlie & Thshakkori, 2009 = 2017: 109）ためである。
- 3) 仮説探索段階で質的研究・量的研究双方を利用した場合は、検証段階で質的・量的どちらかまたは双方を利用することになる。その場合、類型としては検証段階での質的方法利用なら説明的順次型、量的方法利用なら、検証的順次型、両方を利用するなら、完全統合型になると考えられる。しかしながら、混合研究法は研究法自体の洗練よりも、いかにリサーチクエスチョンに答えうるかというプラグマティックな観点が重視される。したがってあまり精密に場合分けしたすべての可能性をここでは論じないことにする。
- 4) アクセスパネルは、会社によって特性があるが、謝礼目的の人のほか、ひまつぶし目的の人もいと指摘されている（杉野・小内, 2020）。
- 5) もちろん、自発的に調査に登録してくれるとはいえ、4) で見たように動機は様々であり、中には複数の会社にモニター登録するプロの回答者も存在している。協力的である点はメリットであるが、同時にデータの質に影響を与えるこのような存在であり、注意が必要である。

## 参考文献

- Groves , Robert M. , and Lars Lyberg , 2010, “Total Survey Error : Past ,Present, and Future.” *Public Opinion Quarterly* 74 (5) :849-79.  
DOI:10.1093/poq/nfr057
- 星野崇宏 ,2009, 『調査観察データの統計科学』 岩波書店
- Hughes , Thomas P. , 1983, *Networks of Power Electrification in Western Society, 1880-1930* , The Johns Hopkins University Press , 市場泰男訳  
1996 『電力の歴史』 平凡社
- Kohler ,Ulrich, 2019, “Possible Uses of Nonprobability Sampling for the Social Science” *Survey Methods : Insight from the Field*,DOI:10.13094/SMIF-2019-00014
- 三輪哲・石田賢示・下瀬川陽 ,2020, 「社会科学におけるインターネット調査の可能性と課題」『社会学評論』71 (1) :29-49
- 森岡清志 編著 2007 『ガイドブック社会調査 第2版』日本評論社.
- Roger Tourangeau ,Frederick Conrad ,and Mick Couper 2013 *The Science of Web Surveys*, Oxford University Press ,2019 大隅昇・鳩真紀子・井田潤治・小野裕亮 訳『ウェブ調査の科学』朝倉書店.
- 総務省 『令和2年 情報通信白書』
- Salganik , Matthew J. , 2018, *Bit By Bit* (2019 瀬川裕貴・常松淳・阪本拓人・大林真也 訳 『ビット・バイ・ビット デジタル社会調査入門』有斐閣 )
- 杉野勇・小内透 , 2020 , 「特集に寄せて ―インターネット上での社会調査を再考する」『社会学評論』71 (1) :18-28
- Teddlie ,Carles & Thshakkori, Abbas ,2009, *Foundation of Mixed Methods Research*,
- 土屋敦・八田太一・藤田みさお監訳 (2017) 『混合研究法の基礎 社会・行動科学の量的・質的アプローチの統合』 (西村書店)
- 轟亮・歸山亜紀 , 2014, 「予備調査としてのインターネット調査の可能性」『社会と調査』12:46-61
- Wang, Wei ,David Rothschild, Sharad Goel and Andrew Gelman,

2015, “Forecasting Elections with Non-Representative Polls “ ,  
*International Journal of Forecasting* ,31 (3) 980-91.DOI:10.1016/  
j.ijforecast.2014.06.001

吉村作治,2020, 「ウェブ調査の結果はなぜ偏るのか」『社会学評論』71 (1)  
65-83

吉岡洋介,2020, 「インターネットパネル調査の利用可能性 —大学生を対象とした調査事例を通して」『社会学評論』71 (1) :50-64