輿論科学協会創立 65 周年記念講演

ウェブ調査とはなにか？
ー可能性、限界そして課題ー（その 1）

大隅　昇
（統計数理研究所 名誉教授）

◎ウェブ調査 ― 激動のこの十年

さて前置きですが、実は2001年にもここでお話しさせていただいております。つまり、インターネット調査がまだ初期の頃にお話しをさせていただきました。それから10年ほど経ったのですが、この10年はかなり激動の時期であったと思います。特にウェブ調査はすでに実用場面で広く普及、利用されておりますが、いろいろな面で賛否両論があります。市場調査分野の一部の方々からは、我々調査方法論の研究者たちからもう学ぶものはないと指摘されることもありました。言い方を変えると、調査方法論研究に関わる人たちは、ウェブ調査にあまり関心を持っていないのではないかという懸念も持っております。一方、本来の調査方法論から逸脱した所で議論しているのではないか、という気もいたします。実証研究が必要なのだけれどもなかなか進んでいない、とも考えています。

国内ではあまり使われない用語でしょうが、ここでまず“調査方式（調査モード；survey mode）”という言葉を使います。ウェブ調査は、一言で定義すれば、単なる“調査方式の一つ”です。電話や郵送に並び称される調査方式の一つに過ぎません。IT を使っているから優れていると思われる向きもあるかもしれませんが、私の印象では、システム化、電子化が進んだことで非常に複雑になっただけ、むしろ話しが繁雑になっているとみております。つまり見えない部分がたくさん出てきている。ですから、いまだ検証が十分でないと認識しているのですが、商品化が先行し、いわばウェブ調査は即マーケティング・リサーチの商品という形で認知されている、そういう中で議論が進んでいるわけです。しかもこれから私がお話しするウェブ調査、インターネット調査は1995年代半ばに登場のウェブ（WWW）、そしてインターネットの登場以降に進展した非常に若い未成熟のツールなのです。つまりいろいろと問題があって当然です。一番の懸念要素は、敢えて申しますが、多くの場合は“人を対象”にした調査であることを忘れているのではないかと思われることです。人という非常に微妙なもの、扱いにくい対象であるのに、測定技術だけはどんどん進んでいるようにみえるということです。

今日は、“調査方式”としてウェブ調査はどういう位置づけにあるかをお話ししようと思います。

特に大事なことは、2つの要素、つまりまず誰をどう選ぶかということ、もう一つは調査方式と

本日のトーク

- どのような**“調査方式（調査モード）”**なのか，独自に行った**“実験調査例”**で俯瞰する．
- **“ウェブ調査とはなにか”**の基本的な部分を述べる．
- 重要な2つの要素：**“「調査対象者選出」と「調査方式」”**の関係について概観する．
- とくに**“調査方式”**をどう考えるか．
- ウェブ調査の位置づけ，仕組み，特性（利点・欠点）．
- 公募型と非公募型，または**“確率的アプローチと非確率的アプローチ”**という見方．
- **“標準化された実験調査による検証”**の重要性．
- 日頃，理想論を述べていると言われてきたが，事例を見ていただきウェブ調査をどう考えるか，を問いかけたい．

してウェブ調査をどう運用する、データ収集方式としてどう使うか、この２点だと思っています。当然のことなのですが、調査方式なるものがどういう位置づけにあるかがあります（後述）。さらにここで重要なキーは、確率的アプローチであるか、非確率的アプローチかということです。結論を先に申せば、ウェブ調査は“多くは”非確率的です。確率的アプローチに近づきたいのだけれども、なかなか近づけない状態にあるということです。もう一つ重要なことは、“実験調査が必要だ”ということです。ここでいう実験調査とは“標準化”が必要だということです。標準化ということは、平易に言えば、調査票や質問文をそろえ、調査時点をそろえ、対象をどう選んだかもはっきりさせて行うということですが、こうした調査が国内では非常に少ないということです。

今日のお話しに関係する資料を、ここに 10 点ほどあげさせていただきました（注 1）。これは後でお読みいただくことにして…、いろいろある話題に関係することですが、この私のサイト（http://wordminer.comquest.co.jp/index.html）を参考にしていただくと、関連するペーパー、資料がダウンロードできるようになっておりますので、ぜひご覧いただければと思います。

さて、呼称なのですが、これがいままで多少混乱しておりました。初期の頃はメール調査・電子メール調査という言い方をしておりましたが、ウェブの登場によりインターネットと電子メールが混用されるようになった。またオンライン調査、オンライン・リサーチという言い方もあります。ここでは、ウェブの登場以降の調査方式ということで、“ウェブ調査（Web-based survey、Web survey）”という言い方をさせていただきます。ネット調査、ネットリサーチなどの言い方が新聞記事などでは好まれますが、これは一つの方言で、外国の研究者に「ネットサーベイ」と言ってもおそらくは通じませんし、こういう言葉はあまり使ってほしくはありません。ともあれここでは、ウェブ調査は“データ収集方式（data collection modes）の一つ”にすぎないということをお話しするわけです。

適用分野を考えると、大別して、市場調査（マーケティング関係の顧客調査など）、世論調査・社会調査、そして我々が行うような学術研究と、いろいろな利用面があると思います。大きく分けるといくつかのキーがある、特に時間と経費です。それから、正確であるかどうかということ、

調査の代表性、さらに“調査の品質”に関わるということです。これらがキーなのですけれども、インターネットはこれらをどう考えるか、ということがきわめて軽いというか、注意しなければいけないことです。結論を申せば、オンライン・パネルの実態は玉石混交でよく分からない、よって比べなければならないが、それもかなり遅れているということです。はっきりしていることは、我々の立場からいえば、調査の品質をどう担保するかということです。私の今日のお話しは、ここにすべて行き着きます。調査品質の裏返しが調査誤差ですから、こうした誤差をどう取り扱うかということにも関連します。

主な適用分野	相対的に見た主な関心事項と重要度			
	調査時間	調査経費（コスト）	調査の正確さ	代表性
市場調査 消費者・顧客調査	◎	◎	△	△，×
世論調査や社会調査	○，△	△	◎	○，◎
政府，公的機関 学術研究	○，△	△	◎	◎

◎調査方法論の観点からウェブ調査を考える

調査方法論を考察するときに、一般的には大きく２つのアプローチがあると考えております。まず調査方法論の観点から眺めることと、それと、テクニカルな面から電子調査システムをどう構築するかと、という考え方があると思います。ここでは一貫して調査方法論から眺める、つまり、調査データとしてどう利用できるか、という観点から考えてみたいと思います。

なお、調査方法論研究者の視点からみた電子調査システムのあり方という点で、非常に興味ある本としてミシガン大学調査研究センター（SRC：Survey Research Center）のクーパー先生が書かれた本があります（注1、④）。私のサイトに、100ページほどの抄録を作ってありますので、関心のある方はご覧ください（注１、⑤）。

技術的側面から見ると、ご存知のようにウェブ調査は見かけ上いろいろなことができます。見かけ上は、いままでの従来型の調査方法でこんなことができたら、と考えていたことの大半ができるようになったということです。例えば“回答の制御”です。無回答が見かけ上減るわけです。論理的に誤った回答選択肢を選ぶと、「あなたの選んだ選択肢は間違ってます」とウォーニングを出す、あるいは無記入があると「未記入だから書き入れなさい」といった指示ができる。それがよいか悪いかは横において、とにかくいろいろと制御ができる。これを容認する立場からは、いろいろとできますということだが、調査を受ける側から考えると、この項目のこの質問には回答したくないのに回答しろと言われている、ということも起こる。またセンシティブな（微妙な）質問で、「ドラッグをやっているか」と聞かれて、これはちょっと答えたくないなという質問であるのに何かを回答しなければいけない。そうすると、自分が回答したくはない選択肢を無理に選

んだりする“黙従傾向（acquiescence）”のような現象も起こるわけです。このような制御の影響が出てくると、これをどう見るかがあります。こういう研究は、欧米ではかなり進んでいるが、日本ではあまり検討されてはいない、簡単な例をあげればこういう状況です。

調査票デザインでいろいろと綺麗なことができるのですが、そういうことではなくて、調査測定上どういう問題が生じるかということです。繰り返しますが、“調査誤差”（注：総調査誤差、例えばカバレッジ誤差、標本誤差、無回答誤差、測定誤差など）からみたらどうなっているのだろうか、特に非標本誤差（測定誤差、無回答誤差など）です。調査計画時点で検討できる標本誤差ではなくて、問題の多くが非標本誤差の方に移っているということです。これは非常に重要なことです。いまいろいろと言われていることをスライドに書いてみました。まず、IT を使っているから信用できる、これはもはや多くの皆さんそうは思わないでしょう。でも信用している人もかなりいる。回収率が低い、ウェブ調査は一般にこれが低いと言われているのですが、のちに例で見るように、これはケースバイケースです。それに、いろいろなメディアが使っている。他の調査でも回収率が低くなってきているのだから、ウェブでいいじゃないか、という意見が非常に強くなってきています。比べてみたら従来とあまり変わらない結果になりそうだからいいじゃないか、というわけです。私からみると、かなりおかしな話し、誤信です。同時に、大きなリスクを背負わなければいいじゃないかという意見もある。大きなリスクを背負うことがないと、何億円もする調査での失敗は問題だが、数百万円だったらやむをえない、だめだったと諦める、こういう傾向もなくはない。これは研究者にも多いようですが、つまり我々の世界の話しですが、ウェブ調査だからとはじめから馬鹿にしている。また、迅速・廉価・簡便というのがキーワードで、これはどこのサイトのホームページにも必ず書いてあるわけです。これを丸呑みのクライアントが多かったのですが、最近はアメリカのＰ＆Ｇの例にみるように、ウェブ調査はどうもおかしいのではないか疑問を持って用心するようになってきた、という例もあります。

◎ウェブ調査は有用か？

以上を前置きにして、ここでウェブ調査は有用かと問われれば、こういうことをやっている立場上ノーとは言えない、ここはイエスと答えるでしょう。つまりは、否定する決定的な理由もないわけです。いろいろと技術も進んできて魅力的な要素も多々あります。しかも、調査環境の悪化で、従来のような確率的アプローチができなくなっているから、代替措置として使えるのではないか、という言い方になってきています。ですから検証が必要なのですが、ここ十何年かを見ていて、この検証が十分かというと、どうもそうではないと感じられます。

あとで事例を見ていただきますが、およそどういう特性を持った調査方式かということは分かってきています。つまり、どこをどう押さえ改善すればよいかという必要条件は分かってきた。しかしその答えを一言で言えば、それを行うにはお金もかかるということです。つまり「ウェブ調査は安い」と言ってきたことに反してしまうわけです。

こうした点をどう考えるか、後半でちょっと、改善策と言えるものを提示します。ともあれ、継

続的な調査、比較調査、実験調査が必要なのだけれども、一つのキーは“標準化”です。つまり、ただ丸投げの調査をやっても仕方がないわけで、何かある標準化した調査設計でやらなければならないということです。これについても少し考えてみたいと思います。

まさに釈迦に説法となりますが、ここで少し私見を述べます。調査に関して日本は特殊な状況にあったと思っています。実際に外国の研究を見ていると、住民基本台帳や選挙人名簿からサンプリングすることは特殊なケースです。言い方を変えると、日本はこれらが利用できるきわめて優れた状況にあったために、急速な調査環境悪化への対応が十分ではなかったのではないかという風に見えます。北欧の国では、スウェーデンとかフィンランドなどのように住民登録簿が徹底した国々でも、かっての日本のようには自由にこうした情報が使えるわけではないようです。ところが、国内ではご存知のように個人情報保護法などが施行され、マーケティング・リサーチの世界などでは住民基本台帳などほとんど使えず、利用できても研究目的やメディアが限られた中で使うということになってしまったわけです。しかも調査経費が高騰してきた。さらにはいろいろな不祥事も発生したということもあり調査への信頼度はますます低下しております。ということで、全体として私の見るところでは、確率的アプローチから非確率的アプローチへ移行せざるを得ない状況にあるとみえるわけです（後述）。電話調査がよい例だと思います。電話調査を完全な確率的アプローチと思っている人はいないと思います。また市場調査の８割以上がウェブ調査や定性調査に移っているという報告もあります。住民基本台帳はほとんど使わなくなり（というか利用できない）、エリア・サンプリング、ポスティング、電話勧誘、それからスノーボール・サンプリングなどパネルの構築方法はさまざまです。いま話題のウェブ調査の“リバー・サンプリング”という方式、別紙に資料を作っておきましたが、こういうものも登場しております。つまり寄せ集めパネルをそのままにしておくと劣化が進むから、魚釣り方式でその都度使い捨てようというわけです。あまり適切な言い方ではないが、釣っては使って、釣っては使ってということです。当然、これがよいかどうかはいろいろと議論があります。

こうしたことに加え、急速な調査方式の変化、ご存知のようにCATIとかCAPIなど、コンピュータ支援（CA: Computer-assisted）のいろいろな調査方式が出てきているわけです。しかも、多くの場合調査協力度は低いです。つまり、なぜその調査が行われているか、対象者の理解が徹底しないということです。国勢調査を含む官庁統計などでいろいろと市場化テストなども行われています（注：2010年国勢調査では東京都ではウェブ調査も利用された）。このようにいままでのような確率的アプローチには限界があるから、どういう形で別の方向にシフトするかを考えることが一つのキーだと思います。

◎ウェブ調査の現状は…

ではここで、ウェブ調査の現状がどうなっているかを考えてみます。当然、技術要素の急速な変遷があります。特に90年代半ば95、96年頃、ウェブ（WWW）登場以降に急速に進んだ技術です。調査業界も、いわゆるネット調査専業社と、従来方式の調査を行っていて次第にウェブ調

査業務にも参入した企業とに分かれます。技術要素の変化が急速なのですが、実験調査で分かりますが、こうした調査企業のスキルにはかなり差があるようです。例えばウェブ調査用の調査票の作り方にしても、HTML で非常に重いプログラムを書くケースと的確にページが作れるケースが出てくるわけです。こういうところをどう評価するかというと、調査を企画する側も回答する側にも、一向に分からないわけです。つまりどう作られたかに関わりなく見た目は同じ画面（調査票）を見ているわけですから。しかし、目に見えない部分でいろいろな要素が入ってきて、調査の品質が変動しているわけです。さらに、ネット調査企業の再編成の加速現象もあります。例えば最近のメディア記事によると、ヤフーバリューインサイトがマクロミルと併合…、元々インタースコープなどいくつかの企業が一緒になったのにまた一緒になってと、これがどんどん広がるとどうなるのでしょうか。どういうことが起こるかのごく一部を後に少し例でご覧いただきますが、非等質なパネルの融合であり「成り立ちが違うからこそ魅力的な部分もあって一緒になった」との記述は、裏を返せば、どういう属性でどういう特性のサイトが併合されたのかということがなかなか見えにくいということです。そういうパネルが一緒になったという状態が起きている。これは看過できないことです。我々から見ると、非等質な登録者情報を一緒にしてしまってよいのか、どう混ぜるのか（そもそも一緒にできるのか）、現状はどうもバラバラに使っているらしいとも言われている、等々です。

さらに、情報の開示が十分ではないようです。最近はネット専業社でも登録者属性情報などを開示しているのですが、我々から言わせれば、不十分だと思います。同時に調査の大衆化というか、皆さんでもすぐウェブの調査票が作れるようなソフトがたくさん出ております。ちょっと困ったことは、回答するソフト、つまりほぼ自動的に回答してくれるソフトもあるそうです。アメリカなどでは非常に問題になっているようです。こういうことも含めて、さまざまな事象を日本でも後追いで追いかけている、2 サイクル遅れくらいで追いかけているように見えます。米国世論調査協会（AAPOR）や ESOMAR が、これでは困るということで AAPOR タスクレポート（注 1、⑧）やガイドラインを出しており、そこでは非確率的オンライン・パネルを使う時には注意して使うこと、きちんとした調査をやりたかったら利用は控えること、パネルの内容を開示すること、というようなことをレコメンデーションしています。レコメンデーションですから、どう守られるのかは、ちょっと分かりません。もちろんコンプライアンスとか社会的責任などの問題も出てきているわけです。

さて、今日のお話しで使ういくつかのキーワードがあるので、2 点ほど言葉を用意します。まず、目標母集団あるいは“調査母集団”を何にするかです。つぎに“枠母集団”、“標本抽出枠”をどうするか、より具体的にはこれが適切に作れるのかということです。つまりこれが“うまく作れるのか”がウェブ調査の最大の問題なのです。それから、カバレッジ誤差の問題もあります。携帯電話の急速な普及で、電話調査におけるカバレッジ誤差が増えているということに似たことで、ウェブ調査でももちろんそういうことが起きているわけです。ウェブ調査も携帯電話に浸食されて、カバレッジが落ちています。つまりは対象者の選び方と調査方式をどう設定したらよい

か、が最大の問題になるだろうと思います。いわゆる“代表性”、つまり調査母集団に対する標本抽出枠と計画標本の関係、計画標本と回収標本の関係、これらそれぞれのギャップをどう考えたらよいかです。ここらはウェブ調査では十分に議論されておりません。

新聞紙面で以下のような記事を頻繁に見ることがあります。この例では「20 才以上の“禁煙を始めた人”、男女 500 人が回答した」とあります。ではここで、何で 500 人なのか、人口統計学的変数はどうなっているのか、そもそもどうやって集めた誰に聞いたのか、何も書かれていません。しかしそれらしい説明は加えられている。普通の人はこれを見ると、「この調査ではこんなことが分かったのだ」とさほど疑問も持たずそう思うでしょう。ここではそこまで指摘させていただき、それ以上は皆さんに考えていただくことにしましょう。繰り返しますが、重要なキーワードは、母集団の設定と枠母集団あるいはサンプリングフレームの作り方です。計画標本と回収標本の関係、特に回収したあとに計画標本とどう照合するかなどですが、ここらはウェブではほとんど精査されていません（回収結果が報告されるだけ）。調査品質をどう考えるかに関係してくるので、ここでは名前だけあげておきますが、こうしてここにあげたような調査をどう行うかを念頭に、設計しなければいけないのだ、ということだけ指摘しておきます。研究面で大事なことは、調査方法論研究のパラダイム・シフトもあって、調査主体側の調査設計が適切であればうまい調査はできるのではないかという段階から、回答者行動の認知心理学的アプローチ、つまり回答者側がどう調査に参加し回答してくれるかを適切に把握する段階に研究がシフトしてきたと思われることです。これも重要なキーで、調査誤差の発生や測定に関係して状況が変わってきているということです。最近よく言われる“混合方式（ミックスモード；mixed-mode）”などもこういうことの延長線上にあると私は考えております。

◎ウェブ調査の最大の問題 ― 誰をどのようにして選ぶか

まあそういうことで、日本の調査事情全般を前準備として、ウェブ調査をどう考えるかということですが、ご存知のように、誰を選ぶかということが第一の課題、これがウェブ調査の最大の問題です。二番目はどうやって測定するかと、これは測定の問題ですから、どんな手段で回答を集めるかという調査方式の問題です。この 2 つのキーは不可分なのですが、特にウェブではよく意識しないといけないだろうということです。これはあえてコメントしておきたいと思います。対象者の選出方法は、住民基本台帳などきちんとした優れたリストが使えればいいのですが、それはまずだめだということから出発しなければならない。ウェブ調査は、これが曖昧だということです。同時に回答者について、例えば A というパネルを使ったのと、B というパネルを使ったのとでは、回答行動、結果が違うということです。後で例を見ていただきますが、違うことははじめから明らかなわけです。明らかなことをどう測ってよく知るかということが課題なのです。その測り方の点で非常に問題があり、研究が遅れているということです。回収率向上が一つのキーなのですが、後で述べるように、回収率だけではだめで「いろいろな指標が必要」になることもウェブ調査の特徴です。

そもそも、なぜ調査に回答してくれないのかということを考えなければいけないわけです。

もう一つは、こちら側の「調査方式」についてですが、これはどういう手段でデータを集めるかということ、つまり、データ収集方式のお話しです。言い換えれば、標本抽出枠をどう作って、そこからパネルをどう作るか、私の言い方ですと“登録者集団”の作り方です。何度も繰り返しますが、確率的アプローチが次第に困難となっているから、そこをどうクリアするか、できるのかです。そもそも確率的アプローチが難しいことが分かっている中で、それでもなお確率的アプローチに近い場面を念頭に（忘れずに）考えること、調査過程の中で、ここはかろうじて確保できるけれど、ここはやむを得ないなど、そこをきちんと識別するということです。確率的アプローチがどこまで可能かを考えようということです。繰り返し述べているようにIT技術が進んできましたから、P＆P（ペンシル＆ペーパー）や面接から、コンピュータ支援にと、どんどん移ってきて、さまざまな調査研究の方向がある、例えばCADAC（Computer Assisted Data Collection）とか、CASIC（Computer Assisted Survey Information Collection）です。電子化の方向に進むということが、全体の趨勢なわけです。日本は分かりませんが、欧米、少なくとも米国ではそうなっています。

そして日本は、こういう視点から調査方法論を研究しようということが希薄だったのではないかと、個人的にも反省しているわけです。そういう中で、いきなりベンチャービジネスとしてネット調査が登場してきたので、本来あってほしい望まれる調査環境との間にギャップが出てしまったのだと思います（注：ここらの事情は欧米に類似している）。

◎調査方式の類型化とウェブ調査の位置づけ

ここで、調査方式全体がどうなっているか、その中でウェブ調査がどういう位置づけにあるかを大雑把にかつ教科書的に見ておきます。こうした調査方式の類型化は必要なことだと考えるからです。まず、コンピュータ支援の有無、調査員の有無、自記式か非自記式（他記式）か、面接か面接でないか、といった二分法的に分けてみる考え方があります（図「調査方式の類型化（1）」）。もう一つは、時間軸にそって、特に電子技術の進歩に応じてどういう変遷をたどってきたかという見方があります（図「調査方式の類型化（2）」）。

特に、対面による直接的な面接、つまり面接聴取から、電話を通して間接的に対象者から回答を得る電話聴取、それのコンピュータ支援モードである最近の電話調査という非自記式の流れがあります。一方、自記式は、郵送、DBM（Disk by Mail）やファクシミリなどを経て、かなり電子化の進んだ今のウェブ調査にいたる流れがあります。いくつか名前をあげてみると、CATI、CAPIはもとよりCASIとかAudio CASIといろいろな使い方がある。ウェブ調査は自記式であり、最近の携帯電話やスマートフォン（高機能携帯電話）なども使った調査票を用いる方式も、たぶん（図の）ここらあたりにあると、そういう理解です。特に、コンピュータ支援があるかないかで大きな違いが生じます。これをもう少し詳しく述べてみます。

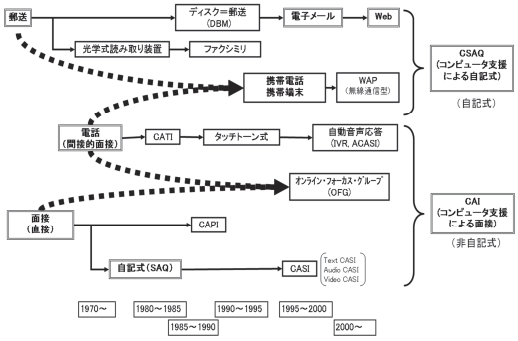
調査方式の類型化(1):形態で分ける

		コンピュータ支援（CA：computer-assisted）の有無			
		なし		あり（CASIC, CADAC）	
		面接員の関与		面接員の関与	
		あり	自記式	あり（CAI）	自記式
調査方式（調査ﾓｰﾄﾞ）	面接	面接 FG	訪問留置	CAPI	CASI Text CASI, Audio CASI, Video CASI
	郵便	…	郵送	…	DBM (Disk by Mail)
	電話	電話	ﾌｧｸｼﾐﾘ	CATI IVR, ACASI	…
	ｲﾝﾀｰﾈｯﾄ	…	…	OFG	ウェブ調査，電子メール 携帯電話，高機能携帯（ｽﾏｰﾄﾌｫﾝ，ﾓﾊﾞｲﾙPC）

CASIC: Computer Assisted Survey Information Collection　CADAC: Computer Assisted Data Collection
CATI: computer-assisted telephone interviewing　CAPI: computer-assisted personal interviewing
CASI: computer-assisted self-interviewing　OFG: online focus group

(†)吉村・大隅の分類をあらためて整理要約した.

調査方式の類型化(2):時間軸で変遷をたどる



(†)Groves et al.(2004), Marcer (2003), Biemer & Lyberg(2003)などを参考に作図

これは「ウェブ調査のルーツはなにか」を考えて作った図です。繰りかえしになりますが、“郵送調査”は間接的な自記式です。この自記式の発展形が“ウェブ調査”だという見方があります。よって、郵送とウェブを使った比較調査は、我々を含めてたくさんの人が試みております。もう一つの流れは、“面接”があるかないかということです。面接からみれば非自記式ということで、“電話”が中間に位置し、これが“面接”と関連していることに注目したいと思います。特に電子化に向かって、（図では）左から右に向かって進んできた、つまり（米国における）70年代の電話調査から始まって、今の電話調査はまったく様子が違うということです。また、ウェブ調査を少し整理してみると、まったく新しい調査方式で、90年代半ばくらいに登場した新しい方式だと言えます。もちろんインターネット環境を使う、そして電子的に回答を集めるから、電子調査システム構築というインフラが必要だということです。なおここで、“おもに”インターネット・ユーザを対象に行う…、としました。この“おもに”という所に引用符をつけた理由は、かりに非ユーザであってもインターネットを介して回答を取ろうという試みもあるということです（後述）。それから、ウェブ調査は「電子調査票を使った間接的な自記式調査」という点が重要です。マルチメディアを使うなど、いろいろなことができますから、調査票作成の自由度が非常に高い。先ほどのクーパー氏の本などを読んでいただくと、いろいろなタイプのいろいろな調査票が作れることが分かります。我々が過去に、こうあったらよい、ああなったらよいと想像していたことがほとんど実現できてしまいます。しかし「それがよいことか？」ということが重要なのです。確かにリアルタイムでデータが取れますから、先ほど述べたような回答制御や記録の保持などトラッキングにより、回答のトレースを行うこと、そういうことも可能です。いわゆる“パラデータ”と言いますがこれの使い方については後で少し例を示します。

ここでは調査システムの細かい仕組みの説明は必要ないでしょう。要点は、いくつかの機能を分担するサーバ・コンピュータといったような物理的な用具があったり、そこをインターネットを介してプログラムでサポートするということです。調査客体つまり調査を受ける側は、ブラウザを介して質問文（調査票）だけを見ているけれども、裏でいろいろなことが行われているということです。このことが重要で、「電子的に…」という一言でくくられてしまうのですが、さまざま

なソフトを使い、またHTMLが基本的に使われているのですけれども、さらに回答者がクリックしたデータをどう取得するかなど、いろいろな箇所でさまざまな細かいことが使われている。これは我々には見えないわけですが、さまざまなノイズが当然入っていると考えるべきです。さらにはシステム保守の問題やセキュリティの問題が出てくるわけです。これは非常に問題がある、プライバシーの問題もある。こういう中で電子調査システムは動いているのだということです。

見かけ上はいろいろな部品を使っていろいろなことができます。特に回答制御や回答分岐処理、それからパイピングといって情報を埋め込んでいく処理、例えば「あなたは先ほどの質問で○○と答えました。」この「○○」を受け取って「ではその○○についてどう思いますか？」といった聞き方もできる。さらに警告画面を出すとか、自動的にロジカルチェックをやる、ページ遷移、改ページやスクロールの制御、プログレス・インジケーター出して「あなたの回答は今ここまで進みましたよ」などといったこともできる。いままで、こうあったらよいなと思っていたことがほとんど実現できる。回答選択も、ラジオ・ボタンで選択肢を一つだけに制限するとか（注：郵送や留置の自記式で選択肢を「ひとつだけ選ぶ」と指定しても複数選ばれてしまう、といったようなことは起こらないということ）、チェック・ボックスを使い複数回答とする、しかもいずれも選ばないと警告を出して回答を促すなども行う。これは非常にシンプルな例で、実際はさらに複雑な調査票設計も可能です。自由回答もテキスト・ボックス内に回答者が自由に書き込み、結果、自由回答が電子的に容易に取得できる等々、です。さらに、使おうと思えば、イメージ（動画、静止画）や音声を用いることも可能です。しかしこうした自由度が高く柔軟な対応ができることが問題なのです（注：前出のクーパー氏の著書でいう、調査票の見かけのデザインと本来あるべき調査票設計とは別物、ということ）。

◎従来の調査方式との違いが鍵

そういうことで、ウェブ調査の測定環境は自由度が高いように見える。こうしたいこと、ああしたいことができるように見えます。調査方法論、我々の立場からいえば、見かけは同じ調査票でも実は測定には違いがある、というわけです。これはちょっとリスクが大きいのではないか、調査の品質を保証することにはならないのでは、となります。そこでウェブ調査の“利点、欠点”を少し整理してみたものが次なのです。

ウェブ調査の利点，欠点とされる主な事項

利点とされてきたこと	欠点とされてきたこと
• 簡単にできる（簡便性） • 調査期間の短縮化（速報性・迅速性） • 廉価（調査経費の低減化） • 登録者集団の作り方で回収率が上がることがある • 回答行動の電子的追跡（トラッキング）が可能 • 調査不能の抑制が可能なことがある（回答制御による警告通知など） • 回答制御の有効利用が考えられる（分岐回答，パイピング等） • 地域性，地理的距離の解消（実際は都市圏に偏ることが多い） • 間接的自記式であるので面接者・調査員の影響を受けない • 自由回答設問設計とその回答取得が容易とされる • 微妙な質問への回答取得可能性が高い（「本音」で答える） • 適切なパネル管理で登録者の高い協力度が期待できる • 双方向的（インタラクティヴ）な利用可能性 • 調査票設計時のカスタマイゼーションの多様性 • 調査票設計の多様性，マルチメディア機能の有効活用（（測定誤差の回避を考慮） • パネル内の登録者との情報授受の容易性（登録者ページを作るなどしてモニタリング可能，登録者との密な情報交換）	• 母集団（目標母集団）が曖昧，分からない • 誰をどう選んだかが曖昧（標本抽出枠が不透明） • 誰を調査したのか（回答の代表性が疑わしい） • 一般に回収率が低い（状況による） • 回答者の顔が見えない，回答者同定の困難性 • 虚偽，代理など不正回答の混入のおそれ • 謝礼目当てのプロ回答者の存在，その混入のおそれ • 回答の制御・強制が起こりうること • 調査誤差の評価が難しい，十分に徹底していない • 有効回答の確定が難しい • 標本設計の困難性（統計的アプローチが困難） • 回答者との信頼性の確保が希薄（合意形成の曖昧性） • 調査不能・無回答の扱い（確認と処理）が複雑となる • ネットワーク利用から生ずる種々のハードウェア上の障害（通信障害，サーバダウンなど） • 回答者のコンピュータ・リテラシーのバラツキの影響 • 回答者の PC，インターネット利用環境のバラツキ（PC 性能，OS，通信回線速度，利用ソフト，…） • パネルの疲労やパネル管理状態が見えない，調査主体・調査対象者間のなれ合い現象など • マルチメディア機能の誤用，濫用の可能性（時には回答誘導のおそれがある）

基本は、従来の調査方式と比べてどういう差があるのかということと、似ているのか似ていないのか、どこがどう違うのかということです。またウェブ調査特有の現象として何か生じるのかといったことです。典型的な例が“調査不能”です。調査不能がいろいろ別の形で出てくる。後で“完答率”という言葉も用いますが、回答したいのに回答できない人もいます。答えたくないのではなくて、答える意思があったのに回答できない、そのようなことも出てきます。ここに表にまとめましたので、見ていただければと思います。要点は、「簡単、迅速、廉価」だけではなくて、いろいろ固有の特性があるが、それらをどうやって組み合わせて使うのか、またそれらを調査方法論から見たときに、“誤差を低減する”という方向で役に立つのか立たないのか、といった判断が大事です。

そこで問題は、繰り返しとなりますがパネルの作り方、これが一言で言ってしまえば、純粋に確率的に、従来のように明確な抽出枠があって、サンプリングをして、誰が選ばれたかがよく分かっていて、どう実査を進めるかというプロセスが「完全には実現できない」ということです。しかしそこから出発しなければならないわけです。これを否定してしまえば、そこで議論は終わりとなります。ずっと調査に関わってきた方々は、私も含め内心忸怩たる気持ちもあるわけですけれども、ウェブ調査をよく知るには、もっと研究しなければいけない要素が、そこらじゅうにあるわけです。純粋にこれだ（完全に実現できること）はないので、確率的な操作がどこまで使えるかを、少し頭の隅に置いていただくということです。

◎登録者集団の作り方 ―「公募型」と「非公募型」

非確率的なアプローチによる登録者の集め方、つまりこれが世の中でいう“ネット調査”です。いわゆるセルフセレクション（自己参加型）といって、“ボランティア・パネル（volunteer panel）”を使う典型例です。つまりこの指とまれ方式で集まった人たちを調査対象とする場合で

す。私はこれに“公募型”と名付けたのですが、これに対してもう１つは“非公募型”、要するに「公募型とは違う」という含みで、また部分的に確率的アプローチが可能か、あるいはそうできそうかを考えると言う意味で非公募型です（少し奇妙な呼称ですが）。この両者の違いをここに一覧にしましたが、要は「確率的アプローチがどの部分でどの程度まで可能か」ということです。のちの例で少しみていただくように、考え方はそう難しいことではありません。

非公募型と公募型の主な特徴

	非公募型	公募型
適用範囲と主な勧誘方法	• 少数だが採用している調査機関がある（例：博報堂 Hi-panel, Knowledge Networks, Gallup） • 従来型の標本抽出法で集め，合意・応諾をとる（エリア・サンプリング，電話，郵送，ポスティングなどによる勧誘）	• 大半のネットリサーチ，ネット調査企業 • 市場調査などで急速に普及し商品化が進む • 勧誘・公募（バナー広告など） • 自己参加，ボランティア／オプトイン
主な特徴	• 従来型標本抽出法でパネルを構築 • そのパネルにインターネットで調査 • 統計的推論（確率的アプローチ）の可能性 • パネルの人口統計学的特性の偏りの原因の推測可能性 • 一般に登録勧誘への応諾率は低い • 代表性に注意，パネルの作り方に依存 • 応諾後の調査協力度は高い • 回収標本は計画標本に近い • 廉価にはできない，高度な基盤整備が必要 • 従来の調査会社・機関がウェブ調査に参入 • ネット調査専業社との調査の考え方の違い	• パネルの作り方が多様で情報が不透明なことが多い • パネルの人口統計学的特性の偏りがある • 統計的推論が難しい（非確率的アプローチ） • 登録者の顔がよく見えないことがある • 一般に回収率が低い，変動が大きく偏りがある • 計画標本と回収標本の差違が大きい • ネット調査専業社が多く調査方法論の知識に懸念要素がある • ネット調査専業社の再編成・併合の傾向 • 登録者情報の名寄せ，重複者管理，個人情報保護など懸念要素の増大
共通したこと	• 個人情報保護，プライバシーなどへの十分な配慮が必要なこと • 従来型の調査機関とネット調査専業社とのコラボレーションや受託が増えてきたこと • 電子調査票作成の技術的水準の向上があるが，ウェブ調査向きの方法論研究が希薄	

公募型とはいわゆるネットリサーチ（ネット調査）ですが、先ほどからお話ししているように、人の集め方、パネルの作り方、誰が対象者でどう選んで、どういう割付でどういう設計で行ったかということがどう行われているかです。先ほど「500 人に聞いた」という例を見ていただきましたが、なぜ 500 人かというと「打ち切り型」により 500 人で回収を打ち切っているからなわけです。またそもそも何人に発信したのか…、もうこの時点で回収率という概念がないことが分かる、こういう状態で使うことはよいのかです。

こういう状況なので、少し焦点がずれた議論もあるわけです。その一つが、登録者数が多くなればよいだろう、スノーボール式にどんどん集めて増やしていけばよい、といった意見です。パネルサイズのビッグ化も進むわけです。登録者数が多くなれば信頼できるという誤信です。後で例を見ますが、登録者数が少なくても高い回収率をあげることはできます。なので、これはちょっと疑問が残ります。

インターネット・ユーザが、仮に 6000 万人、7000 万人になろうとも（注：昨年の時点で 9000 万人を越えていると言われている）、利用者のレベルやスキルがそれぞれ違う、関心も異なる、であるから、こういう問題はそう簡単には解消されないと私は思っています。やはり繰り返し述べるように、調査誤差をどう考えるか、そこにあります。これにどれだけ意識を持って対応するかですが、たぶんほとんど考えられていない。

公募型のスキームを描いてみると、まずインターネット・ユーザが目標母集団であることは明

らかです。そこからどうやって抽出枠を作るか、登録者集団のパネルをどう作るかということです。図の中央にたくさんの楕円が描いてある理由は、A 社、B 社、C 社…と“作り方”がたくさんあるということです。ですから、当然調査結果の答えは違って当たり前なのです。そう思わなければいけないわけですから、1 社だけ調べてもだめなのです。調査依頼して回収した標本、つまり図で曖昧だと描いてある部分が、照合してみると、ズレが顕著に出ることは分かっているのです。多くの場合ここらをきちんと調べていないが、我々の実験調査ではすべて調べております。調べてみると、ズレが出てくるということが分かります。つまり、選出方法をどうしたか、登録方法をどうしたか、パネル疲労がないか、どう作ったのか、誰を選んだのか、誰が回答したのか、を明示的に示す必要があると思うのですが、現状はここらの情報開示は非常に限られているのではないか、ということです。

ここで非公募型の一つの改善案を紹介します。これは実際に実験的に試行し、一部ではそれを吟味した上で実用化しているわけです。つまり調査対象者の選び方くらいはきちんと選ぼうということです。かつては住民基本台帳を使っていたのですが、最近はこれも困難であるから、エリア・サンプリングにより構築した抽出枠から抽出してパネル登録を行うというものです。実際に訪問により確認をとり、つまり「調査に協力してくれますね」という応諾をしっかりとるということです。この方式により不完全ではあるものの部分的にでも確率的なアプローチを担保するということです。ですから登録の仕方によっては普通の人、インターネット・ユーザも非ユーザもいるわけです。その両方にアクセスして応諾をとって、自前の PC を持っていれば使ってもらう、PC の非保有者には貸与するというやり方もあります。このユーザだけでなく“非ユーザ”までを含めてという点がよかったのですが、経費がかかるという理由でできなくなってしまった。実際にこれを行っていた企業が既にやめています（単なるウェブ調査に移行してしまった）。新聞社の A 社さんがやっていたのですけれども、できなくなってしまった。これは復活可能なら本当はやるべきことではないかと残念に思っております。

つまりユーザは誰か、普通の人とは誰なのか、ということをまず考えなければいけないけれども、そこはなかなか保証されないわけです。しかしこれも欠点があって、この方式は、応諾を得た後は、後で例を見るように回収率は非常に高くなり、ロイヤリティも高いわけですが、当然応諾率が低いから、全体で回収率をどう見るかがあります。つまり、統計的推論のみで母集団推論を行おうとした時に、依然として問題がなくはない、解決されません。ただ登録者の顔が見えること、誰を選んだかは分かる。しかも廉価にはできないけれども、従来型よりは経費は安くなることは分かっています。

抽出枠の設計時に、エリア・サンプリングを使ったり、（かっては）住民基本台帳を使うなどして、登録者を勧誘して登録者集団とした非公募型の例としては、例えば博報堂系のパネル（Hi-panel）、電通リサーチのパネル（Rnet-Premium）、アメリカのナレッジネットワークス（Knowledge Networks Inc.）と、若干は見られます。ナレッジネットワークスは米国の学術研究でよく使われているようですがこれは電話調査の抽出枠を利用しているようです。またここは、

調査用具として（一時は）ウェブ TV などを配付していました（注：考え方としては上述のA社の方式に類似）。要するに、なるべく登録者の顔がみえる“抽出枠を自前で作る”こと、インターネット・ユーザだけでなく、“少しでも一般対象者に範囲を拡げる”ことに意味があるわけです。

◎実験調査に必要な「標準化」

まずは、以上のような状況下で調査をどう行うかを考えてきたのですが、さらに重要なキーが“標準化”だと思っています（表「実験調査の指針（標準化の目安）」）。電話調査が標準化された手順にそって行われるように、ウェブ調査もこれが大事なわけで、特に電子調査システムであることを考えると、いろいろなことを標準化しなければならない。

ただし、技術改善の速度が早いし調査機関間のスキルの差違が大きい。これは実際にやってみると分かりますが、こちらがリクエストしてもある調査企業はできない、あるところはできると答える。こういうことの調整・管理も含めて標準化が必要になります。同時に、従来型調査との比較もしなければならない。絶えず比較しておくことは必要になるだろうということです。さらに、誰かが類似した実験調査を追試したい時には、なるべく同じことが再現されなければいけませんから、私の場合はほとんどの情報を開示しており、特に調査設計仕様まですべてを開示しています。

特にウェブ調査の最大の特徴は、調査に関わる諸要素をそれぞれかなり共通化できることなのです。調査時点をまったく同じにするとか、調査票をまったく同じにするとか、いろいろな点で非常にメリットがある、そこを最大限に活かす、またできるだけ繰り返して行う、これも比較的容易なわけです。さらに大事なことは、お願いする調査機関との“合意形成”が重要ということ、つまり調査目的も説明せずに丸投げで調査を行うのではないということです。過去の自己体験や他の類似調査を調べることで得た情報をもとに、次のページ（スライド）を作ってみたのですが、総じて、委託や丸投げで行った実験調査ではまずいのではないか、という印象です。また既に述べたように、パネルの構築方法がさまざまであるから、1社の調査ではだめで複数パネルで比べる必要があるということも明らかです。比較する別の調査方式としては例えば郵送調査です。郵送と比べてどうなっているのかをやってみる必要があり、同時に標準化を行う必要があると考えています。

実験調査の指針（標準化の目安）

実験調査の原則と指針・ガイドライン（概要）	ウェブ調査に有利な点
参加機関との間で諸条件について合意形成を行うこと．調査の意図の説明なく，単なる丸投げ・委託の調査は避ける．	△
可能なかぎり単一機関でなく複数機関の参加が望ましいこと（複数の登録者集団を用いた比較分析の重要性）．	○
登録者集団の構成を明らかにする（誰をどのように集めたか，どのような人口統計学的構成か，どのように管理しているか等の情報を可能なかぎり得る）．	
計画標本の作り方をなるべく統一化する（共通したサンプリング方式の設定）．	△→○
ウェブ調査の特性を活かし，（ほぼ）同時期・同時日程で実施する．	○
調査票，質問文形式などを（ほぼ）揃える（同じ調査票形式とする）．	△→○
調査対象者・回答者の回答行動がみえる調査設計とする．	○
回答制御をできるだけ標準化，共通化し，また何を行ったかを正確に記録する．	◎
回答者行動を電子的に追跡する（トラッキングによるパラデータ収集と詳細分析）．	◎
回収標本と登録者集団，計画標本の照合を行う（各種指標の確認）．	
現場の日常的手順も勘案する（理由：調査経費に影響）．「出来ること」「出来ないこと」を明らかにする．調査経費と調査の品質はトレードオフの関係にある．	
継続的に反復検証を行う，同じ登録パネルで時系列的に追跡すること．	○
できれば従来型の調査方式（郵送，面接など）との比較検証を行うこと，とくに郵送調査との対比は重要．	△

この合意形成について、いくつか例があります。委託した複数の調査会社に何のためにその調査をやるかを伝えないで、比較を行っている、これはちょっとまずいのではないかということです。要するになるべく調査条件をそろえなければいけません。繰り返しますが、同一時点で調査を行うとか、パネルからの抽出方法を共通化するとか、いろいろなところでです。特に一番大事なことが調査票のそろえ方とデータの取り方です。回答行動のトラッキングにより電子的に回答を追跡し“パラデータ”を取るわけですが、これの取り方も仕様で決めることです。確信を持って言えることですが、とにかく調査条件を可能な限りきちんと決めることが肝要ですが、現状はほとんど行われていないと思います。特に研究者が学術研究の一環としてウェブ調査を使うのであれば、絶対にここは守らなければいけないことです。もちろん、この標準化については、いろいろな議論が電話調査の方でもなされてきたわけです。電話調査も非常に高度に電子化が進んでおり、ウェブ調査との類似点もあるのですが、電話調査は見方によっては間接的な面接調査とも考えられます。一方、ウェブ調査は間接的な自記式調査なわけです。この違いがあるのですが、回答者との不自然な相互行為、妙な質問を作るとか、そういうような部分で、当然同じことはあり得るだろうということです。また、標準化したことで、質問のワーディングや設計の全体の文脈がおかしくなったりしたらまずいわけですから、標準化が必ずしもよくないのではないかという意見もずっとあるし、よく言われてきたことです。それから、回答誘導になるのではないかという指摘もある。特にウェブ調査の場合は回答制御が可能ですから、どこまでならやってもよくて、どこまでならだめなのだとの限界がよく見えない。職業選択などで専業主婦を選ぶと、他の質問のその条件下では選んではいけない選択肢を選べないように調査票の画面表示でマスキングをするわけです。そんなことが自由にできてしまうわけです。それをやってよいのか、いろいろな点で、慎重に議論しなければいけない。しかも複雑な調査票設計も行う、例えば前述のパイピング処理などです。後で少し参考例を見ますが、そういうことが、とくに測定誤差がどう介入するの

か、ときには測定できないのではないかといろいろあります。こういうことで、あらゆることを標準化することがよいかどうかも注意する必要があるということだけを指摘しておきます。（次号へ続く）

（注1）関連する主要な資料

①大隅昇，前田忠彦（2008）：インターネット調査の抱える課題－実験調査から見えてきたこと－，「よろん」（日本世論調査協会報）2007；第100号：58－70／同2008；第101号：79－94.

②前田忠彦・大隅昇（2006）：自記式調査における実査方式間の比較研究－Web調査の特徴を調べるための実験的検討－，エストレーラ2006； No.143, 12－19.

③大隅昇（2006）：インターネット調査の抱える課題と今後の展開，エストレーラ2006；No.143: 2－11.

④Couper, M.P.（2008）：*Designing Effective Web Surveys.* New York: Cambridge University Press.

⑤大隅昇（2009）：M.クーパー著『効果的なウェブ調査の設計』を読んで，「よろん」（日本世論調査協会報），第104号，p50-60.（④の書評）

（*）http://wordminer.comquest.co.jp/smr/pdf/f00017b.pdf

⑥Groves, R.M., Fouler, F.J., Couper. M.P. et al.（2004，2009）, *Survey Methodology.* New York: John Wiley.（朝倉書店から2011年春訳本刊行予定）.

⑦M. Couper, Online Research in the U.S.　Yesterday, Today, and Tomorrow -, JMRA 41st Topics Seminar, September 21, 2010,Tokyo.（日本マーケティング・リサーチ協会第41回トピックスセミナー；2010年9月21日開催），セミナー配布資料.

⑧AAPOR Online Panel Task Force Report (2010): AAPOR Report on Online Panels.（約90ページからなるレポート）.（*）http://www.aapor.org/Home.htm

⑨林文，大隅昇，吉野諒三（2010）：ウェブ調査から何を読み取るか―基底意識に関する実験調査―，日本日本行動計量学会，第38回大会，2010年9月23日（埼玉大学），予稿集，30-33.

⑩矢口博之，大隅昇（2010）：電子書籍と読書行動についての実験調査，日本日本行動計量学会，第38回大会，2010年9月23日（埼玉大学），予稿集，26-29.

（※）上の多くは以下のHPから閲覧可能

http://wordminer.comquest.co.jp/index.html

調査方法論関連情報資源：Survey Methodology Resource (SMR)

http://wordminer.comquest.co.jp/smr/index.html

追加文献：

1）De Leeuw, E.D., Hox, J.J., and Dillman, D.A. (2008): *International Handbook of Survey Methodology*, Lawrence Erlbaum Associates.

（*）ここの第6章として“Mixed mode surveys, When and why, pp.299-316”がある.

2）Dillman, D.A., Smyth, J.D., and Christian, L.M. (2009): *Internet, Mail, and Mixed-mode Surveys*, Third Edition John Wiley & Sons.

（おおすみ・のぼる）

輿論科学協会創立 65 周年記念講演

ウェブ調査とはなにか？
ー可能性、限界そして課題ー（その２）

大隅　昇
（統計数理研究所 名誉教授）

◎複数の調査方式を比較分析することが重要 ― 実験調査例の紹介

ここから少し急いで、調査例を順番に見ていきたいと思います。ここ十年くらいやってきたことです。ここまでにお話ししたことから、複数の調査方式を比べた方がよいこと、1か所で調べるだけではだめではないか、ここらを念頭に、実験調査を開始した当初から進めてきたわけです。さいわい科研費や共同研究費をいただく、企業のサポートをいただくなどして進めることができました。一覧にあるように、初期はウェブ調査とオムニバス、面接、訪問留置、郵送と比べることを、また郵送とウェブを比べるなどいろいろあります。

今日はそれぞれの調査の抜粋を、特に関心のある項目をいくつか設定して述べますが、分かりやすくするために、調査Ⅰ、調査Ⅱ、調査Ⅲ…のように調査番号をつけてあります。要はここでの一連の実験調査は、いずれも複数のサイトで比較実験すること、参加協力機関との合意形成を行うこと、共通仕様を作ること、その仕様を守ってもらうことなど、「標準化して行った」ということです。特に、仕様に示したデータの提供が可能な調査機関とそうでない調査機関も分かります。つまり比べる必要があるわけです。どれくらいのスキルであるかを知る、少し意地悪な言い方ですけれども、そういうことです。それから、調査票も共通化し、（ほぼ）同時点調査を行うことは当然です。以上のことは全実験調査を通じて共通で、これから述べることにはすべて含まれています。

調査Ⅳは、これは少し特殊な調査です。ここでは、特別な条件を満たす世帯だけを取り出して、その対象者に日記形式でほぼ１週間調査を続けております（注：http://wordminer.comquest.co.jp/smr/pdf/j00001a.pdf から資料ダウンロード可能）。市場調査などでレアサンプルを抽出する時に使っている方法に類似します。調査Ⅴ（郵送調査とウェブ調査３パネルとの比較）、調査Ⅵ（ウェブ調査のみ）は、お声を掛けていただいたので一緒に参加させていただいた調査研究です。

調査協力機関・企業に仕様書を示し、合意書を交わすときに、取得情報を調査後に、ここまでは公開してよいという約束で行っています。ここにあるような協力いただいた調査機関・企業と組んで、調査Ⅰ、調査Ⅱ、調査Ⅲ、特に調査Ⅲは複数のオンライン・パネルを用い同時的に比較

実験調査の要約(1)

調査ｺｰﾄﾞ 調査方式	調査ﾌﾟﾛｼﾞｪｸﾄ名	ﾌﾟﾛｼﾞｪｸﾄ組織者	参加機関名	実施年次	実施回数	計画標本数 回収標本数
(調査Ⅰ) 訪問留置自記式 郵送調査 ウェブ調査	第1次実験調査	統計数理研究所	リクルート・リサーチ	1997年5月～11月	延べ12回	延べ計画標本数(184,638人) 延べ回収標本数(51,696人)
	第2次実験調査		NTTナビスペース 電通リサーチ ハイパーリサーチ リクルート・リサーチ	1999年3月～4月	延べ17回	
	第3次実験調査		電通リサーチ AIP	2000年4月～5月	延べ6回	
	第4次実験調査		電通リサーチ 博報堂－東京サーベイ・リサーチ・グループ 日本リサーチセンター	2002年3月～6月	延べ13回	
(調査Ⅱ) (A調査～D調査) A，D は郵送調査 B，C はウェブ調査	調査方式比較調査	統計数理研究所 博報堂	博報堂－東京サーベイ・リサーチ・グループ	2005年1月～3月	延べ4回	延べ計画標本数(9,322人) 延べ回収標本数(6,356人)
(調査Ⅲ) ウェブ調査 (e) ﾊﾟﾈﾙ間比較	Web実験調査 (E調査)	統計数理研究所 博報堂	アイブリッジ インテージ 東京サーベイ・リサーチ マクロミル ライフメディア 楽天リサーチ	2006年3月	同時的に全ｻｲﾄで1回	延べ計画標本数(26,897人) 延べ回収標本数(5,231人)

実験調査の要約(2)

調査コード 調査方式	調査ﾌﾟﾛｼﾞｪｸﾄ名	ﾌﾟﾛｼﾞｪｸﾄ組織者	参加機関名	実施年次	実施回数	計画標本数 回収標本数
(調査Ⅳ) ウェブ調査	親と子の生活行動と健康に関する調査 (親子調査)	(財)健康・体力づくり事業財団	博報堂－東京サーベイ・リサーチ NECビッグローブ	2007年3月	1回	割当法で割付： ﾊﾟﾈﾙA (380人)，ﾊﾟﾈﾙB (625人) を計画標本，合計1,005 (人) 回収標本： ﾊﾟﾈﾙA (322人)， ﾊﾟﾈﾙB (511人)； 合計833 (人)
(調査Ⅴ) ウェブ調査 郵送調査	基礎意識構造の統計的研究 (伝統的価値観)	林文教授 (東洋英和女学院大学)	博報堂－東京サーベイ・リサーチ ｱﾄﾞﾚｽｺﾐｭﾆｹｰｼｮﾝｽﾞ 日経リサーチ	2010年2月 (2/19～2/25)	1回	割当法で住民基本台帳から推計， 割付，延べ計画標本数(6837人)，回収標本数 (2271人)．
(調査Ⅵ) ウェブ調査	本と読書調査	矢口博之准教授 (東京電機大学) 博報堂－東京サーベイ・リサーチ	博報堂－東京サーベイ・リサーチ・グループ	2010年1月 (1/8～1/12)	1回	住民基本台帳から推計， 割付 計画標本数 (756人)， 回収標本数 (405人)．

しています。標準化しまた合意形成のもとに行ったこのような調査は、国内ではあまり例がないと思います。また調査Ⅳ、Ⅴ、Ⅵも、この一覧にあるような調査主体と調査機関で行いました。

皆さんの関心があることとして、まず属性つまり人口統計学的変数かと思います。例えば性年齢がパネルによってどうズレてしまうかということです。それから回収率の問題があります。回収率の観察で具体的に見えてくる登録者集団、計画標本と回収標本の間のギャップ、整合性の問題があります。これが分析の1、2です。分析3で、いろいろな質問のうちのごく少数、世論調査型の質問についての回答分布の傾向や、回答選択肢別の選択率のパネル間比較、つまり各質問の各回答選択肢の比率分布がパネルでどう異なるか、我々が日常的に行う比較分析です。それから完答率、これは聞き慣れない用語かもしれません。ウェブ調査では（回答する意思があるのに）完答できない人がいるわけですが、それを完答率や回答遷移パターンで観察するということです。つまり回答行動のトラッキングで取得したパラデータの分析で調査票の各ページをどう閲覧し、どう回答したかが分かるわけです。さらに、この中から“初頭効果（primary effect）”を確認する、回答所要時間と回答選択肢の関係を調べるなどあります。最後に我々が一番関心

のある重複登録の問題（登録率と重複登録率）、つまりどれくらいのサイトで登録者がどう重なっているかを推測してみた例です（調査Ⅲ）。これらを駆け足で眺めてみます。

こうした分析で、ボランティア・パネルの利用の限界がみえるわけで、またいわゆる“プロの回答者（professional respondents）”の存在が見えてくるわけです。たくさん登録している人は何十サイトにも登録していますし、インセンティブを目的に答えるわけですから、こういう行動をどう考えるかという問題があるわけです。

なお、調査方式が異なれば回答傾向も当然異なるわけですから、他の調査方式も使ってみるとどうなるかもありますが、ここらは少し省略します。

実は、似たような研究が米国にあることはクーパー氏などと時々連絡をやりとりして、聞いてはいるのですけれども、調査方式による違いがあるかどうかも調べています。日本ではあまり例がないのですが、欧米でもそうたくさんあるわけではない。主な実験調査例を2ページほどにまとめました。広告業界や研究者グループ、ESOMAR、AAPORでの発表などです。他にもありますが、ここではほぼ確率的パネルと非確率的パネルを比べるという例、ここで言うほぼ確率的というのは、アメリカは電話調査を基本に、これとオンライン調査の比較です。イギリスにおけるウェブ調査、電話調査、面接調査の比較などもあります。いろいろですがたくさんはなくて、特に興味あるARF（Advertising Research Foundation：米国広告調査協会）の実験調査結果報告は有料配付ですが、何とか一部を閲覧する機会を得ました。ここでは17のオンライン・パネルと電話調査、郵送調査と比べています。特にこのARFの特徴は、私が後でお見せするような間接的なやり方ではなくて、アドレスを直接確認、分析して、パネル間オーバーラップの推計を出していることです。つまり、個人が複数パネルに登録と自己申告した答えではなくて、物理的に重なっている様子を推計しているわけです（こうした実験例は国内にはないと思う）。

繰り返しますが、標準化を行ってなるべく同等の条件下での比較分析が必要なのですけれども、あまり例がないことと、とにかく“丸投げ調査をいくらやってもだめだ”、ということだけをここで述べておきます。

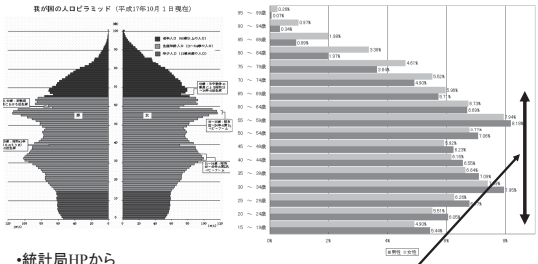
◎分析1：基本分析、探査 ― 登録者集団

では、まず分析1から始めます。ここで大事なことは、登録者はどう作ったのか、ということです。最近はさすがに各調査サイトを見ていただくと、「我が社のパネルはこういう属性構成になっています」などと書いてあります。男性が何％、女性が何％で年齢比がこうなっている、という人口統計学的情報はあるのですが、ご覧になるとすぐに分かるように、比べてみるとかなり違うことと、また当たり前ですが国勢調査に類似した構成比分布が再現されることはほとんどないわけです。大事なことは、「どういうパネルを作っているか、登録者の集め方は」ということを知る必要があるのですが、通常この情報はほとんど得られません。

念のために国勢調査情報を見ておくとこういう分布です（スライド）。これは2005年の例ですが、多くの調査は、マーケティングなどでも、年齢の上下を切って、ある層でしか聞いていない

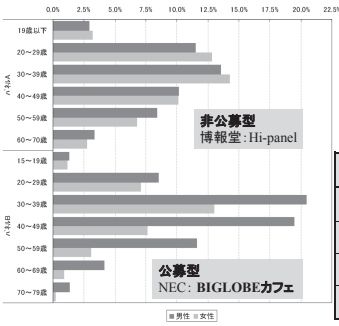
わけです。60 代後半あたりから上の層と、若年層は聞かない。世論調査であったらふつう 20 歳以上です。例えば、先ほどの調査Ⅳのものすごく登録者数の大きいパネルと、それが小さなパネルをみます。これは公募型と非公募型の典型例でもあるのですが、これを比べると、男女比が際立って違う。このギャップは当たり前で、ここで大事なことはむしろ“回収率”は登録者数が小さいパネルのほうがほとんど同等か高いということです。

国勢調査情報（2005年）の確認



- 統計局HPから
- 右側の図：年齢区分の上下をトリミングした
- よく知られているように“**双峰性**”の分布となる
- **ウェブ調査でこうなることはほとんどない!!!**

調査Ⅳ：親子調査の登録者集団：性年齢区分

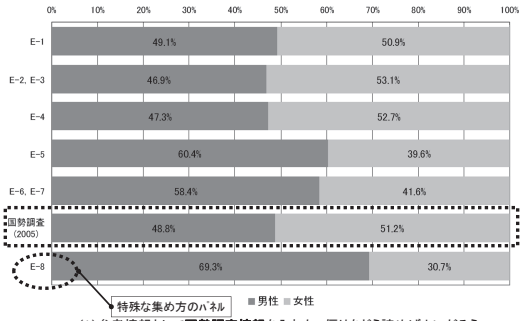


- 2パネルに差違がある.
- パネルAは関東・関西都市圏
- パネルBは全国登録者
- パネルA性比（約50:50）
- パネルB性比（約67:33）
- 回答分析に影響する.

2パネル構成	パネルA	パネルB
登録者数（人）	**15,803**	**1,604,776**
該当者数（人）	2,065	150,104
計画標本（人）	380	625
回収標本（人/%）	322 (84.7%)	511 (81.8%)

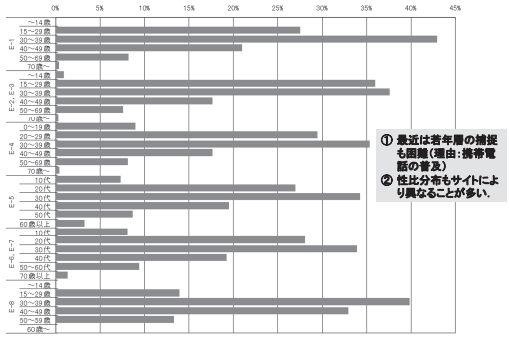
今度は調査Ⅲですが、これはほとんどがネット調査専業社サイト、6 社 8 パネルで比べた例です。ただしここで、E － 8 は特別な集め方をしています。これは主にガソリンスタンドを訪問する人をサンプリングしたものです。男性が多いから、当然男性登録者が多くなっています（注：これは実験調査当時の情報、現在は、別の集め方も行っているので属性に変化がある）。残りの 7 パネルは、一般に公募したパネルで、回収率がこれだけばらつきますが、これは（いままでお話ししたことから考えても）当たり前でこうなるということです。我々が日常的に行うようにデモグラフィック要因で分類して回答分布を観察すると、これらが違うことはほとんど自明なわけです。それから、E 調査（調査Ⅲ）ですが、これは少し最近の特徴が出ているのですが、よく見ると若年層が少ないことがわかる。つまり、携帯電話への移行で若年層の捕捉が難しくなっている表れです。もちろんウェブ調査を携帯電話で行うという例も多く出てきていますが、このケースは普通の PC 経由で調査したもので、結果登録者が偏っていることが分かるわけです。

調査Ⅲ：E調査（ウェブ調査）の登録者集団（性別）



（†）参考情報として**国勢調査情報**を入れた．偏りをどう読めばよいだろう．

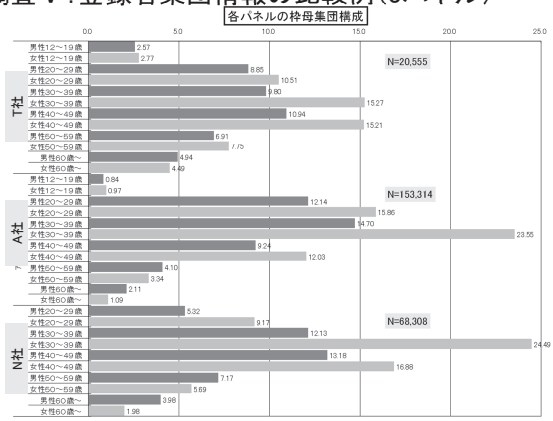
調査Ⅲ：E調査の登録者集団（年齢区分）



（†）双峰性の分布にはならないこと，若年層に偏ること，サイト間の違いがあること．

これは調査Ⅴの例です。3社のオンライン・パネルを使ってみると（2社が公募型、1社が非公募型）、明らかに2つは女性が多く、特にある年齢層が際立って多い。非公募型の1社は何となくバランスが取れているけれども、やはり女性が多い。これらが特徴です。こんなことをみても、よくあるようにただウェブ調査の結果から、男性が何％でこう答えた、ああ答えたという議論はナンセンスなはずなのですが、どうもそのあたりが私には理解できない。

調査Ⅴ：登録者集団情報の比較例（3パネル）



◎分析2：回収標本の特性、特徴 ― 望まれる“調査要約表”

回収標本の傾向を見ましょう。回収率が一つのキーです。ウェブ調査は一般に回収率が低いと言われていますが、やりようによってはそうでもないこと、つまり、“パネルの作り方で回収率が違う”ということが分かります（かなり常識的な結果）。それから、公募型と非公募型でも違うこともあります。複数パネルの比較は大事だということ、このように比べないと分からないということです。それからこれは、登録者集団と計画標本と回収標本の不整合が起こる例です。これは従

来調査でもみられたわけで、これが顕著に出てきたわけですが、通常はここに関心をもつ人はいないようです（また分析可能なデータを集めていない）。その他、調査不能などいろいろな問題がありますが、ここでは項目だけを上げておきます。現状は精密に調べているとは思えないということです。

ところで多くの場合、ウェブ調査の回収率がきちんと表記されていないことに気がついておられると思います。回収率を精密に出そうとすると、非常に面倒で、場合によっては今までの調査以上に厄介ということが分かります。これがそうです。私はこれを“調査要約表”と名付けております。JMRA などはこれの抜粋版を調査報告につけなさいとレコメンドしていますが、あれは実は我々がやった実験調査の結果の変形です。要するに、いろいろな測定項目、指標をここに列記してあります。皆さんのお手元の資料だと細かい文字が分かるかと思います。“有効回収数（率）”を決めるために、無回答、アクセス数、接触数、未着（数）の有無、完答数など、いろいろな項目を測定列記して、これらをまとめて評価しないと有効回収数（率）が決まらないのです。つまり、いろいろな指標を集めておくことが必要だということです。しかしおそらくは多くの場合は行われていないでしょう。我々の過去実験の際にはすべて、仕様書の中にこれら調査要約表を提出するようにとお願いしているわけですが、こうした情報をしっかり出せないサイトも出てきます。この項目は取れませんでした、これはここまでしか取れませんでしたという感じで情報が抜けているわけです。ということで、こうした調査要約情報が一種の調査品質表示として必要ではないか、というのが私の主張です。

調査要約表　ウェブ調査（調査Ⅲ：E調査）から

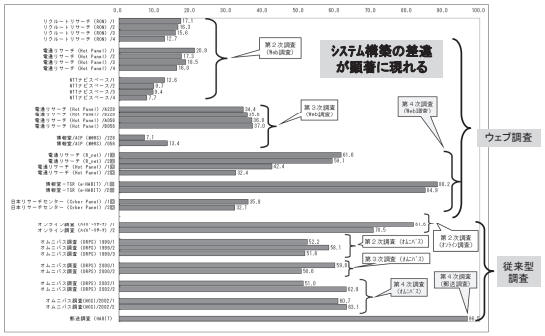
調査機関／調査条件	E-1		E-2		E-3		E-4	
調査期間	2006/3/23 ～ 2006/3/29		2006/3/24 10:00 ～ 2006/3/30 10:00		2006/3/24 10:00 ～2006/3/27 22:08		2006/3/23 10:00 ～ 2006/3/29 10:00	
調査テーマ	インターネット調査特性研究・実験調査		普段の生活やインターネットなどについて		普段の生活やインターネットなどについて		普段の生活やインターネットなどについて	
謝礼	250円		回答者全員に150円相当		回答者全員に150円相当		回答者全員に20p(200円相当)	
サンプリング方式	系統抽出		リソース内サンプリング方式 （層別無作為抽出）		リソース内サンプリング方式 （層別無作為抽出）		リソース内サンプリング方式 （層別無作為抽出）	
登録者数（サ）マスキングした								
計画標本	500（サ）		1,191		1,221		3,429	
依頼配信前除外	22577（サ）		0		0		1	
調査依頼配信数 以下の(%)の母数は調査依頼配信数	1,437		1,191		1,221		3,361	
有効回収回答数（%）	713	(49.6)	737	(61.9)	602	(49.3)	943	(28.1)
無効回収回答(PQ1 終了)（%）	7	(0.5)	4	(0.3)	6	(0.5)	4	(0.1)
無回答（%）	717	(49.9)	450	(37.8)	613	(50.2)	2,414	(71.8)
未着（%）	19	(1.3)		不明		不明	67	(2.0)
無接触（%）	680	(47.3)	426	(35.8)	582	(47.7)	2,302	(68.5)
アクセスのみ（%）	6	(0.4)	14	(1.2)	22	(1.8)	62	(1.8)
途中中断（%）	12	(0.8)	10	(0.8)	9	(0.7)	50	(1.5)
回答送受信異常（%）		不明		不明		不明		不明
アクセス数（延べ数） 以下の(%)の母数は総アクセス数（延べ数）	738		873		717		1,059	
アクセスのみ（延べ数）（%）	6	(0.8)	35	(4.0)	46	(6.4)	62	(5.9)
回答送信（延べ数）（%）	732	(99.2)		不明		不明	997	(94.1)
回答送受信異常（延べ数）（%）				不明		不明		不明
回収回答（延べ数）（%）	720	(97.6)	741	(84.9)	608	(84.8)	947	(89.4)
重複回収回答（%）	生起しない		生起しない		生起しない		生起しない	
非登録者回答			生起しない		生起しない		生起しない	
ID不明回答			生起しない		生起しない		生起しない	
調査依頼未達（延べ数）	19		不明		不明		67	
登録者数確定日時	2006/3/23		2006/3/23		2006/3/23		2006/3/23	
督促回数（リマインダー数）	0回		0回		0回		0回	

※パネルの比較・評価にはこうした要約表が必要（作れるかどうか）

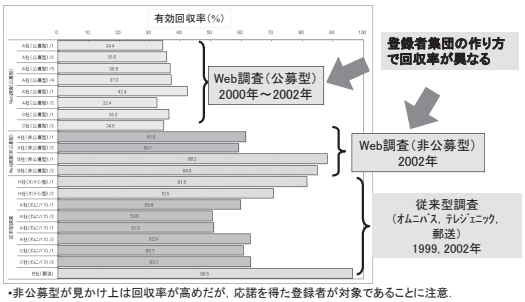
次は少し古い話しですのでここは簡単に述べます。例えば回収率をサイト別に比べると、従来型の調査方式でやったものとウェブ調査でかなり違っていることが分かると思います。次が大事で、公募型と非公募型に分けると、さらに明らかに差があり、非公募型の回収率が高めです。ここでは何となくそろっているように見えますが、たまたまです。ただし先ほどお話ししたように、

非公募型は登録者を集めた後のロイヤリティが高いわけで、しかし計画標本の応諾率が低いわけですから、これが全体の回収率を評価する際には注意が必要ということになります。

調査Ⅰ：回収率比較（第2次～第4次調査，実施年次で要約）



回収率比較（第2次～第4次調査，公募型・非公募型で要約）

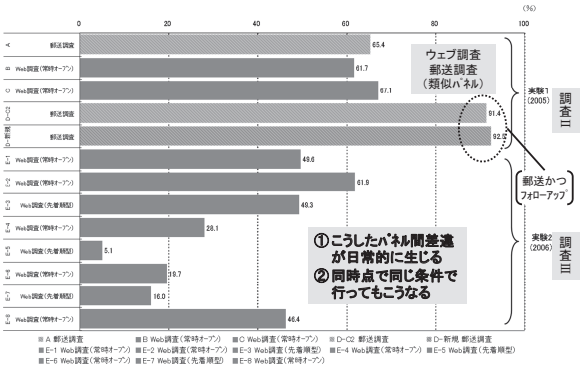


•非公募型が見かけ上は回収率が高めだが，応諾を得た登録者が対象であることに注意．
•公募型，非公募型とも，まずカバレッジ誤差が問題になる．

ここから少しゆっくりと話しします。調査Ⅱは、ウェブ調査と郵送調査の比較を行った結構大規模な実験でしたので、本当はここだけを取り上げてお話ししたい要素がたくさんあります。特に、調査の一部で登録パネル内の同一人に調査時点を変えてまったく同じ調査内容で調査方式を郵送とウェブで調査を行ないました。それは、いわゆる“加重調整（weighting adjustment）”の実験研究のためです（注：つまり調査方式を変えることで生じるであろう誤差、特に測定誤差や無回答誤差などの非標本誤差の検証と調査方式間の補正分析）。そういう背景もあるのですが、ここでは回収率だけみると、薄い緑（注：図の斜線網部）が郵送で、ここはフォローアップすれば回収率は上がることが見えます。この例では、ウェブ調査とあまり変わらないことも分かります。一方、これとは別の実験調査（調査Ⅲ）の結果である、同時に行った８つのパネルの回収率ではこれだけ差がつくということです（全体に回収率が低い）。これは特別な例ではなく日常的にこうです。後で同じサイトで行った少し違う例を見ます。

次にこれは調査Ⅴです。ここでは３つのサイトで同時にほぼ同じ調査票で行っております。ここで３パネルの登録者数がかなり違うことに注目してください。またここで計画標本数とは、対象とした実際にこれだけに配信しましょうという割当数です（注：各パネルともクォータ法で住民基本台帳に合わせて共通の割付を行った）。この結果の有効回収率は３パネルでかなり違うことが分かります。つまり、パネルの大きさが違うだけでなく回収率もかなり違う、それもパネルの大きさが小さい方の回収率が高いという現象がここでも見られます。よく言われるように、パネルの登録者数と回収率が高低の関係があるとは限らないし、しかも回収率にも差があるということが、数字として見えてきます。

調査Ⅱ，Ⅲ（郵送，ｳｪﾌﾞ；2005），回収率比較



調査Ⅴ：調査の概要（回収率ほかの比較）

機関名	T社	A社	N社
実施時期	2010年2月9日～2010年2月25日（同時点実施）		
勧誘方法	非公募型	公募型	公募型
登録者数	20,587	152,248	159,948
計画標本数	962	1,840	68,308
依頼発信前除外	0	0	不明
調査依頼発信数	962	1,840	4,049
有効回収回答（%）	**581 (60.4%)**	**623 (33.9%)**	**1,067 (26.3%)**
無回答（%）	**381 (39.6%)**	**1,217 (66.1%)**	**2,982 (73.6%)**

・パネルごとの回収率の差違が大きい
・登録者数が大きいから回収率は高いとはならない

参考までに、T 社の同一オンライン・パネルで、通年で行った一連の定例調査の 2 つの時系列的な回収率の比較結果の図をご覧いただきます（注：2005 年に実施の 75 回分のウェブ調査データ）。これは配布ができない情報ですので、ここだけで見ていただきます。2 つの大きな調査課題につき年間を通じて実施しているのですがこんなに回収率が異なり、上から下まで差がある（注：30 数%から 70 数%までに分布）。つまり、調査時点、調査内容とボリューム、対象者のその調査課題への関心度と、いろいろな要因が関係して、こうした差違が生じるようです。

◎回収率の特徴のまとめ

ここで回収率の特徴の大事なことをまとめてみます。以上に見るようにパネルによってさまざまで、これは当たり前のことなのです。それから、有効回収率の定義が（従来型調査以上に）繁雑で、測定誤差はどうかあるいは無回答誤差は、といろいろなことを評価しないといけないことが分かってはいるが、ほとんど考慮されていないということです。

ここに、ある学会が日経新聞と共同で行った調査を紹介した日経新聞の記事があります。ご覧になった方もあるかと思います。記事見出しは「財政支出拡大、経済学者の 51%『分野絞るべき』本紙・経済学会がアンケート」とあります（注：2010 年 11 月 4 日記事）。3450 人の学会会員に対して、540 人が回答、わずか 16%の回収率です。調査方式はウェブ調査というが、ウェブを使える人が何人いて、どういう人たちだったかということは一切書かれていない。どのように調査票を配信したかも書かれていない。もちろんより詳しい情報が裏にはあるかもしれないが、この小さい枠記事の情報では分からない。ただ、単純計算でこの回収 16%の内のさらに 51%というとわずか数%、8%足らずではないでしょうか。それで「51%がこう言った」などと記事に書いてよいのか、はなはだ疑問です。しかし、現実にはこうした報道が多いのです。

さらに、ウェブ調査では全体に回収率が高くなったという証拠もない。むしろオーバーサチュレーション、つまりパネルが過飽和状態になって、登録者があっちに登録こっちに登録という状

況にあること（後ろで例をみます）、全体としては劣化の傾向にあるようだということ、アメリカでもこれが顕著であると聞いております。

加えて、情報の開示が十分でないということで、ここで述べたことも、上の記事のように、多くは裏読みをしなければならないわけです。登録者が多いから、回収率や代表性が保証されるということも、どうも誤信らしいということです。

◎登録者集団、計画標本、回収標本の“不整合”

登録者集団と、サンプリングした計画標本と回収標本、それぞれの間の不整合性がいつも出るという例をみます。これは昔からあったわけで、別にウェブ調査特有の現象ではありませんが、これに注目する人はあまりいないようです。また、情報があっても出せないようです。E調査では、参加協力の企業にこれらの情報提供をお願いしましたが、何社かはそれを出すのは勘弁してくれと、提供してもらえませんでした。提供いただいた中の一つを見ると、いくつかの属性項目（年齢や年間収入）で不整合がみられます。年齢では、数％くらいです。つまりパネルの登録時情報と、サンプリング時の情報、また回収回答にある情報の属性が一致しない、食い違いがあるということです。これは、回収標本に関してログをたどらないと分からないこともあります。きちんと答えてくれたかどうかです。所得分布などは、マーケティング等での重要指標だから初期登録時に尋ねるわけです。登録時に得た情報がこの表の表側です。年齢についても登録時に聞いています。こちらの表頭が回収時に聞いた年齢区分結果で、このくらいずれております。これはウェブ調査だからということではなくて、ウェブ調査でもやはりそうなのだという例です。

こうしたことからパネル疲労も予想されるのですが、パネルがどういうローテーションでどう使われていて…、といった一番肝心なことがいつも分からないのです（公開されていない）。またこうした事象は日常的に起こると思って使ったほうがよいのです。

こういうベリフィケーションを多数設けると、ウェブは安くできるということにならないのではないか、となります。我々の実験調査でも、協力機関にこうした情報の提供をお願いしたときに、日常的に行っている場合より調査費用が余計にかかったとの報告もありました。また、先ほど廃止されてしまったと言ったA社の例のように、維持費がかかるからできないとなって、脱落してしまうケースも出てきます。

こうしたパネルへの登録時、計画標本作成時、回収時と回収履歴の情報収集の開示がなされることはないのではないか、開示されていたら、ぜひ見たいのですが…。

◎分析3：回答分布に見られる主な特徴

分析3として、世論調査型の典型的な質問についての回答傾向等をみます。一つは支持政党、もう一つは階層意識、よくある典型的な例で、どうなっているのかを眺めます。もちろん細かくは質問文のワーディングが違う、調査方式が違う、登録者・調査対象者のパネルも違う、といろいろな要素があるのだけれども、それでもなお少し比べて見ようということです。それからもう一

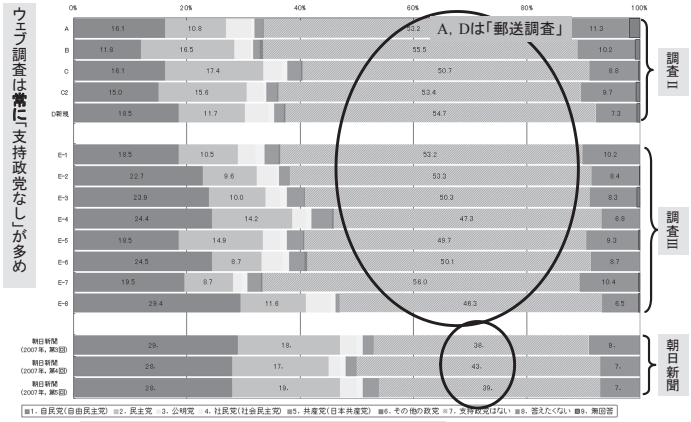
つは、パネル間の回答比率が、つまり、異なるパネルで得た同じ質問についての回答選択比率がどう変動するのか、どれくらい揺らぐのかを比べてみるということです。

支持政党、これは数年前の分析比較例です。上が調査Ⅱの結果、同じ会社の同じサイトの複数のパネルを使ったものです。A と D は郵送調査です。図の下の方の調査Ⅲは同時に行った複数パネルの比較です。こういう風に、微妙に揺らぐわけで、これをどう読むかです。揺らぐということを知っても仕方がないかもしれないが、こうなっていますという例です。（調査Ⅱにおける）郵送調査とウェブ調査の違いは、この場合はあまり顕著でないけれども、（調査Ⅲの）ウェブ調査間の違いが大きいようにみえます（注：つまり、パネル間差違が顕著ということ）。

登録者集団の属性が違うわけだから当然予想されることです。では年齢別にブレークダウンして補正をするなどしてそろえて比べればよいのではないかと言う意見、質問があると思いますから、答えを先に言っておくと、たとえば単純な事後層化法（post stratification）で補正を行ってもあまり合いません。つまり、事前の人口統計学的変数の分布の違いが大きいのではないか、ということです（注：つまり補正を行う意味があまりない）。

この例だと、2006 ～ 2007 年に行った E 調査に調査時点が近い朝日新聞の調査（2007 年）があったのでこの中の支持政党を整理して書いてみたのがこの図です。もちろん朝日の場合は（朝日 RDD 方式によるとされる）電話調査です。比べるとすぐ分かるように明らかにウェブ調査がすべて「支持政党なし」が多い。これは、過去のほとんどの実験調査でこの質問をずっと使ってきていますが、過去の新聞紙面に出てくるどの数値よりもさらに「支持政党なし」が多くなる傾向にあります。つまりウェブ調査では（系統的に）偏っているということで、こうした特徴・傾向は分かっています。

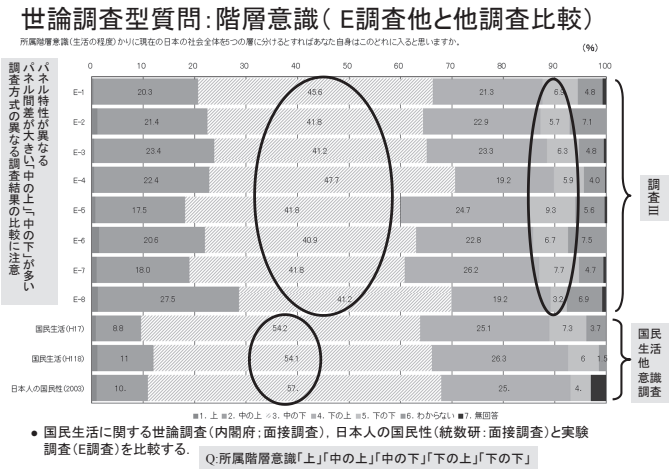
支持政党の比較（朝日新聞は調査方式が電話調査）



こうした現象をただちに調査方式間の差異と言ってよいのかは微妙です。しかし、“黙従傾向”や“社会的望ましさ”などはどうなのか、例えば調査方式が電話調査だと調査員の質問にじかに

は答えにくい、偏りが入ると言われている。一方、ウェブ調査は自記式であるからそういう影響を受けにくいのか、あるいは何なのかということはある、これらはさらに精査しなければ分からないということです。

ここで所属階層意識について、E 調査（8 パネル）の結果と他調査、ここでは「国民生活に関する世論調査（内閣府）」と「日本人の国民性調査（統計数理研究所）」にある所属階層意識の同じような質問を引用しこれを比べてみます（図「世論調査型質問：階層意識」）。ここではウェブ調査のほうは似ているように見えるけれども、下の他の調査（3 つの結果）では不満な層、階層があまり上でないと考える層がインターネットでは増えているようにみえる。実は似たような傾向は過去に行った同じ質問結果から見られます。少なくとも、ウェブ調査は他の調査とは少し違う傾向があるということです。

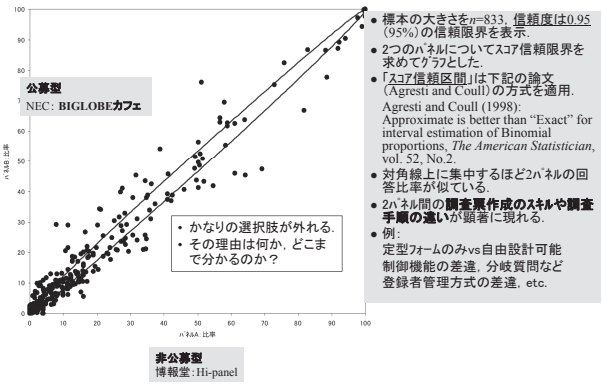


◎回答分布はパネル間でどれくらい違うのか

さらに関心のあることは、回答傾向にパネル間でどれくらいギャップが出るのかということで、ここからしばらくそのお話しをします。パネル別に用いた質問全部の回答比率を並べて、パネル間を比較する散布図を描きます。我々がよくやることです。ここでは横軸がパネル A（非公募型）で、縦軸がこちらのパネル B（公募型）です。ここで各点は各質問の回答選択比率です。もし両者の答え方が似ていれば、これは対角線に近く分布します。また、ここでは信頼度（95%）の信頼区間の曲線を入れてあります。この中に入ってほしいのだけれども、このようにばらついています。この例ではこうしたばらつく理由がおよそ分かっています。こちらのパネル B（公募型）の方が、我々の期待するようには調査票を作れなかった、つまり日常的に用いている定型フォーマットでしか作れず、パイピングや分岐処理、画面遷移、改ページ処理などやや複雑な設計があり、2 パネル間でかなりスキルが違っていたということです（注：つまり顕著な測定誤差の違いがある）。それとパネル本来の特性（登録者の集め方や属性など）の違いももちろんある。分析 1 で

性年齢区分の違い、年齢の違いを見たように（パネルＢは）、男性が多いわけです（注：しかしこの調査は、ある年齢層の女性の世帯を対象とした調査である）。実は、特に図のこのあたりの外れている部分はほとんどがインターネット関連の質問文と選択肢なのです。外れた質問項目は分かる、もちろん、ここらは調べています。ここでは同じ内容の調査を行っても、これだけ違いが出るということを見ていただいたわけです。

調査Ⅳ：2パネル間の質問選択肢回答比率の比較



これは調査Ⅱについて、上と同じ回答選択比率を比べた例です。調査方式は郵送とウェブを用いてます。2つの調査方式を同一サイト内の性格の違ういくつかのパネルも用いて比較するということです。今日は詳しく述べられないのですけれども、（前述のように）“同じ調査対象者”に対して郵送とウェブで比べています。つまり、のちの調査方式間の加重調整を検証するための実験調査です（注：事後層化や傾向スコアによる補正法の研究）。加重調整についてはいろいろと言われていますが、これはなかなか難しいです。世の中に興味ある結果が出てきていますけれども、そこはごく一部の話しで、本当かなというものも多々あります。

調査Ⅱ：郵送とウェブ（同一サイト内の比較）

- 郵送調査とウェブ調査の比較（一部，同一対象者）
- インターネット・ユーザと非ユーザの比較

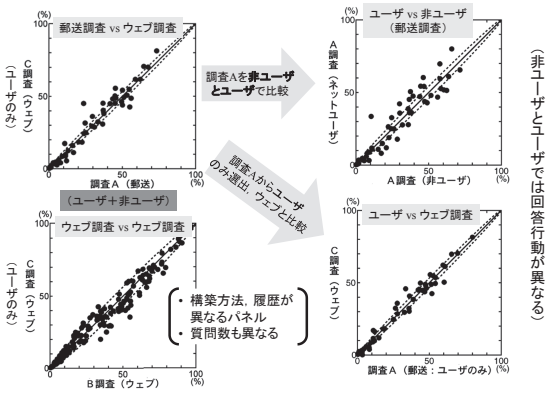
調査要約（一部を抜粋）	A調査	B調査	C調査
対象者（登録者集団と抽出法）	2004年度HABIT ほぼ全数調査	2004年度HABITのうち2005年度Hi-panel参加意向者（注1）	2004年度Ｈｉ－panelより層化無作為抽出
調査方式	**郵送調査**	**ウェブ調査**	**ウェブ調査**
調査時期と回収期間	2005年１月 約3週間	2005年３月 １週間	2005年３月 １週間
依頼数（注2）	4,630	857	2,499
回収数	3,026	529（507）（注4）	1,678
回収率（注3）	65.4%	61.7%	67.1%
備考（実査時の条件など）	定例調査最終回に追加調査として同時的に実施	Hi-panelへの移行前のプレ調査として実施	通常のウェブ調査のひとつとして実施

注1）2005年度パネルHi-panel参加意向者の中には、2004年度パネルHABITに参加した本人ではなく、一部にその家族等が含まれる。A調査とB調査の反復結果を検討する場合には、B調査の集計を本人のみに限定する。
注2）依頼数は郵送調査では調査票発送数、ウェブ調査では依頼メイルの発信数。
注3）回収率の定義は、ここでは回収数／依頼数による。例えばD調査は事前の応諾を経た上での依頼であるので、協力意向を照会した人数をベースとした回収率はこれよりも低い。
注4）B調査は、調査実施期限内の回収サンプル数が507であったが、ここは期間外回収を含めた529サンプルを表記。

ここではその一部を眺めます。（この表には）3つの調査、A調査、B調査、C調査がありますが、このうちのA調査とC調査を取り上げます。ここでA調査（郵送調査）は「インターネット・ユーザと非ユーザが混在」したパネルです。またC調査（ウェブ調査）は「インターネット・ユーザのみ」からなるパネルです。A調査では調査時にインターネット・ユーザと非ユーザを確認してますから、2つに分けることができます。これを分ける前と分けた後でそれぞれ回答選択比率を比較してみます。つまり「（A調査）対（C調査）」と「（A調査のインターネット・ユーザのみ）対（A調査の非ユーザ）」「（A調査のユーザのみ）対（C調査、つまりユーザ）」、これらの回答選択比率を比べた散布図です。ここで明らかなことは、ユーザと非ユーザ間の差異がもっとも大きく、ユーザ同士のパネルBでは、同じパネルからの比較ではよく似ており、パネルの構成が異なる場合はその類似性が崩れるが、しかし「ユーザと非ユーザの違い」ほどは大きくはない、となります。実は、何回かこれに似た実験調査を行っております。さらにもう一つ踏み込んだ調査分析も行っております。インターネット・ユーザをインターネット利用頻度の高いヘビー・ユーザと、そうでないユーザかを質問で問うようにしてますが、この利用頻度情報でブレークダウンすると、ヘビー・ユーザとそうでもないユーザでは答えが少し違ってくることも分かります（注：単にインターネット・ユーザだけで括れないということ）。

さらに、同じウェブ調査の結果、「（B調査）対（C調査）」（つまりどちらもインターネット・ユーザだがパネル構成、作り方が少し違う）、同じウェブ調査でもパネルの作り方が少し違うと、この散布図のように（ただし質問数が異なるのでそのことに注意して）、このように信頼区間などから見ると、結果がすこし違うということも分かっている。ともあれ、こういうことを繰り返し調べる必要があると考えております。

調査結果間の比較（抜粋）



今度は今年実施した3つの調査の例です。行動計量学会で一部発表させていただきました。まずここで調査企業3社のパネルからえた各質問回答選択率を用いて単純にT社、N社、A社のそれぞれ2つ

のパネル間の絶対距離を計算して比べてみます（注:$d_{ij}=\sum_k \left|p_{ik}-p_{jk}\right|(i,j=2$つのパネル;$k=$質問の選択肢)）。

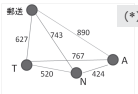
距離の模式図を描くと、T 社（非公募型）は郵送に近い、ところが（公募型の）T 社と A 社はかなり離れている、特に A 社は郵送からかなり離れているといったようなことが分かってきます。つまり、ウェブ調査の 3 パネル間にも差があって、しかもそれぞれの郵送調査との距離もパネルによって異なります。事前に予想されていたことがこうして数字で出てくるわけです。もちろん質問内容にもよりますが、ここでは「すべて同じ質問」を使い、ウェブ調査については「同時点で実施」、郵送調査では同じ質問を使っておりますが調査票のフォーマットはウェブ調査のそれとはやや異なり、また、ウェブ調査とは調査時点が若干ずれています。ともあれ、かなり同じような条件下で（ほぼ標準化して）行った調査ですがこういう傾向が調査方式の差違を示し、また同じウェブ調査でも差違があることが見えるわけです。

調査Ⅴ：質問選択肢回答比率の比較（1）

- 回答比率のパネル間距離を求めてみる（下の表と図）.
- 質問項目選択肢について，選択肢のある2調査（パネル）との回答選択率の差を合計する.
- 明らかに“パネル間，調査方式間の差異”がある.
- 郵送調査からみると，ウェブ調査Cが最も離れ，ウェブ調査Aが近い.
- ウェブ調査間にも差異がある（NとAが近く，TとAが遠い）.

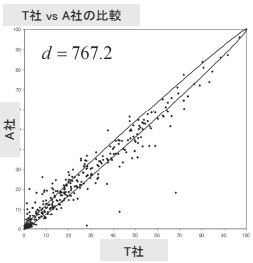
パネル／調査方式	ウェブ調査（T社）	ウェブ調査（N社）	ウェブ調査（A社）
ウェブ調査（N社）	519.8		
ウェブ調査（A社）	767.2	424.2	
郵送調査	**627.1**	**742.9**	**889.9**

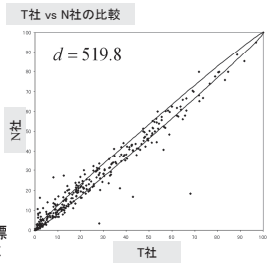
$$d_{ij}=\sum_k \left|p_{ik}-p_{jk}\right| \quad \left(\begin{array}{l} i,j=2\text{つのパネル} \\ k=\text{質問の選択肢} \end{array}\right)$$



（*）林・大隅・吉野（2010）から

調査Ⅴ：質問選択肢回答比率の比較（2）

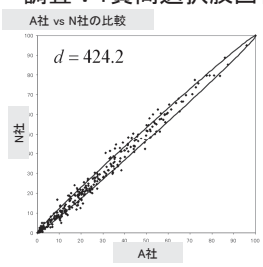


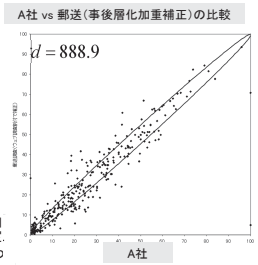


- 計算条件は，信頼度95%，2組の標本の大きさの調和平均を標本の大きさとし，スコア信頼限界を指定.

あと（スライドに）数ページ描いてある図は、ウェブ調査（3 パネル）と郵送の、合わせて 6 通りの組合せに対する散布図を描いたわけですが、これらを見ていただければ、いま述べた相互の類似関係がより分かりやすいでしょう。

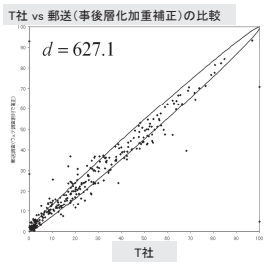
調査Ⅴ：質問選択肢回答比率の比較（3）

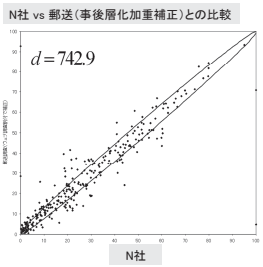




- ここも信頼度95%，2組の標本の大きさの調和平均を標本の大きさと，スコア信頼限界，を指定.
- 郵送をウェブ調査の人口統計学的構成に合わせるために**事後層化加重補正**を行った.

調査Ⅴ：質問選択肢回答比率の比較（4）





- **ウェブ調査間にも差異あり**
- **郵送調査へ近さにも差異あり**
- **予想される常識的な結果**

（説明があとになりましたが）ここでは若干の加重調整をしています。ここで郵送調査は属性構成がウェブ調査のパネルとは異なりますので、形式的に事後層化補正を行い、ウェブ調査の属性に合わせるようにしました（注：この郵送調査の回収率はかなり低かったので事後層化による補正が有効かはやや疑問があるが形式的に適用した）。とにかくこのようなことも調べる必要があるのではないかということです。外れている項目がどういう項目かは、これもいろいろと調べています。

◎分析4：回答行動の分析 ―「パラデータ」取得・分析がウェブ調査の特徴

ここで次に回答者の回答行動を少し見てみます。いわゆる“プロの回答者”や回答行動等に関係することです。このトークのはじめに触れた、2つの区分、調査方法論の側からみるか、調査システムの側からみるかでいうと、調査システム側にに少し足を踏み入れた話しです。

まず“完答率（complete response rate）”があります。これの問題はウェブ調査の場合は完答できるとは限らないということです。必ず完答できない人が出る。もちろん今までの調査でもこれはあるが、違う形で出てくるということです。次に“回答所要時間”、これがほぼ正確に取得できるのがウェブ調査の特徴なのですが、これが早すぎも遅すぎもどちらも問題があって、まあいろいろとあります。例えば、同じ質問をしているのに、数分で答えてくる人がいるかと思うと、数日あとにならないと答えてこない人もいる。しかも、途中でやめられますから、途中まで答えて待機して、また答えてと…、回答所要時間が何時間もかかる人もいる、バラバラなのです。ただ、ある特徴があります。回答所要時間は非常に歪んだ単峰性分布となり、これは共通しています。回答所要時間と質問文の関係が当然予想されるわけですから、これも注目点です（注：質問の回答難易度や調査票設計の複雑さなどに依存）。それから、チェック・ボックス型の複数回答の質問文になると、早く答える人ははじめの方に答えやすいのではないかということが言われている、つまり選択肢の並びの上の方だけ答えやすいのではないかということです。これを少しチェックしてみた実験があります。いわゆる“初頭効果”を調べた結果です。ここらを少し見てみましょう。

ここで完答率から観察します。ウェブ調査の最大の特長は、私も好んで利用するのですが、回答者の行動を電子的にトラッキングし、いわゆる“パラデータ（paradata）”が取れることです。このことを回答者は知らないわけです。調査倫理を厳密に考えた時にこれを行ってよいのかどうかは、少し議論の余地があって、いま注意はしております。ただ今まで行った実験調査ではこれを取得させてもらってます。そこで完答のことですが、質問量や調査票デザインなど、いろいろな要素の影響があります。特に分岐がたくさんあるとか、調査画面のデザインが複雑になると、1回で完答できない人も出てきます。それが具体的にどうなっているかを少し追いかけてみます。

これはもちろん、回答制御方式にも関係し、そのレベルが関わります。結論を言うと、複雑な調査票設計はやめたほうがよい、なるべくシンプルに作るということです。まだパラデータの分析方法の研究開発が遅れていて、私も、また米国のクーパー氏もこの研究を進めているところです。また、これをはじめるとお金もかかるわけです。ウェブ調査は安いということは、こうしたことには目をつむっているからです。実際に私がお願いしてきた調査機関では、「（やむなく）やり

ますけれど、これでは赤字ですね」と言われています。しかしこれは非標本誤差の観察に大きく関係しているのです。

では完答率を見てみます。回答意思があるのに回収不能になってしまう、これは完答、中断、途中で脱落…とさまざまで、回答がスムーズにできるとは限りません。つまり回答者のコンピュータ・リテラシーや使用コンピュータの問題、使っているブラウザの問題、…とあるわけです。回答は１回で済んだのか、あちこち調査票内のページを迷いながらなのか。このことは、郵送調査などで経験するように、どこか書いてまた戻って消して書き直すなどというように、従来もみられた現象です。同じことがウェブでも当然起こるわけですが、それを電子的に追跡するということです。例えば、質問文のワーディングなどが悪かったら、迷ってあちらこちら、こっちを選んだら、いやこちら、…と迷って答える。そこらをできるだけ追いかけてみるということです。アクセスだけして、少し見て止めてしまう人もある。完全な回答拒否ではないけれども、積極的には回答に参加しない（注：項目無回答となる）。調査内容への関心はあってアクセスはしてくれたのだなということは分かるけれども、そこから先は回答してくれない。これらはみな、PC の環境やスキル、リテラシー、回答行動の差違から起こる測定誤差、無回答誤差などに関係するということです。

例：ある割合で完答できない人がいる（調査Ⅲ）

		N	回答完了者のページごとの延べ送信数カウント 00	01	02	03	04	05	06	前に戻ることなく最後まで回答	
E-1	常時オープン型	713	732	722	718	717	716	713	718	672	
E-2	常時オープン型	737	737	737	737	737	737	737	737	737	←延べカウント不能
E-3	先着順型	602	602	602	602	602	602	602	602	602	←延べカウント不能
E-4	常時オープン型	943	951	953	947	946	950	950	943	912	
E-5	先着順型	612	616	617	614	616	619	613	612	595	
E-6	常時オープン型	597	600	598	599	597	599	599	597	587	
E-7	先着順型	595	596	596	597	597	598	597	595	585	
E-8	常時オープン型	432	432	430	430	434	430	435	432	415	

このサイトは記録出来なかった

N 【サンプル数に対する延べ送信ページ数の増加率%】

		N	00	01	02	03	04	05	06		
E-1	常時オープン型	713	2.7	1.3	0.7	0.6	0.4	−	0.7	94.2	
E-2	常時オープン型	737	−	−	−	−	−	−	−	100.0	←延べカウント不能
E-3	先着順型	602	−	−	−	−	−	−	−	100.0	←延べカウント不能
E-4	常時オープン型	943	0.8	1.1	0.4	0.3	0.7	0.7	−	96.7	
E-5	先着順型	612	0.7	0.8	0.3	0.7	1.1	0.2	−	97.2	
E-6	常時オープン型	597	0.5	0.2	0.3	−	0.3	0.3	−	98.3	
E-7	先着順型	595	0.2	0.2	0.3	0.3	0.5	0.3	−	98.3	
E-8	常時オープン型	432	−	1.4	0.9	0.5	1.4	0.7	−	96.1	

“完答率”がこれ
※過去例では70%〜80%程度
※この例は完答率が高い

*(参考)各ページの内容
00 開始ページ／職業チェック質問
01 A.「普段の生活での気持ち」について
02 B.「政治」について
03 C.「科学技術」について
04 D.インターネットの利用状況等について
05 E.「情報源」について
06 F.ご自身について

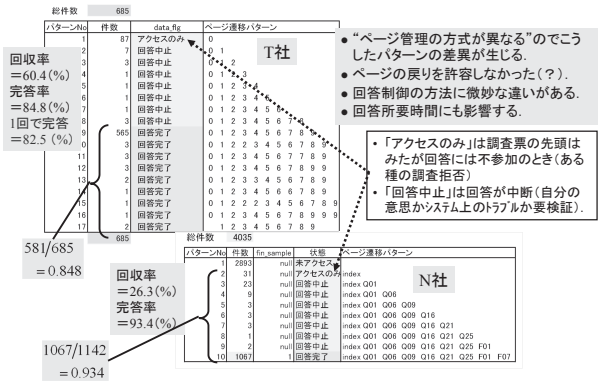
*E-1は職業チェック質問が独立したページ(01ページ)となっているが、他社とそろえてカウントした

過去の例でいうと、完答率100%はまずない。7割から8割、高いとこのくらい（90数%）です。当然これが回収率に関係し影響します。先ほど調査評価表を紹介しましたが、そこでも見たようにさまざまな指標が必要になるということです。あのような表を作っておかないと分からないのです。繰り返しになりますが、いわゆる有効回収率にも関係しています。このＥ調査の場合、データが取れなかったサイトもあります。取れたサイトだけみると、完答率は90数%です。調査Ⅲですとほぼ完答ですが、ごくわずか数%完答できない人が出ています。

次は少し完答率がやや悪い例です。これは調査Ⅴの３パネルのうちの２社（Ｔ社、Ｎ社）です。調査票を見ると、単純であまり複雑な部品を使っておりません。ごく簡単にラジオ・ボタ

ンやチェック・ボックスで済ませてます。これがページ遷移の履歴です。アクセスだけして、1ページまで見てやめてしまった人、途中までで回答が止まってしまった人、完答した人といろいろあります。下のN社の例はきれいに見えますが、これは正確にデータが取れなかったということです。ここだけ訪問したところは取れたのできれいに並んでいるだけで、上のT社がこれと違う理由は、質問文ページを往き来した履歴がしっかり取れていることです。例えば、1→2→2→2→2→3→4など、同じページを繰り返し見るなど、完答した人でもいろいろ遷移しているわけです。回収率をみるとT社が60.4%、こちらのN社の回収率は26.3%です。一方、T社の完答率が84.8%で、N社は93.4%でこちらの完答率が高い。これは見かけ上の完答率で、その内訳を見ればまた違ってくるかも知れない。T社の1回で完答できた人が82.5%だけれども、N社は分からない、正確に測定できなかったからです。

回答遷移パターンの例（調査Ⅴ：伝統的価値観）



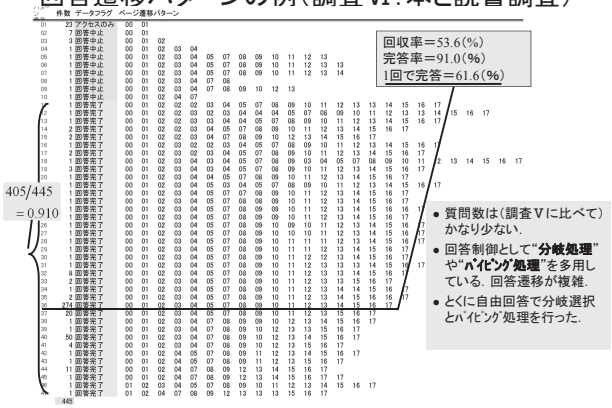
ここで調査Ⅵの例をみます。T社の上とほぼ同じパネルを使った調査です。いま話題の電子書籍に関連した調査を行ったものです。これの調査票はやや複雑で、分岐処理やパイピングを多用しています。例えばある質問で「あなたはどんな本を何冊読みましたか」と聞くと、読んだとして選んだ本のジャンルに合った部分だけを次の画面に表示し冊数の記入を促す、といった処理を行う。これの実物はサーバに繋がっていないと見られませんから、イメージのみでご覧いただきますが、例えばここに本が、どの本を読みましたかということでクリックすると、仮にここで3つクリックしたとすると、次のページ進んで、その選んだ3箇所の本のジャンルだけが表示されるわけです。この"パイピング処理（piping あるいは fill）"は、例えば自由回答の例をみると、「あなたは今後書籍の進歩に技術の進歩に伴い、紙の書籍の時代が終わって、電子書籍になると思いますか」という質問にある選択肢、「非常にそう思う」から「そう思う」「いずれ共存すると思う」「まったくそう思わない」「わからない」としておいて、これのどれかを選ぶと選んだワーディングをそのまま次の画面に移して（埋めて）、例えば「あなたが"非常にそう思う"と答えた理由

をお聞かせください（自由回答）」とすることができます。こうしたパイピング処理ができます。

とにかくさまざまなことができるのだけれども、あれこれ多用すればそれだけ調査票は複雑になります。あちこちで分岐やパイピングを使うと、ページを訪問するパスが複雑になって、このように一度で完答できた人は445人中274人、つまり61.6%が一度で完答できたけれども、迷ってしまって1割くらいが脱落しています。また回収率は53.6%となってます。この例は一つ前の2パネルの回答遷移パターンに比べ、非常に複雑になってます。

こういう測定が必要なのかというと、私は必要だと思っております。なぜなら、調査票設計の出来映え、測定用具としての性能評価になるからです。また、途中まで答えた人たちが完答者とくらべてどういう答えをしてくれたのかも事後の分析で分かるわけです。

回答遷移パターンの例（調査Ⅵ：本と読書調査）



◎回答行動と回答所要時間の関係 ― 初頭効果の観察ほか

少し先を急ぎます。調査Ⅲで回答してくれた回答所要時間と回答の仕方が何か関係があるのかをみます。実は、日本でも若干そういう研究を行っている人が出ています。つまり質問文、質問の状況、ワーディングを含めて、選択肢の並べ方、配列のレイアウト、デザインですね、それから使うツール、…でどうなるかということです。私もクーパー氏と一緒に行った実験に、プルダウン・メニューで10個の選択肢を逆順に入れた場合、ランダムに入れた場合などとすると、日本とアメリカで結果が違ってしまったりもするのです。詳しくは分からないのですが、国民性の違いもあるのかなとか、もう一度やってみようとか言っていますけれど、そのように違ってくる。こういう研究、これは参考で書いておきましたけれども無数にあって、先ほど紹介したクーパー氏の本（注1、④）を見ると、600編くらいの文献資料が引用されていますから、ぜひご覧になってみてください。いろいろなことが行われていますが、多くの場合あまり複雑なことはやらないでなるべく単純がよいだろうというのが、我々が得ている現時点の結論です。

それに多くの場合は、従来の郵送調査などでも見られたことが形を変えて現れている可能性

もある。ここらの吟味はまだ十分でないと考えております。

少しこれに関係したことを行ってみたのでこれを述べます。意図的にたくさんの選択肢からなるチェック・ボックスを使う複数回答質問を用意します。こうした複数回答質問を８つほど挙げてあります。選択肢数はかなり多いものからそうでもないものまで用意してあります。これらの各質問と（段階尺度化した）回答所要時間とのクロス表を作ります。かりに初頭効果があるとすると、つまりはじめの方が選びやすければ、回答所要時間との関係で有意になるはずです。統計的な検定をどこまで信用するかということがありますが、またこれらのクロス表はスパースなセルも多いので、その影響も無視はできないですが、まあ全体に有意になっていますし、念のために対応分析も行いチェックしていますが、有意になっているようです。回答所要時間が早いと選択肢の前の方を選びやすいのではないかなということです。このような現象があるので、最近はランダム化を行います（注：この例ではまだこうした方法を適用することまで考えていなかった）。つまりシャッフルして選択肢の並び順を変えて回答者に配信します。しかしこの方式が本当に初頭効果の解除になっているのかというのは、さらに検証の必要があると思います。実は先ほどの読書調査では、これに類したこと、つまりある選択肢の並び順をまったく逆順にした場合と初めの並び順の２通りを作り、これをランダムに配信するとしましたが、結果は効果があったのかどうかちょっと分からない微妙な結果でした（注：回収数が 455 人と少ないこともある）。これは今後の研究の余地があります。しかし最近はほとんどこうしているはずです。例えば、ラジオ・ボタンも配列をランダムに入れ替えるとか、プルダウン・メニュー内の選択肢をランダムに入れ替えて表示するなどです。いま見ていただいた例はランダム・シャッフルしていない場合で、ここでは回答所要時間が短いと、選択する反応数がつまりチェックした数が少ないのかもしれない、少しは関係がありそうだということです。

8質問（MA）の回答所要時間とのクロス分析

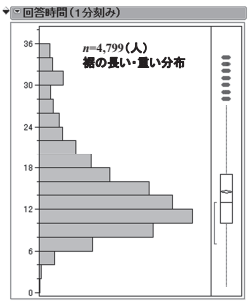
- DQ3：利用している検索エンジン（22個）
- DQ6：インターネットを利用した調査に「協力してもよい」と思う条件（11個）
- DQ7：登録サイトは（41サイト）
- DQ8：情報感度（11個）
- EQ3：耐久財の世帯保有状況（10個）
- EQ4：缶・ペットボトル飲料の情報源（36個）
- EQ5：薄型大画面テレビの情報源（38個）
- EQ6：携帯電話の情報源（35個）

- いずれも選択肢が多いことに注意
- 最近はランダム化（シャッフル）を行う
- これで“初頭効果”が解除できるかは疑問

質問	有意確率		自由度
	尤度比カイ二乗統計量	ピアソンカイ二乗統計量	
DQ3：利用している検索エンジン（22個）	<0.0001	<0.0001	72
DQ6：インターネットを利用した調査に「協力してもよい」と思う条件(11個)	<0.0001	<0.0001	44
DQ7：登録サイトは	0.0185	0.0200	136
DQ8：情報感度(11個)	<0.0001	<0.0001	44
EQ3：耐久財の世帯保有状況（10個）	<0.0001	<0.0001	40
EQ4：缶・ペットボトル飲料の情報源(36個)	<0.0001	<0.0001	140
EQ5：薄型大画面テレビの情報源（38個）	<0.0001	<0.0001	140
EQ6：携帯電話の情報源（35個）	<0.0001	<0.0001	132

(†)すべてで有意であることに注意．対応分析の結果も勘案すると，“回答所要時間が早いと選択する反応数が少ない”という傾向があるらしい．

回答所要時間の分布の例（調査Ⅲ）



回答所要時間（1分単位）のﾋｽﾄｸﾞﾗﾑとﾎﾞｯｸｽ・ﾌﾟﾛｯﾄ図

分位点

100.0%	最大値	34.000
99.5%		34.000
97.5%		32.000
90.0%		24.000
75.0%	4分位点	17.000
50.0%	中央値	13.000
25.0%	4分位点	10.000
10.0%		8.000
2.5%		6.000
0.5%		4.000
0.0%	最小値	1.000

回答所要時間が非常に短いからかなり長いまでさまざまとなるのが特徴

モーメント

平均	14.487185
標準偏差	6.6096573
平均の標準誤差	0.0954121
平均の上側95%信頼限界	14.674236
平均の下側95%信頼限界	14.300133
N	4799
重みの合計	4799
合計	69524
分散	43.68757
歪度	1.1541866
尖度	0.9160426
変動係数	45.624167
欠測値 N	0

各種の統計値

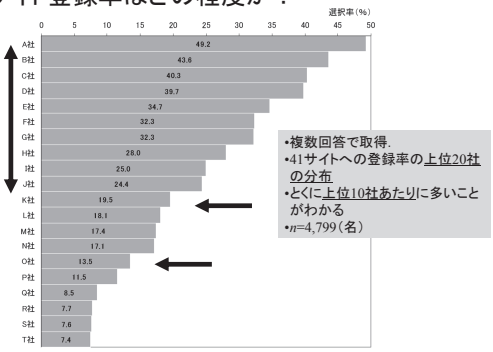
回答所要時間の分布は、質問数に応じて分布のモードの位置や、平均値の位置は変わりますが、過去調査ではいずれも単峰性のロングテールの裾の長い分布になります。早い人は数分で答えているが、遅い人はさらに長い時間かかる。この例だと最大が約 34 分で最小が約 1 分、わずか 1 分で答えている人もいるわけです。ここで用いた質問ボリュームで、1 分で答えられるかということはかなり疑問ですが、とにかく答えているわけです。

◎分析5：モニター登録情報の分析 ― 重複登録はどれくらいあるのか

最後の分析として登録者の登録率と重複登録について観察します。先ほどの ARF 調査では、より正確なやり方で回答者のアドレスを分析し登録情報を分析しています。しかし我々はここまではできなかったので、この調査時点でリストアップしたオンライン・パネルの国内サイト百数十社から、登録者数が多いサイトを調べ 41 社を拾い出しました。これを「ところで、インターネット上ではさまざまな会社がアンケートモニターを募集しています。あなたご自身がモニターとして登録しているものをお知らせください（あてはまるものをいくつでも）」の複数回答質問とし、41 社の中から回答者の登録しているサイトをすべて選んでもらうという方法で回答を集めました。つまり自己申告による回答という意味です。ARF 調査のようにメールアドレス分析から得られたものではないので、さらに検証が必要ではあるが、参考情報として、各サイトの“登録率（participation rate）”はどれくらいか、サイト間の“重複登録率”はどの程度かを推計してみようということです。これは複数サイトで同時に行った実験調査であるから分かったことで、おそらく類似の研究は国内ではないのではないかと思っております。

まず登録率を、上位 20 社につき並べてみます。どことは書きませんが、名前を言えば皆さんがよくご存知のサイトが上の方に並んでいます。ここで一番上の 49.2%は、ほぼ 2 人に 1 人が登録しているということです。上位 10 社くらいがずーっと続きあと少し落ちて、非常に小さいサイトまであります。ここで、回答者数は 4799 人です。さらにこの情報を 2 社ずつの組合せでどう重なっているかを計算した割合、2 サイト間の重複率を調べました。例えば、A 社と B 社に両方に登録の人が 23.9%、A 社と J 社は 14.9%、この表の各要素をそう読んでください。下三角行列部の重複率では、一番大きい値が約 24%、一番小さい値が 10.4%、等々です。同時にここで条件付き確率を使い簡単な計算で、「上位 5 サイトのどれか一つに登録」が 30.0%、「上位 3 サイトのどれか一つに登録」とすると 8 割近くが登録、「上位の 3 サイトのいずれにも登録」は 12.9%、……と、いろいろな求めたみたものです。つまりかなり重なっているのではないか、ということが予想される。上位の数サイトがよく似ているようにも見えるので、この重複率の表に MDS（多次元尺度構成法）を適用し分析し布置図を作ると、上の 4 社が似ていて、登録者に特徴のある E 社、これは「アンケートモニターになってお小遣い稼ぎを」とうたっているよく知られたサイトです。それで E 社は他と違った特徴があります。この他、先ほど新聞記事で紹介した併合となった数サイトがあります。少し異質のパネルがここに吸収されているようです。A 社と B 社が一緒になって、その後さらに他社が併合ということです。このように、この重複登録率の情報から、登

サイト登録率はどの程度か？



重複登録率の傾向：2サイト間の重複率，他

サイト名	A社	B社	C社	D社	E社	F社	G社	H社	I社	J社
A社	*	0.097	0.055	0.192	0.107	0.165	0.050	0.245	0.169	0.137
B社	0.2386	*	0.096	0.215	0.179	0.211	0.092	0.156	0.255	0.300
C社	0.2119	0.1990	*	0.344	0.169	0.349	0.269	0.256	0.249	0.193
D社	**0.2428**	0.2253	**0.2428**	*	0.277	0.403	0.301	0.275	0.285	0.263
E社	0.1961	0.1934	0.1792	0.2023	*	0.229	0.213	0.153	0.259	0.275
F社	0.1977	0.1898	0.2103	0.2207	0.1630	*	0.317	0.260	0.316	0.262
G社	0.1705	0.1619	0.1917	0.1971	0.1592	0.1736	*	0.172	0.143	0.160
H社	0.1932	0.1569	0.1694	0.1719	0.1300	0.1452	0.1267	*	0.226	0.186
I社	0.1594	0.1636	0.1534	0.1596	0.1398	0.1446	0.1094	0.1140	*	0.350
J社	0.1494	0.1700	0.1390	0.1521	0.1407	0.1315	0.1109	**0.1042**	0.1259	*

下三角行列が重複率
上三角行列は相関係数

- 有効回答数は7ﾊﾟﾈﾙでn=4,799(名).
- 登録率の高い上位10サイトにつき，すべての対の相互重複率を求めた(つまり45通り).
- 最小値は0.1042(10.4％). 最大値0.2428(24.3％)の重複率.
- 単純平均で“約16.95%の重複”
- これは多いか，少ないか.

組合せ	重複の割合(%)
上位3サイトのいずれにも登録	**12.9**
上位2サイトのいずれにも登録	**23.9**
上位3サイトの少なくとも1つに登録	41.9
上位3サイトのどれか1つに登録	**81.0**
上位5サイトのすべてに登録	7.3
上位5サイトのどれか1つに登録	30.0
上位10サイトのすべてに登録	1.9
上位10サイトのどれかに登録	20.7

録者が複数のサイトに登録することと、単純平均でざっと17%弱が重なっているらしいということ、その登録の仕方に何か特徴があることが見えてくるわけです。

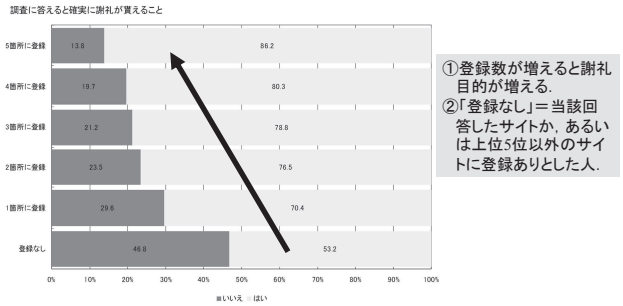
もちろん、ここでは自己申告の回答から推計したものですから、正確とはいえないかもしれない、ARF の実験調査のようにアドレスから逆探知して解析したものではありません。

最後にいくつか例を示します。謝礼を目的としている可能性があるわけですから謝礼と登録サイト数の関係を調べてみます（注：「調査に前向き“協力してもよい”と思う条件」として複数回答で提示の選択肢の一つとした「調査に答えると確実に謝礼が貰えること」と「登録サイト数」を6段階尺度に分類の回答とのクロス表を分析）。図にみるように「調査に答えると謝礼が確実にもらえること」に「はい」と答えた人の割合が、登録サイト数が増すにつれて増える、自明のこと当然そうなるだろうということです。

最後に同じく「登録サイト数（6段階尺度）」と「インターネットを利用した調査に、どの位の頻度で回答するか」（「年に1～2回」から「週に10回以上」まで8段階）とのクロス表を眺めてみます。これがそのクロス表で、これも当たり前でしょうが回答頻度と登録サイト数の間には

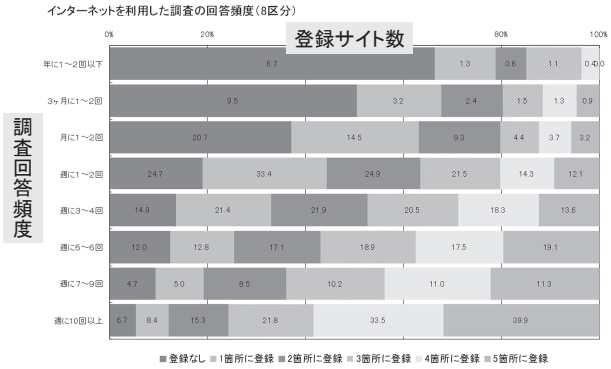
調査Ⅲ：登録サイト数と謝礼の関係

- 「登録サイト数」と「調査に答えると確実に謝礼が貰えること」の関係はどうか？
- 「登録サイト数」の多い人ほど「謝礼を目的」として調査に回答しているという傾向が顕著. ⇒謝礼目的で調査に回答
- 別の時点で行った調査では異なる傾向もある(例：調査Ⅱ).



調査Ⅲ：登録サイト数と調査回答頻度

- 「登録サイト数」と「インターネットを利用した調査の回答頻度（8区分）」の関係．（かなり自明な結果）



強い関係があるということです。以上のような情報から見て、やはり何となく“プロの回答者”がいるのではないか、ということが予想されるのです。

◎今後の動向：調査方法論研究のパラダイム・シフトにあわせて

ここまで雑多に実験調査で得たいろいろな例をみてきました。特に何か解答を示したわけでもなく、こうなっているらしいということをお話ししてきたわけです。ウェブ調査の回答者は、必ずしも“普通の人たち”ではなさそうだと、何となくそう思っております。特に“プロの回答者も混在している”のではないかなとも推測しております。ではどうすればよいのか、オンライン・パネルと言うが、これはうまく管理されているのかどうか、過飽和になっていないか、劣化していないのか、であればそこからどう抜けだしたらよいのか、ということです。回収率が落ちているというけれども、どうしたら改善できるのか、もあります。登録者集団のサイズをどんどん大きくする傾向にあるが、例えば250万人いますとか、当社は100万人いますとか、宣伝しているわけですが、そういうことでいいのかなとも感じます。繰り返しますが基本的には調査の品質をどう考えたらよいかということでしょう。いま見たようにネット専業社の併合もある、等質でないパネルの統合をどう考えるのか、我々の役割はどこにあるのか、と複雑な気持ちになります。調査の科学などと言われてきたけれどこれは大丈夫なのか…、疑問符がいっぱいあって、ここで軽々に正解を皆さんに示すことはできないだろう、ということです。

しかしパラダイム・シフト、研究の方向の変化は感じるわけです。特に、“調査環境の電子化”は確実に進むだろうということです。ここで双方向メディアと書いたとき、ふと思い出したことがあります。上村修一氏（注：元NHK放送文化研究所）が「市場調査」誌に書かれた「機械調査の誘惑」（輿論科学協会「市場調査」no.280～283（2009年no.1～4））という大変興味深い報告がありますが、これはぜひお読みいただきたい。つまり、こういう方向への展開の可能

性、先ほど私がユーザ・非ユーザ共に対象にしたウェブ調査、これに類似した部分もあります。上村氏のペーパーでは、（私の記事の読み違いかもしれないが）何となく機械調査に対してネガティブに書かれているようにもみえます。NHK でいろいろ試みられてきたことのいくつかは、非常に意味がある（あった）ことではないか、ということです。つまりは少し登場時期が早すぎたのではないか、これは技術的要素や基盤整備にかかる経費なども含めてでですが、そのように思えるのです。ともあれ、今後の調査環境の電子化に向けてこうした数々の実験調査の記録は非常に貴重だということです。

もう一つは、このところ急速に普及してきた、スマートフォン（高機能携帯電話）、iPad などの新たな情報機器類の調査システムへの影響があるだろうということです。表示画面も大きくなって、いままでの携帯電話の小さな画面でなく、大きい画面で調査ができるようになってきた。例えば、iPad で実際にテストしてみましたが、スムーズにできるわけです。そうなると（比較的価格の廉い）iPad を回答者に配っておけばよいわけですし、また実際に関心があるのですが、こうしたことの可能性が高いわけです。するとここでまた、対回答者のインターフェースが違ってきます。新しい機器・装置が登場すると、今までと違った問題が生じる可能性があります。こういうことが重要になってきます。

最後に今話題の“混合方式（mixed-mode）”について少し触れます。これはここで述べてきたことの延長線上にあるわけで、唐突に出てきたわけではないのです。混合方式とは、調査実施過程の時間軸の流れの中で、異なる調査方式をどう切り替えて使い分けるかということです。その目的は、回収率の向上、さまざまな調査誤差の評価と削減、調査期間の短縮、そして結果として調査の品質向上にあるとされてます。皆さんも我々も（そうとは意識せずに）行ってきたように、郵送調査を主調査方式として、郵送で未回収となった対象者にフォローアップとして電話で質問を行うなども単純な混合方式の一つです。これが電子化の進歩で複雑な調査環境となってきたが、一方調査方式の選択肢の組合せは非常に増えてきた、これらをどうやって複合的に使おうかという話しが出てきているということです。

要約すると、大きくは 6 つくらいの要素に分けて考えられます。①調査過程のどの段階で適用するか（適用場面）、②適用の時点・時期、③調査対象集団とその数、④調査回数・実施期間（1 回で済むか複数期間にまたがるか）、⑤調査方式の選択権（これが調査主体にあるか、回答者側にあるか）、⑥調査票の選択（一つの調査票を複数の調査方式で使い分ける、調査方式に合わせて調査票を変える）、などです。つまりいずれもが、調査品質に関係するわけです。これからの重要な研究要素だと思っています。新しい電子デバイスを使うこともこういう線上で考える必要があるでしょう。これに関する試みはすでにたくさんの本が出ています。例えば Dillman や de Leeuw など（追加文献）、さまざまな試みがなされています。

◎新しい流れ、方向、試行錯誤が続く ― さらに研究が必要なウェブ調査

新しい方向としては、特にマーケティング・リサーチなどでは別の世界への移行展開、いわゆるソーシャル・メディアやオンライン・コミュニティへの関心、そういう方向でいろいろと調査研究を始めているようです。例えば最近刊行された、またESOMARも薦める「オンライン・リサーチとソーシャル・メディア ハンドブック（R. Poynter （2010）: The Handbook of Online and Social Media Research, John Wiley）」という本をここで紹介しておきます。こういったものをご覧いただくとオンライン・リサーチとソーシャル・メディアをどう考えるかといったことが書かれています。流行のブログ、ツイッターなども見方を変えれば、定性情報、テキスト・マイニングやフォーカス・グループをどう使うかなどに通底することでしょう。

もう一つのキーは、ご存知のように各企業は調査会社を通さないで自分でオンライン・パネルを保有するようになってきたことがあります。例えば私もこの自分のコンピュータの登録をしていますが、時々アンケートが届きます。要するに顧客パネルの囲い込みが進んでいるということです。しかしこれは、パネル疲労劣化や回答者の偏った回答の提供となる恐れがあります。ロイヤリティは高いようにみえるけれども、潜在的な顧客の全体を代表するとは限らないから、やや危ない使い方だろうと思います。ともかくウェブ調査という調査方式はさまざまな使われ方、流れがあるということです。

私の関わる世界ではないので言いにくいのですが、ウェブ調査を「世論調査でどう使えるのだ」、「使えそうではないか」という意見があります。オンライン投票などもそうですが、こうした実験的な試みの一つが、先週開催の日本世論調査協会の発表にあったと聞いてます。その情報源はギャラップが電話調査で試みているGallup Dailyのオンライン版のようなものですがこれで何が分かるのか、非常に興味がありますがちょっと私には分かりません（参考：萩原雅之氏の記事「世論調査の弱点を補うネットによる世論“観測”の試み」、ジャーナリズム誌、2011年248号掲載）。

日本国内にも無数のオンライン・サーベイのサイトが登場しておりますし、小規模の調査会社は自前のシステムを作れないので、オンライン・パネルを借りたり、先ほど述べたように併合による規模の拡大もあります。そうではなくて、例えばコンソーシアム的にきちんと作ったパネルを共有化するという手もあるかなと思います。こうすれば標準化の問題もクリアされる、…といろいろありましょうが、先がよく見えません。

いずれにしても基盤整備が必要となるでしょう。調査は品質と費用がトレードオフなことはほぼ常識でしょうが、現状は“廉価で高い品質”を期待するという矛盾、ともかくもっと研究が必要だろうということです。

アメリカなどでは産官学の協働が非常に進んでいるのですが、残念ながら日本は、特にウェブ調査の認知度や理解度は非常に低いのが現状です。研究者として感じていることですが、このあたりはぜひ、皆さんも意識し、理解していただければと思います。

そういう事でウェブ調査については「さらに考えるべきことはたくさんある」という何ともまと

まりのない結び、お話しとなってしまい申し訳ないかぎりですが、私のお話しはこれで終わりにしたいと思います。ご静聴ありがとうございました。

（注1）関連する主要な資料
①大隅昇，前田忠彦（2008）：インターネット調査の抱える課題－実験調査から見えてきたこと－，「よろん」（日本世論調査協会報）2007；第100号：58－70／同2008；第101号：79－94.
②前田忠彦・大隅昇（2006）：自記式調査における実査方式間の比較研究－Web調査の特徴を調べるための実験的検討－，エストレーラ2006； No.143, 12－19.
③大隅昇（2006）：インターネット調査の抱える課題と今後の展開，エストレーラ2006；No.143: 2－11.
④Couper, M.P.（2008）：*Designing Effective Web Surveys.* New York: Cambridge University Press.
⑤大隅昇（2009）：M.クーパー著『効果的なウェブ調査の設計』を読んで，「よろん」（日本世論調査協会報），第104号，p50-60.（④の書評）
（*）http://wordminer.comquest.co.jp/smr/pdf/f00017b.pdf
⑥Groves, R.M., Fouler, F.J., Couper. M.P. et al.（2004，2009）, *Survey Methodology.* New York: John Wiley.（朝倉書店から2011年春訳本刊行予定）.
⑦M. Couper, Online Research in the U.S.　Yesterday, Today, and Tomorrow -, JMRA 41st Topics Seminar, September 21, 2010,Tokyo.（日本マーケティング・リサーチ協会第41回トピックスセミナー；2010年9月21日開催），セミナー配布資料.
⑧AAPOR Online Panel Task Force Report (2010): AAPOR Report on Online Panels.（約90ページからなるレポート）.（*）http://www.aapor.org/Home.htm
⑨林文，大隅昇，吉野諒三（2010）：ウェブ調査から何を読み取るか―基底意識に関する実験調査―，日本日本行動計量学会，第38回大会，2010年9月23日（埼玉大学），予稿集，30-33.
⑩矢口博之，大隅昇（2010）：電子書籍と読書行動についての実験調査，日本日本行動計量学会，第38回大会， 2010年9月23日（埼玉大学），予稿集，26-29.
（※）上の多くは以下のHPから閲覧可能
http://wordminer.comquest.co.jp/index.html
調査方法論関連情報資源：Survey Methodology Resource (SMR)
http://wordminer.comquest.co.jp/smr/index.html

追加文献：

1）De Leeuw, E.D., Hox, J.J., and Dillman, D.A. (2008): *International Handbook of Survey Methodology*, Lawrence Erlbaum Associates.

（*）ここの第6章として“Mixed mode surveys, When and why, pp.299-316”がある.

2）Dillman, D.A., Smyth, J.D., and Christian, L.M. (2009): *Internet, Mail, and Mixed-mode Surveys*, Third Edition John Wiley & Sons.

（おおすみ・のぼる）