

ウェブ調査とはなにか？ －可能性、限界そして課題－（その 2）

大隅 昇
(統計数理研究所 名誉教授)

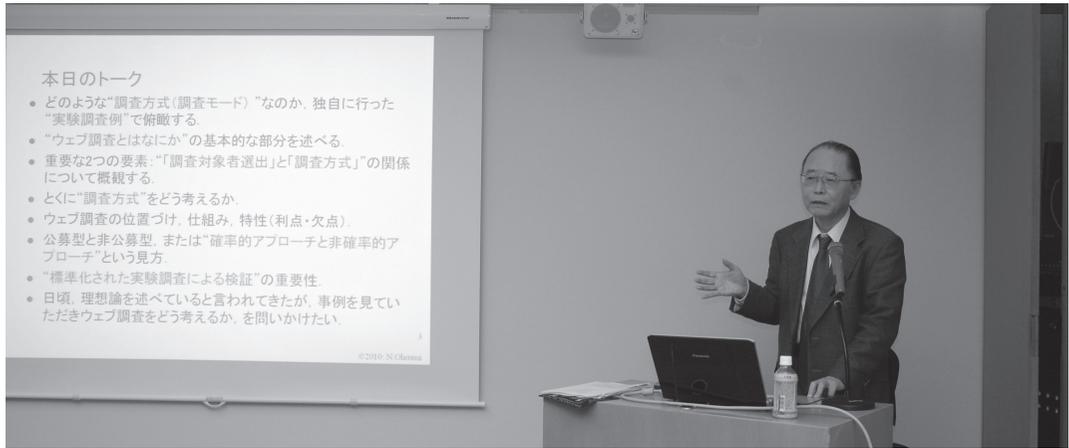
◎複数の調査方式を比較分析することが重要 — 実験調査例の紹介

ここから少し急いで、調査例を順番に見ていきたいと思えます。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はほぼ 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・パネルを用い同時に比較



実験調査の要約(1)

調査コード 調査方式	調査プロジェクト名	プロジェクト 総括者	参加機関名	実施年次	実施回数	計測標本数 回収標本数
(調査Ⅰ) 計測標本自記式 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	延べ計測標本数 (364,628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 ハイムリサーチ 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, Bは郵送調査 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	延べ計測標本数 (9,322人) 延べ回収標本数 (6,356人)
(調査Ⅲ) ウェブ調査 (4) A' 特別比較	Web実験調査 (B'調査)	統計数理 研究所 博報堂	アイブリッジ インテリジ エンテンス 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	同時に 全4社で1回	延べ計測標本数 (26,497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	調査プロジェクト名	プロジェクト 総括者	参加機関名	実施年次	実施 回数	計測標本数 回収標本数
(調査IV) ウェブ調査	親と子の生活行動 と態度に関する調査 (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	郵送法で郵付: A' 344 (380人), A' 348 (625人) を計測標本 合計1,009 (A) 回収標本: A' 344 (322人), A' 348 (511人); 合計833 (A)
(調査V) ウェブ調査 郵送調査	基礎意識構造の 統計的研究 (仮想的価値観)	林文教授 (東京英和女学院大学)	博報堂-東京サー ベイ・リサーチ 7月 2010年2月 日経リサーチ	2010年2月 (2/19 ～2/25)	1回	郵送法で住民基本台帳 から推計: 郵付: 延べ計測標本数 (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・ グループ	2010年1月 (1/8 ～1/12)	1回	住民基本台帳から推計: 計測標本数 (756人), 回収標本数 (460人)

しています。標準化し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えます。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が。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が、調査方式による違いがあるかどうか調べています。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が、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

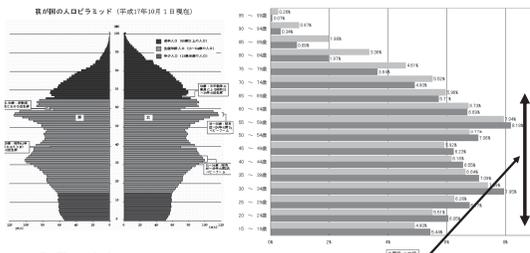
では、まず分析1から始めます。ここで大事なことは、登録者はどう作ったのか、ということです。最近はずすかに各調査サイトを見ていただくと、「我が社のパネルはこういう属性構成になっています」などと書いてあります。男性が何%、女性が何%で年齢比がこうなっている、という人口統計学的情報はあるのですが、ご覧になるとすぐに分かるように、比べてみるとかなり違うことと、また当たり前ですが国勢調査に類似した構成比分布が再現されることはほとんどないわけです。大事なことは、「どういうパネルを作っているか、登録者の集め方は」ということを知る必要があるのですが、通常この情報はほとんど得られません。

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。

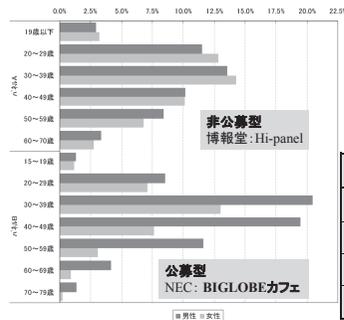
国勢調査情報(2005年)の確認



- 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!!

多くの場合、調査対象年齢がある範囲に限られる

調査Ⅳ:親子調査の登録者集団:性年齢区分



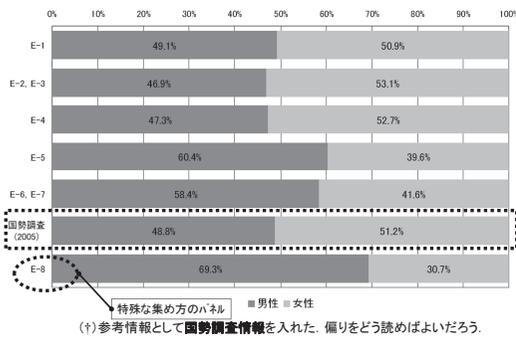
- 2パネルに差違がある。
- パネルAは関東・関西都市圏
- パネルBは全国登録者
- パネルA性比(約50:50)
- パネルB性比(約67:33)
- 回答分析に影響する。

2パネル構成	パネルA	パネルB
登録者数(人)	15,803	1,004,778
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%)

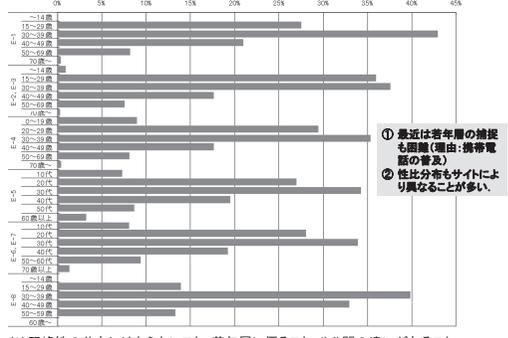
今度は調査Ⅲですが、これはほとんどがネット調査専業社サイト、6社8パネルで比べた例です。ただしここで、E-8は特別な集め方を行っています。これは主にガソリンスタンドを訪問する人をサンプリングしたものです。男性が多いから、当然男性登録者が多くなっています(注:これは実験調査当時の情報、現在は、別の集め方も行っているので属性に変化がある)。残りの7パネルは、一般に公募したパネルで、回収率がこれだけばらつきますが、これは(いままでお話ししたことから考えても)当たり前でこうなるということです。我々が日常的に行うようにデモグラフィック要因で分類して回答分布を観察すると、これらが違うことはほとんど自明なわけです。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)

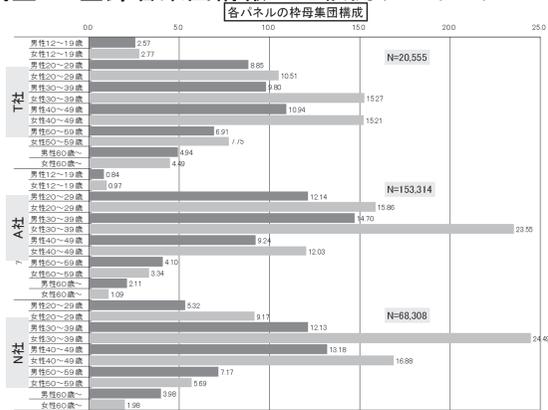


調査Ⅲ：E調査の登録者集団(年齢区分)



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、やり方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えます。“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/調査事件	E-1	E-2	E-3	E-4
調査期間	2006/3/23 ～2006/3/29	2006/3/24 10:00 ～2006/3/29 10:00	2006/3/24 10:00 ～2006/3/27 22:00	2006/3/23 10:00 ～2006/3/23 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 (件) マネジメント				
計画標本	600(件)	1,191	1,221	3,429
依頼配信前除外	225(7件)	0	0	1
調査依頼配信数	1,437	1,191	1,221	3,361
有効回収回答数(%)	715 (49.8)	737 (61.9)	606 (49.6)	243 (7.2)
無効回収回答(PO)終了(%)	7 (0.5)	4 (0.3)	5 (0.4)	4 (0.1)
無回答(%)	717 (49.9)	450 (37.8)	619 (50.7)	2,414 (71.9)
未着(%)	19 (1.3)	不明	不明	67 (2.0)
無回答(%)	698 (47.9)	454 (38.1)	614 (50.3)	2,347 (70.7)
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)	8 (0.7)	50 (1.5)
回答未受領者数(%)	398	875	976	2,966
アクセス数(総べ数)	758	875	717	1,056
リソースの身元は総アクセス数(総べ数)				
アクセスのみ(総べ数)(%)	6 (0.8)	95 (10.9)	46 (6.4)	62 (5.9)
回答未受領者数(総べ数)(%)	392 (51.7)	875 (100)	976 (136)	2,904 (27.5)
回答未受領者数(総べ数)(%)	不明	不明	不明	不明
回収回答(総べ数)(%)	720 (94.7)	741 (84.3)	606 (84.5)	247 (23.3)
重複回収回答(%)	生起しなし	生起しなし	生起しなし	生起しなし
非常時未回答				
非常時未回答				
調査依頼未達(総べ数)	18	不明	不明	67
登録者数標準日数	2006/3/23	2006/3/23	2006/3/23	2006/3/23
有効回収数(ウェブページ)	0回	0回	0回	0回

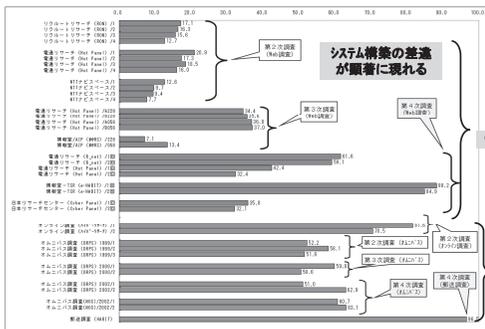
※パネルの比較・評価にはこうした要約表が必要(作れるかどうか)

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、

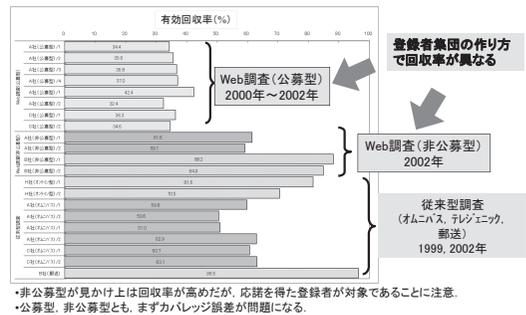


非公募型は登録者を集めた後のロイヤリティが高いわけで、しかし計画標本の応諾率が低いわけですから、これが全体の回収率を評価する際には注意が必要ということになります。

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）



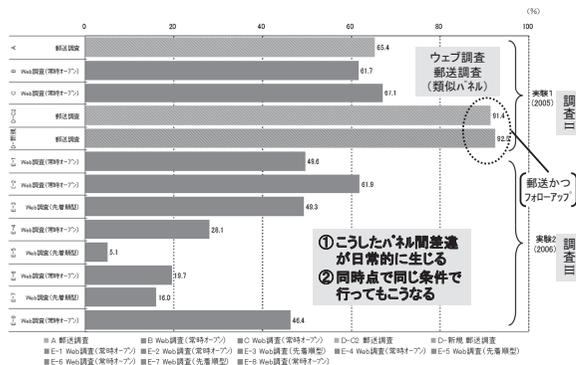
•非公募型が見かけ上は回収率が高めだが、応諾を得た登録者が対象であることに注意。
 •公募型、非公募型とも、まずカバレッジ誤差が問題になる。

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よ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,ウェブ;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ください(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)複雑で、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イケないことが分かっているが、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が。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%というわずかな数%、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、登録者があつちに登録こつちに登録という状



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、パネルがどのような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、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

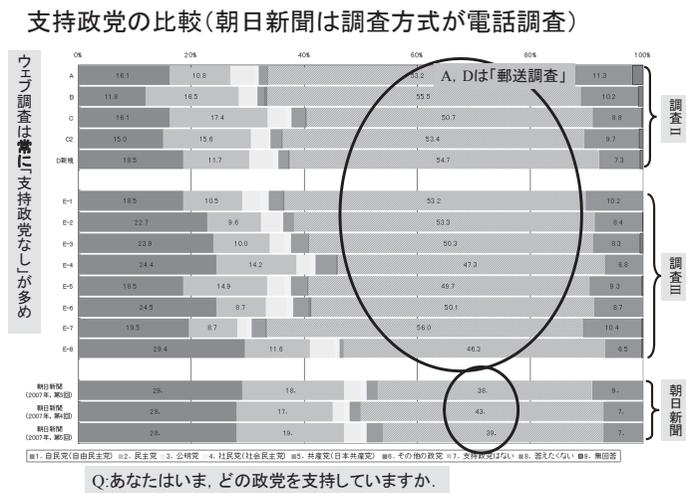


つは、パネル間の回答比率が、つまり、異なるパネルで得た同じ質問についての回答選択比率がどう変動するのか、どれくらい揺らぐのかを比べてみるということです。

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。では年齢別にブレイ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りません)。

この例だと、2006～2007年に行ったE調査に調査時点に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と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。

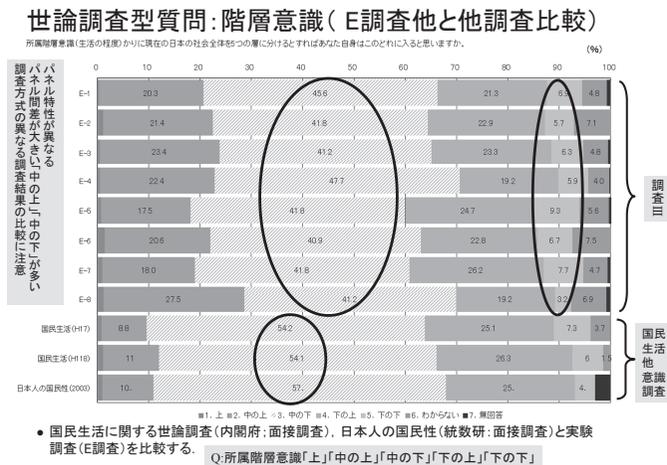


こうした現象をただちに調査方式間の差異と言ってよいのかは微妙です。しかし、「黙従傾向」や「社会的望ましさ」などはどうなのか、例えば調査方式が電話調査だと調査員の質問にじかに



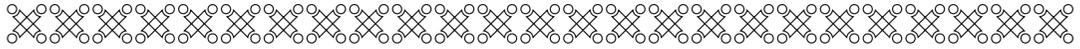
は答えにくい、偏りが入ると言われている。一方、ウェブ調査は自記式であるからそういう影響を受けにくいのか、あるいは何なのかということはある、これらはさらに精査しなければ分からないということです。

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思う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。



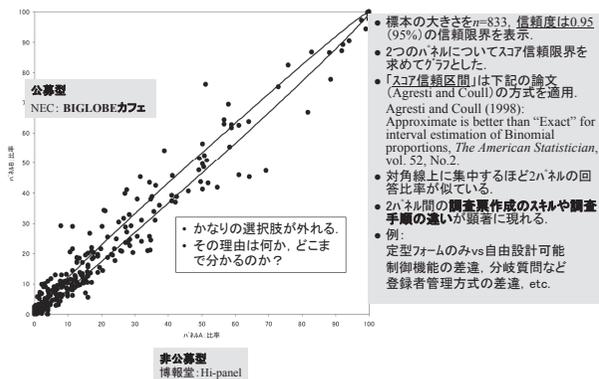
◎回答分布はパネル間でどれくらい違うのか

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でも、このようにばらついていきます。この例ではこうしたばらつく理由がお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が（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います。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事です。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されていますが、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

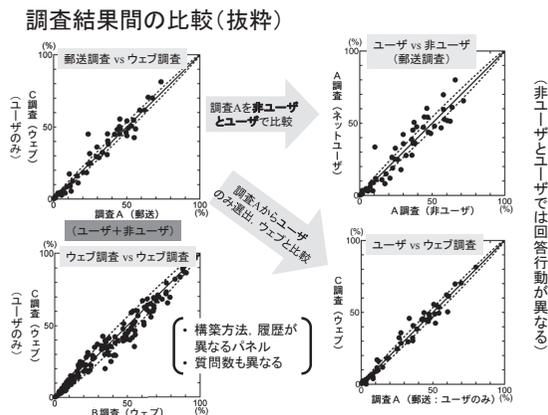
調査要約(一部を抜粋)	A調査	B調査	C調査
対象者(登録者集団と抽出法)	2004年度HABITほぼ全数調査	2004年度HABITのうち2005年度Hi-panel参加意向者(注1)	2004年度Hi-panelより層化無作為抽出
調査方式	郵送調査	ウェブ調査	ウェブ調査
調査時期と回収期間	2005年1月約3週間	2005年3月1週間	2005年3月1週間
依頼数(注2)	4,630	857	2,499
回収数	3,026	529(507)(注4)	1,678
回収率(注3)	65.4%	61.7%	67.1%
備考(実査時の条件など)	定例調査最終回に追加調査として同時の実施	Hi-panel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ールの発信数。
 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B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。

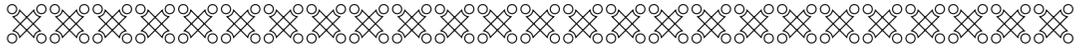


ここではその一部を眺めます。(この表には)3つの調査、A調査、B調査、C調査がありますが、このうちのA調査とC調査を取り上げます。ここでA調査(郵送調査)は「インターネット・ユーザと非ユーザが混在」したパネルです。またC調査(ウェブ調査)は「インターネット・ユーザのみ」からなるパネルです。A調査では調査時にインターネット・ユーザと非ユーザを確認してですから、2つに分けることができます。これを分ける前と分けた後でそれぞれ回答選択比率を比較してみます。つまり「(A調査)対(C調査)」と「(A調査のインターネット・ユーザのみ)対(A調査の非ユーザ)」「(A調査のユーザのみ)対(C調査、つまりユーザ)」、これらの回答選択比率を比べた散布図です。ここで明らかなのは、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いますが、この利用頻度情報でブレークダウンすると、ヘビー・ユーザとそうでないユーザでは答えが少し違ってくるとも分かります(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。



今回は今年実施した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: $d_{ij} = \sum_k |p_{ik} - p_{jk}|$ ($i, j = 2$ つのパネル; $k = 質問の選択肢$))。

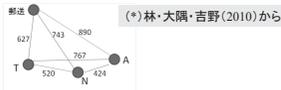
距離の模式図を描くと、T社(非公募型)は郵送に近い、ところが(公募型の)T社とA社はかなり離れている、特にA社は郵送からかなり離れているといったようなことが分かってきます。つまり、ウェブ調査の3パネル間にも差があって、しかもそれぞれの郵送調査との距離もパネルによって異なります。事前に予想されていたことがこうして数字で出てくるわけです。もちろん質問内容にもよりますが、ここでは「すべて同じ質問」を使い、ウェブ調査については「同時点で実施」、郵送調査では同じ質問を使っていますが調査票のフォーマットはウェブ調査のそれとはやや異なり、また、ウェブ調査とは調査時点が若干ずれています。ともあれ、かなり同じよ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

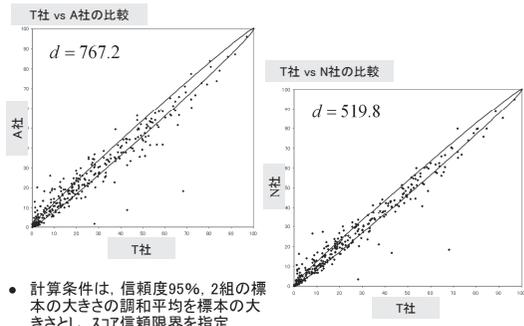
- 回答比率のパネル間距離を求めてみる(下の表と図)。
- 質問項目選択肢について、選択肢のある2調査(パネル)との回答選択率の差を合計する。
- 明らかに“パネル間、調査方式間の差異”がある。
- 郵送調査からみると、ウェブ調査Cが最も離れ、ウェブ調査Aが近い。
- ウェブ調査間にも差異がある(NとAが近く、TとAが遠い)。

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

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

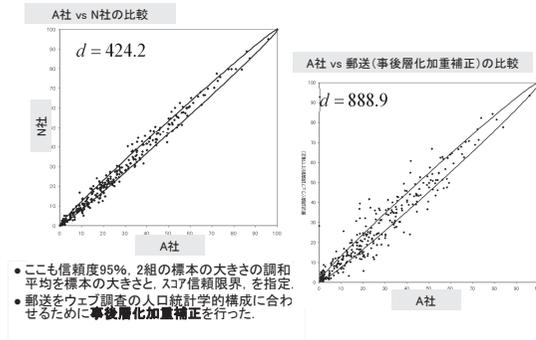


調査V: 質問選択肢回答比率の比較(2)

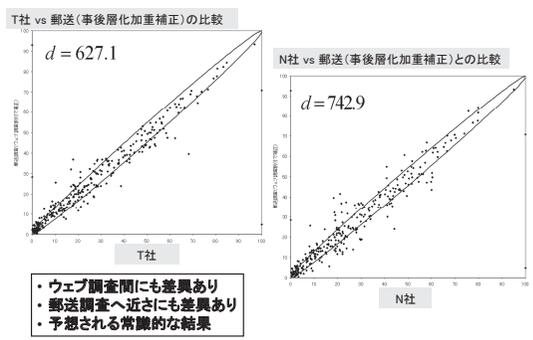


あと(スライドに)数ページ描いてある図は、ウェブ調査(3パネル)と郵送の、合わせて6通りの組合せに対する散布図を描いたわけですが、これらを見ていただければ、いま述べた相互の類似関係がより分かりやすいでしょう。

調査V: 質問選択肢回答比率の比較(3)



調査V: 質問選択肢回答比率の比較(4)





(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、数日あとならなると答えてこない人もいます。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。こころ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らせてもらってます。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、また米国のクーバー氏もこの研究を進めているところです。また、これを始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ができます。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から、ということはあるけれども、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)

		回答完了者のページごとの戻り回数カウント							前に戻ることなく最後まで回答	
		00	01	02	03	04	05	06		
E-1	質問オープン型	713	732	722	718	717	716	713	718	612
E-2	質問オープン型	737	737	737	737	737	737	737	737	737
E-3	先着順型	602	602	602	602	602	602	602	602	602
E-4	質問オープン型	943	951	953	947	946	950	950	943	943
E-5	先着順型	612	616	617	614	616	619	613	613	595
E-6	質問オープン型	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	質問オープン型	432	432	430	430	434	430	435	433	415

		N [サンプル数に対する戻り回数ページの増加率%]							戻り回数平均	
		00	01	02	03	04	05	06		
E-1	質問オープン型	713	2.7	1.3	0.7	0.6	0.4	-	0.7	54.2
E-2	質問オープン型	737	-	-	-	-	-	-	-	100.0
E-3	先着順型	602	-	-	-	-	-	-	-	100.0
E-4	質問オープン型	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	質問オープン型	597	0.5	0.2	0.3	-	0.3	0.3	-	96.3
E-7	先着順型	595	0.2	0.2	0.3	0.3	0.5	0.3	-	96.3
E-8	質問オープン型	432	-	1.4	0.9	0.5	1.4	0.7	-	99.1

※(参考)各ページの内容

- 00 開始ページ・結果チェック質問
- 01 A「製品の用途で実用性について
- 02 B「数量について
- 03 C「耐久性について
- 04 D「インターホンの録音機能について
- 05 E「機能性について
- 06 F「価格について

※E-1は複数チェック質問が複数ページにわたって行われており、他社とそろえてカウントした

このサイトは記録出来なかった

「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

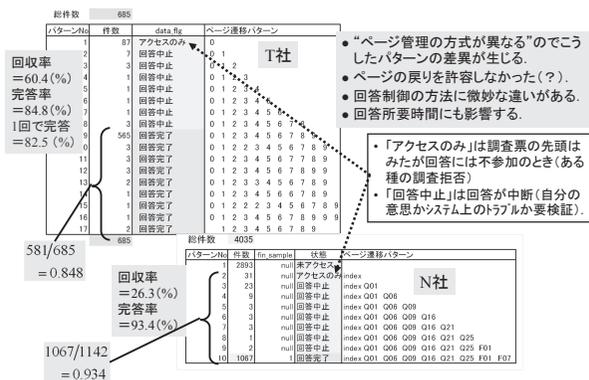
過去の例でいうと、完答率100%はまずない。7割から8割、高いとこのくらい(90数%)です。当然これが回収率に関係し影響します。先ほど調査評価表を紹介しましたが、そこでも見たようにさまざまな指標が必要になるということです。あのような表を作っておかないと分からないのです。繰り返しになりますが、いわゆる有効回収率にも関係しています。このE調査の場合、データが取れなかったサイトもあります。取れたサイトだけみると、完答率は90数%です。調査Ⅲですとほぼ完答ですが、ごくわずか数%完答できない人が出ています。

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、完答した人といういろいろあります。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来るかも知れない。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)



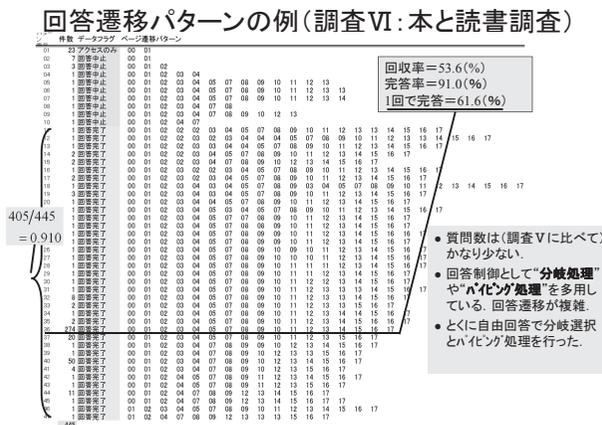
ここで調査VIの例をみます。T社の上とほぼ同じパネルを使った調査です。いま話題の電子書籍に関連した調査を行ったものです。この調査票はやや複雑で、分岐処理やパイピングを多用しています。例えばある質問で「あなたはどんな本を何冊読みましたか」と聞くと、読んだとして選んだ本のジャンルに合った部分だけを次の画面に表示し冊数の記入を促す、といった処理を行う。これの実物はサーバに繋がっていないと見られませんから、イメージのみでご覧いただけますが、例えばここに本が、どの本を読みましたかということでクリックすると、仮にここで3つクリックしたとすると、次のページ進んで、その選んだ3箇所の本のジャンルだけが表示されるわけです。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%となっています。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います。

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ふ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いった実験に、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、そのように違って来る。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が。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。このような現象があるので、最近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、これに類した事、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。しかし最近ほとんどこうしているはず。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。

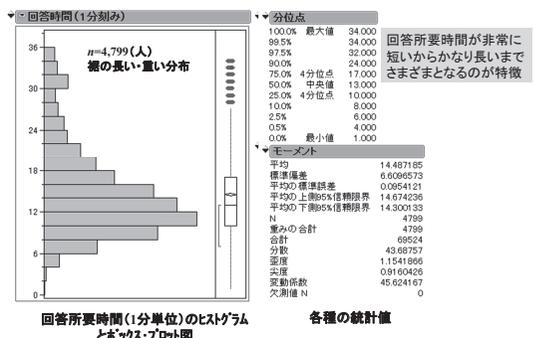
8質問(MA)の回答所要時間とのクロス分析

- 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	ピアソニック二乗統計量	
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(41個)	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

(注)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、わずかに 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。

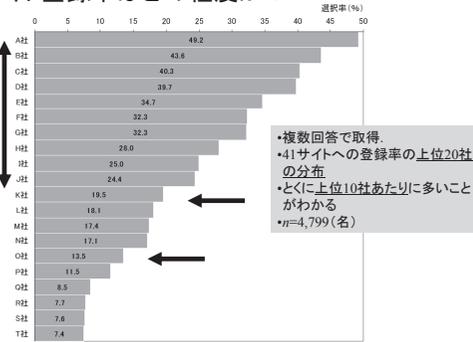
◎分析5: モニター登録情報の分析 — 重複登録はどれくらいあるのか

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。しかし我々はこまめではできなかったのもので、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考えております。

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、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。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	*

上三角行列が重複係数

- 有効回答数は7ハネルでn=4,799(名)。
- 登録率の高い上位10サイトに付き、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3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	61.0
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1つに登録	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ませんが、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。

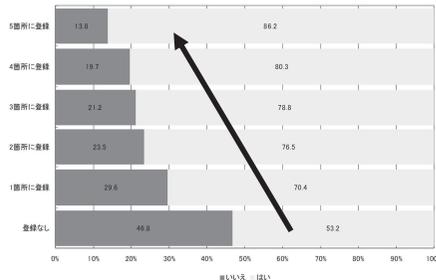
最後にいくつか例を示します。謝礼を目的としている可能性があるわけですから謝礼と登録サイト数の関係を調べてみます(注:「調査に前向き「協力してもよい」と思う条件」として複数回答で提示の選択肢の一つとした「調査に答えると確実に謝礼が貰えること」と「登録サイト数」を6段階尺度に分類の回答とのクロス表を分析)。図にみるように「調査に答えると謝礼が確実にもらえること」に「はい」と答えた人の割合が、登録サイト数が増すにつれて増える、自明のこと当然そうなるだろうということです。

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)。

調査に答えると確実に謝礼が貰えること

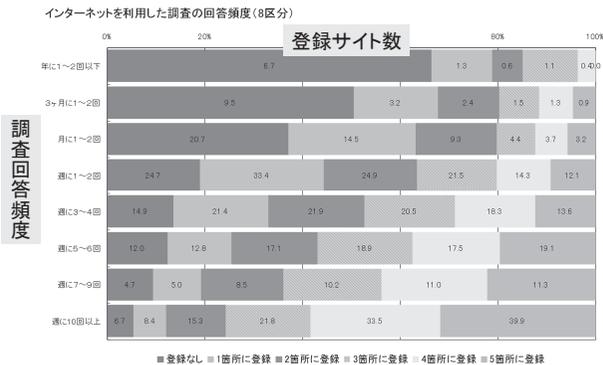


- ①登録数が増えると謝礼目的が増える。
- ②「登録なし」=当該回答したサイトか、あるいは上位5位以外のサイトに登録ありとした人。



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、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。回収率が落ちてき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專業社の併合もある、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(高性能携帯電話)、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、スムーズにできるわけです。そうなる(比較的価格の安い) iPad を回答者に配ってあげればよいわけですし、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います。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、…いろいろありまじょうが、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、お話しとなつ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

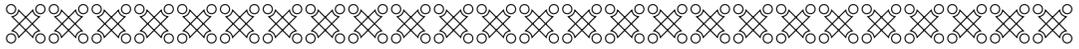
- ① 大隅昇, 前田忠彦 (2008) : インターネット調査の抱える課題－実験調査から見えてきたこと－, 「よろん」 (日本世論調査協会報) 2007; 第100号: 58-70/同2008; 第101号: 79-94.
 - ② 前田忠彦・大隅昇 (2006) : 自記式調査における実査方式間の比較研究－Web調査の特徴を調べるための実験的検討－, エストレーラ2006; No.143, 12-19.
 - ③ 大隅昇 (2006) : インターネット調査の抱える課題と今後の展開, エストレーラ2006; No.143: 2-11.
 - ④ Couper, M.P. (2008) : *Designing 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>
- ⑥ Groves, R.M., Foulter, 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) .
 - ⑦ M. 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.
 - ⑧ AAPOR 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一, 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)