

オンラインリサーチシステムの構築

○柳原 良造

〔株式会社マーケティング・サービス〕
〔ハイパーリサーチ株式会社〕

吉村 春彦

〔ハイパーリサーチ株式会社〕

1. オンラインリサーチシステム構築の動機

かつて、60年代以前に於いては、調査業務は、典型的な労働集約型のものであり、企画・実査・集計（但し、ソーターは使用された）は、専ら人力に拠って行われた。当時は、相対的に労働力は低廉であり、かつ、データの価値づけは高めで、発注者、受注調査機関双方が納得するコスト合意が成り立っていると見えた。しかし、その後の高度成長期に入り、条件が一変する。その大きい変化の一つは労働力の高騰である。これは調査コストを大きく押し上げるものであった。ところが、社会調査の世界では予算前例主義のため、又、産業界では、競争の激化によるコストダウン圧力のため、その後には、ターゲットマーケティングの敷衍による一製品当たり調査予算の縮小が生じ、調査コストの削減が求められることとなった。折しも、コンピューターの普及があり、調査機関側は、これを導入し、合理化を計ったが、調査コスト全体に占める集計費用はさしたるものではないため、発注者側の要求には到底及ばない程のものであった。データ使用者の持つ調査データへの価値づけと実コストとの間に大きい乖離が生じたわけで、この風潮の中で、“まがい調査”が横行し、被調査者の非協力も世の動きにつれ増加、調査の質の低下が、心ある人々から憂慮される事態に立ち至ったのである。調査データの価値づけは、発注者側の判断によって決せられる。従って、調査機関側は、許されたコスト内で調査を立派に仕上げねばならないが、そのためには、調査コストの大きい部分を占める調査員コストをいかに切り詰めるかが鍵となる。最も安易な手段は、調査員手当の切り下げであるが、労働価値が高騰し、仕事のつらい調査員にならずとも他にいくつもの収入取得の選択肢がある世の流れの中であって、これを行うことは、即、調査の質的低下をもたらすものであった。

調査の質をしかるべき水準に保ちながら、実査コストの削減を計れる道はないか。

この問に対する殆ど唯一の可能性は、調査の機械測定法への転換であると考え、この考えの一端を1984年版「サービス産業年鑑」に記した。

2. システム構築の経過

調査の機械測定法とは、具体的には、コンピューター端末を回答者宅に配布設置し、これをモニターパネルとして調査を行おうとするものである。

この方法の利点としては、調査員手当・交通費が通信費のみで済むため、実査コストが大幅にダウンし、又、データ入力コストが不要となること、回答者が回答の為の時間制約から解放されること、これは同時に個別面接聴取調査では獲得不可能な回答者の回答を得る機会が増大すること、通常調査で大なり小なり伴わざるを得ない人間が直接的に介入することのバイアスを避けられること、そして、あわよくばではあるが、モニター契約によって、又、ゲーム感覚で対応できる可能性によって、しかるべき水準の回収率が維持できること等がある。

一方、当然ながら大きい欠点も存在する。まずコストダウンについては、コンピューター端末が、特に当時に於いては、非常に高価で、又、モニター依頼・機器設定費も高くつくため、必ずしも期待通りのコストダウンは達成できないこと、一般の回答者がコンピューター操作、特にキーボード操作に馴れていないこと、住民基本台帳から無作為抽出されたランダムサンプルの全てからこのモニター参加への応諾が得られそうにはなく、結果として、モニターが何らかの偏りを持つであろうこと、モニターであることによる学習能力アップや調査擦れが生じるであろうこと、人間が介入しないことによる回答の無責任性が危惧されること等々である。

これらの欠点のうち、端末操作問題については、昭和59年9月より三鷹地域で行われたNTTのINS実験に参画し、ビデオテックス（当時の呼称はキャプテン）を用いた操作能力実験を行い、一定の成果を得た。

その後、キャプテン活性化の国策に基づき、キャプテンサービス（現NTTビジュアル）がリサーチパネルを設定することとなり、縁あって、筆者がその設計とサンプリングを行うこととなった。その際に、この調査法を科学的調査法であらしめるために、留意し、こだわったポイントについては後述するが、それが現在に於いても我々のシステム構築の基本におかれている。その後、時代の流れの中で、キャプテンシステムから、徐々に、パソコンネットシステムへと転換しつつある現状にある。

3. システム構築に当たっての遵守事項

本調査システムの構築は、一つの科学的調査法の確立を目指したものである。調査が科学であることを求められる理由は、少数の標本集団のデータをもって、予定された誤差範囲内で、母集団推計が成り立つことにある。そして、科学的調査であること条件は、

- ①母集団が明確に定まっていること
- ②回答者が、ランダムサンプルであること
- ③誤差を小さく定め得るだけのサンプル数であること
- ④母集団代表性を損なわない回収率であること
- ⑤サンプリング誤差以外のバイアス発生を抑えること

に集約されよう。これらの点を考慮して、回答者パネル構築に当たっては、以下の配慮を実行した。

(1) 母集団の明確化

母集団は、調査地域、年齢範囲等は容易に定められる所であり、これを定めた。同時に筆者は、母集団規定に調査時点が加えられるべきとの持論を持ち、この考えに従って、通常調査は3日以内、テーマによっては1日で調査完了できる枠組みを立てた。

(2) ランダムサンプル性

サンプリングは、母集団地域内の町丁から、確率比例抽出法によって、世帯又は個人を住民基本台帳からランダム抽出し、抽出された世帯又は個人に協力依頼し、応諾した世帯又は個人に端末を設置し、パネルに組み込んだ。この場合、全員の応諾は不可能であり、欠損サンプルに対しては、属性を同じくする予備サンプルを用いざるを得なかった。この応諾率は調査の質を左右する。従ってより高い応諾率獲得のため、協力依頼行動を多重に長期に亘って行っている。しかしながらこの努力にもかかわらず、応諾率は、町丁により8割に達することもあるが、平均的には6～7割で、欠損部分は予備標本で代用せざるを得ない実態にある。従って、ランダムサンプルによるクォータサンプリングと考えるべきである。

(3) サンプル数

サンプル数は1,000を用意した。(但し、今回提示するデータの取得時点は、パネル設定進行途中であり、未だクォータが未完状態の750サンプルのものである。)

(4) 母集団代表性を損なわない回収率

この場合の回収率は、応諾率×当該調査回収率で定まる。従って、応諾率の上昇と高い当該調査回収率が獲得されねばならない。ある調査時点では応諾率は既定であるので、残る努力としては、当該調査回収率の向上しかない。そのため一定時刻ごとに未回答者をチェックし、協力依頼電話を重ねている。この結果、3日調査の場合には、大抵は80%台の回収率を得ている。80%台の当該調査回収率であっても応諾率が70%であれば、実回収率は56%に止まり、決して満足できる比率ではない。問題はこの回収率の調査結果が、通常調査の結果と一致するか否かにある。(この点の検証は後述)

(5) サンプリング誤差以外のバイアス源排除

サンプリング誤差以外のバイアス源として大きいのは、質問文と人間の介在であろう。質問文は、ディスプレイ上の制約のため、やや短文調となりやすく、又、見た目には、客観的、冷静な印象を持つ。これがプラスなのかマイナスとして働くのかは明確ではない。しかし、実験で、広告画面に対するイメージをとった場合、パネルへの電話調査による結果と比較すると、認知率、及び、肯定的イメージ言語への差はなかったが、否定的イメージへの反応が高めに現れた。これについては、批判型人間がパネルに多い偏りによるとの解釈が成り立ち得るが、他質問への回答をつきあわせた

場合、必ずしもそうではなく、例え電話で顔が見えなくても人が介在した場合、否定的イメージへの回答が控えめとなるのではないかと考えられた。もし、そうだとすれば、この手法は人が介在するバイアスを排除し、回答者の思い通りに遠慮会釈なく答えられるもので、本音を出しやすい、言い換えると人が介在するバイアスを排除できる働きを持つのではないかと考えられる。

以上の手立てを講じて、モニターパネルとして運用される本システムの対策点としては充分ではない。以下にその他の手当てについて述べる。

(6) 調査ズレサンプルの排除——ローテーション

パネルを用いる場合、遅かれ早かれ調査馴れから調査のすれっからしを生ずる。従って、調査経験回数が事前研究で見出されている耐用限界回数に達するか、それ以前でも回答態度に変化が見られるサンプルは、新しいサンプルと交替されねばならず、このローテーション作業は常時進行している。

(7) データ補正の基準項目の発見

以上の手続きを確保しても、パネル設定時の応諾率のため、純粹のランダムサンプル構成に対し、何らかの歪み、偏りを持つであろうことは避けられない。従って、必要に応じデータ補正の必要が生ずる。その場合の補正基準となる事項が何であるかが明確にされねばならない。性、年齢は割り当てられるので、その他項目からこれを見出す作業を行っている。

4. パネル構成の偏りの実態

キャプテンシステム設定時には、鈴木裕久教授の情報処理能力テストに於いて偏りが出たことは、東大新聞研紀要に於いて同教授が明らかにしている通りである。現在設定中のパネルに於いても、最近2ヶ月間の情報機器購入実績が2.2%、回収数は異なるが選挙後の第4回調査結果では自宅でのパソコン使用が3.9%と高く、情報処理馴れ人間が多い偏りを持つと想定される。

先述の通り、今回報告データを得たパネルは、構築中のもので、現段階では、当然偏りを持っているので、まず、その点を明らかにしておく。まず、性別では有意差は出ないが、年齢については表1に示す偏りがある。

表1. 年齢分布の偏り

	合計	～24才	～29才	～34才	～39才	～44才	～49才	～54才	～59才	～64才
住民基本台帳(平成9年)	12,700,158	13.2	14.5	11.7	9.6	9.4	12.7	10.5	9.9	8.6
ハイパーシステム調査	612	7.0	11.6	14.1	15.4	12.4	13.9	9.0	7.8	8.8

$$\chi^2 \text{乗値} = 56.176 \text{ (****)} > \chi^2(0.05) = 15.507$$

母集団が同一である他データの入手が困難のため検定が難しいが、この手法はオンライン型の留置法に該当すると考え、筆者手持ちの他の留置調査結果から母集団を合わせ、検定した所では、表2～表5に示す偏りが見られる。

表2. 支持政党

	合計	自民	民主	公明	社民	共産	自由	支持政党なし・その他
訪問留置調査	875	16.2	7.7	5.2	1.7	4.8	3.0	61.5
ハイパーシステム調査	612	10.6	10.2	4.2	2.3	4.9	1.3	66.6

$$\chi^2 \text{乗値} = 21.157 \text{ (***)} > \chi^2(0.05) = 12.592$$

表3. 行政の情報公開

	合計	全くそう思う	まあそう思う	あまりそうは思わない	全然そうは思わない	関心がない
訪問留置調査	874	47.4	41.3	7.8	0.9	2.6
ハイパーシステム調査	612	61.1	32.5	3.9	0.2	2.3

$$\chi^2 \text{乗値} = 32.385 \text{ (****)} > \chi^2(0.05) = 9.480$$

表4. 暮らし方

	合計	一生懸命に働いて金持ちになること	まじめに勉強をして名を挙げる	自分の趣味にあった暮らし方をする	その日をのんきに暮らすこと	どこまでも清く正しく暮らすこと	社会のために全てを捧げて暮らすこと	その他
訪問留置調査	289	7.3	3.1	48.8	34.6	2.8	2.4	1.0
ハイパーシステム調査	612	14.4	4.9	48.5	11.4	3.3	2.9	14.5

$$\chi^2 \text{乗値} = 100.895 \text{ (****)} > \chi^2(0.05) = 12.592$$

表5. 最終学歴

	合計	中学卒	高校卒	高専・専修 学校卒	短大卒	大学卒	大学院卒	その他
訪問留置調査	875	5.7	37.7	14.9	8.7	30.9	1.7	0.5
ハイパーシステム調査	612	3.9	32.5	10.8	12.4	36.3	1.6	2.5

χ^2 乗値=28.914 (****) > $\chi^2(0.05)=12.592$

即ち、20代が少なく、3, 40代が多い。学歴も短大卒以上が多く、暮らし方では、「のんき」よりは「一生懸命働き金持ち」が多い。そして支持政党は「民主・社民寄り」で、最もズレが大きい情報公開では「全くそう思う」の強い賛成意見に偏る。

5. 結果の精度

こうした偏りを持つパネルの調査結果ではあるが、確定数値が得られる投票結果とつきあわせると今回都知事選挙結果とは、表6に示す通り、有意差アリとはなっていない。

表6. 東京都知事選挙

ハイパーシステム調査：4月10日調査（投票行動予定及び不在者投票の合計）

	合計	石原 慎太郎	はとやま 邦夫	ますぞえ 要一	明石 康	三上 満	その他
投票結果	5,463,768	30.5	15.6	15.3	12.6	12.1	13.9
ハイパーシステム調査	131	35.1	16.0	12.2	12.2	8.4	16.0

χ^2 乗値=3.701 < $\chi^2(0.05)=11.070$

しかしながら、石原の5%のズレは大きい。補正計算を行った結果は表7で、この差は縮まっている。

表7. 東京都知事選挙（補正計算後）

	合計	石原 慎太郎	はとやま 邦夫	ますぞえ 要一	明石 康	三上 満	その他
投票結果	5,463,768	30.5	15.6	15.3	12.6	12.1	13.9
ハイパーシステム調査	3780	34.7	16.3	11.3	13.0	9.2	15.6

なお、今回選挙では毎日新聞がデータを公表している。残念ながら、それらは電話調査結果であり、適合の基準とするには問題があるが、一応検定をかけた所、同じ質問のあった「支持政党」「前回知事選での投票者」で、有意差は現れていない。「投票予定」では「必ず行く」が毎日調査で65.7%、ハイパー調査では52.9%と差があるが、現実結果とは、ハイパー調査の方が近似している。

6. 結論として

筆者らは、このオンラインリサーチシステムの開発当事者であるので、この手法の有用性判断は客観的な他者の評価にまかせるべきと考えている。その判断評価の一助に若干の情報をつけ加える。その一つは、試験測定項目への反応比率が、ある社の専用となっている個人パネルでは、毎回異なるパネル群に於いても、安定していること。第2には、回答返信が、早朝・深夜に数多く、これらは訪問面接調査での回収不可能者から回答を得ていると見られること。第3には、パネルメンバーからの反応によると、手際よく、ゲーム感覚に近い気軽さで応答できること。そして、この気軽さが無責任回答ではないことが、回答の斉合性から信じられること、第4には、画面の冷静さと人間要素の不介在が、冷静かつ忌憚のない回答を可能としていると見られること、第5には、回答者を回答時刻制約から解放していること。そして、第6には、コストが通常の留置法に比し、(質問数の差は大きい) 1/2以下となること、第7には、回答を保存できるため、多量質問調査、同時実施ではバイアスを生ずる質問のある調査は、同一対象に分割して、日を改め実施できること等がある。勿論、欠点もある。その1は、二相サンプリングとなること、2は、世帯パネルの場合、世帯人数により確率が変わる、拡大誤差の程度が明確でないこと、3は、質問文の長さ、量の限界等である。我々の今後の課題は、この欠点について研究を進め、より科学性を高め、既定の科学的調査法により近接することにある。