

# インターネット調査の何が問題か

— 現状の問題と解決すべきこと —

大学共同利用機関法人 情報・システム研究機構  
統計数理研究所・名誉教授 大 隅 昇

## 1. まえがき

適切な調査環境の確保・維持が次第に難しくなる中で、新たな調査法として登場したインターネットを利用する調査（インターネット調査、ネット・リサーチ）の勢いは益々高まっている。調査が商品化されてから久しいがこの種の調査が登場したことで、これが加速されている。しかもその信頼性や妥当性に疑問があるとされながら、その諸特性の科学的な検証が十分に行われないまま、利便性（迅速、廉価など）だけが強調されてきた。確かに、調査とくに市場調査の分野では、迅速性や廉価性が重要であることは理解できる。また、調査を利用するクライアント側の理解や使い方に大きな変化が現れていることも事実である。しかし、このことと調査の信頼性の確保をどう考えるかは異なる次元の話である。

そもそもインターネット調査（Internet survey）とはどのような調査法であろうか。あるいは、巷間、インターネット・リサーチ（ネット・リサーチ、ネット調査）という言葉が氾濫しているが、これらは一体どのような調査法なのだろうか。ここでは曖昧に「調査法」と書いたが、果たして調査法をどう考えるべきなのか、様々な要素が錯綜して簡単には記述できないし実際にそのような状況にある。

昔見た映画「先生のお気に入り」で、記者養成学校の教師が生徒に対して「新聞記事を書くうえで、もっとも重要な項目を列記せよ」と

と問いかけをする。これに対してヒロインが、その答えは「Exact, exact, and exact」と言うシーンがあったと記憶する。記事にとって重要な要件は何をさておき「事実を正確に」ということであろう。今のインターネットを介した調査に欠落するものは、まさにこの「正確さと透明性」である。現状、この種の調査の実態は、何がどう行われているものかほとんど知ることができない。一部でこの種の調査環境の改善努力がなされつつあるが、やはり不透明な要素が多いのである。しかも、従来型調査の環境も次第に悪化する中であってここで調査環境全体を見直す時期にあると考えられる。このようなとき本誌への執筆の機会をいただいたので、この分野について筆者が考えるところを、産学協同研究実験調査から得た経験に照らし、紙幅のゆるす範囲で述べてみたい。

## 2. 簡単な試み

原稿執筆に先立ち簡単な実験を試みた。それは、ネット・リサーチを行っているといういくつかのサイトのホームページにアクセスし、「偽名で登録が可能か」を試みることである。

そもそも、この種の登録が自由にできることが、インターネット・リサーチの特徴であり問題点の一つとされてきた。最近はこれも巧みに防止・回避出来るとの意見もあるので、実際にどういう事情にあるかを試みたわけである。結果は、いずれも登録可能であり、そ

れぞれ登録済みの通知を得た。これはいわゆる自己参加型 (self - selection) 「この指とまれ方式」のケースである。この方式を採用する調査サイトが多いが、ここには下記のようなキャッチコピーが沢山見られる。

- ・ モニター登録者数が多い (よって代表性がある)
- ・ 登録者の管理を徹底している
- ・ モニター品質、回答結果の品質を保証できる
- ・ 多数の登録項目を取得し徹底した管理を行うので情報が豊か
- ・ 個人情報の保護と守秘も徹底できる

等々である。しかしいずれも科学性のない主張にみえ、またこの種のネット・リサーチはマーケティング活動のツールではあっても調査とはいえないのではなかろうか。

ここで、郵送、訪問留置などの従来型調査にも「代理回答」があったとの反論が予想される (実際にある)。しかしインターネット・リサーチと従来型の違いの一つは、その調査対象者の「選び方」とその「同定性」にある。郵送法や訪問留置法では、多くの場合は“誰を選んだか”が分かっている。一方、インターネット・リサーチでは、前述の登録実験にみるように果たして正しいのかが分からない。一見すると従来型調査と“同じ事象、現象に見えること”が、状況が変わって“別の事象として現れる”のがインターネットを利用した調査の特徴である。とくに、集めた集団の特性を記述するだけでは十分ではなく、「どのように集めたか」が欠落していることが問題である。このような簡単な例だけでなく、無数の未解決の課題があり、その多くが曖昧である。

さらに反論として、回答者全員に謝礼を提供し登録者情報 (例えば銀行口座情報) を正

確に確保すれば、登録者の正確な同定が可能との主張もあるようだ。しかしこれも登録者の家族や知人の名前を利用すれば容易にクリアできる。どのような方式を採用しても「登録者集団 (リソース、パネル)」の質を確保するためには、従来型の調査と同様に、登録者の調査への協力に頼るしかない。しかも匿名や「なりすまし」が見かけ上回避できたようにみえても、回答者の顔が見えないという点で従来型調査以上に事が複雑となっている。

このように、いかに高度な情報処理技術を駆使しても「誰を、どのように選び、確実に捕捉したか、どのように回答したか」という調査の基本事項が満たされないかぎりには空理空論である。またこれらに関する議論を避けては通れないのだが、実際はここらの情報は皆無にちかい。

しかし、ネット・リサーチ専門社のサイトをみると、さすがに数年前のような粗雑な情報は少なくなり、少しは意味のある情報が見られる。例えば、サイト登録者の属性情報 (性、年齢、職業、インターネット利用環境など) を部分的に公開するなどである。だが、前述のように、その種の情報の前に必要となる「どう集めた登録者であるのか」の情報記述がほとんどない。つまり、依然として問題の本質への完全な回答とはなっていないのである。一方、従来型の調査機関・調査企業が行うインターネット型調査の紹介を見ると、開示内容の密度に温度差はあるものの、用いた調査方法、方式の記述等がある。また、調査対象者の捕捉・選び方、調査方式の範囲等の情報を掲載しようという努力の跡が見られる。このように両者は対照的であり、このことが正に「今のインターネット調査の複雑な状況」を映しているといえる。

### 3. 電子的調査情報取得法への移行 調査法 研究の一つの視点

いわゆる「電子調査」(electronic survey)のはしりが登場したのは1990年代の後半からである。そのきっかけは、おそらくはWWW環境の登場とその周辺機能の技術改善と基盤整備に因るものである(ここの歴史的変遷については大隅[64])。また電子調査の呼称はインターネットの初期の頃から多用されている。ここには今のインターネット調査(Web調査、電子メール調査)の呼称の他、ファクシミリ調査やDBM(disk-by-mail;フロッピーに質問票を書き入れ郵送で授受を行う方式)などのハイブリッドな方式も含まれていた。

いずれにせよ調査環境の変化を考えたとき、もっとも大きな変化は、情報技術の普及により“データ取得機構”(data acquisition system)が急速に様変わりしたことである。つまり一般に言われる電子的調査情報取得法:CASIC(Computer Assisted Survey Information Collection)やコンピュータ支援によるデータ収集:CADAC(Computer Assisted Data Collection)を前提とした新たな調査方式の整備検討が必要となってきた。同時に、伝統的な調査方法論と、これらCASIC、CADACとの相互関係を比較し、それぞれの特性を客観的に考察せねばならない。従来の考え方で説明できない要素については新たな概念なり知識を導入し、また伝統的な調査法の豊富な経験や知識を活かすことが求められる。CASIC、CADACという言葉から類推されるように、これは電子的に調査情報を取得するための「調査方式あるいは調査モード」(survey mode)の総称である。ここで「調査方式」とは、調査データをいかに取得するかデータ収集方式(data collection mode)

をいう。例えば、従来型調査であれば面接調査、電話調査、郵送調査、...などを用いたということに相当する。つまり、調査対象を誰とし、それをどう規定するかという標本抽出や標本などの操作、これは「標本調査法」とは少し意味の異なることである(不可分の関係にあるが、違う要素について考えている)。

筆者は、多くの議論は“この両者を混同していること”に一因があると考えてきた。例えば「標本抽出枠として住民基本台帳を用いて層化二段抽出法による無作為抽出で選んだ調査対象者に、訪問面接法により面接員が調査を行った」という言い方がよくなされる。このとき標本調査法に関わる事項と調査方式のそれとが混在している。この記述の前半は標本調査法に関わることであり、この例では統計的な科学性をもって処理されたことを意味し、後者の面接員による訪問面接法とは用いた調査方式を指している。

この観点から、“現状のネット・リサーチ”を眺めると、多くの場合、調査方式のありようを技術面からあれこれと議論し、それをもって調査の正当性、信頼性を主張しているように見える。必然的に技術論に議論が集中しがちである。いわく、迅速で納期が早い、回答制御が可能で無回答がない(?)、均質でゴミが少ない(本当か?)、...等々である。

ここで重要なこととして、調査法研究をどう考えるかがある。調査法研究(survey research)とくに標本調査法(sample survey, sampling survey methods)は調査設計の基礎的事項に関わる研究をいう。たとえば、母集団の構成(対象母集団、枠母集団あるいは標本抽出枠、代表母集団)の検討に始まり、標本抽出法(サンプリング)の設計、ここでは確率的サンプリング技法(無作為、系統、確率比例、層化比例など)と、非確率的サンブ

リング技法（クオータ法、スノーボール・サンプリング、インターセプト法など）の基礎研究が必要である。

一方、調査方式（調査モードあるいはデータ収集方式）とは、“いかにデータを集めるかの方式”、その技術面をいう。面接法、郵送法、訪問留置法、電話調査、インターネット調査（電子メール調査、Web調査）等々はこの調査方式の種類をいう。

日本国内では、住民基本台帳や選挙人名簿などの優れた標本抽出枠が利用できたことから、ほとんど理想に近い確率的な方式（probability-based）が適用できた。もちろん市場調査などでは、インターセプト法あるいはCLT（会場テスト方式）、グループ・インタビューなどが多用されており、これらは定量的な調査に対する定性的調査として扱われてきた。このようなことが、調査方式と標本調査法との関係をそれほど厳密に意識して考える必要がなかった理由の一つと考えられる。

しかし調査方式の多様化に伴い、標本調査法と調査方式との関係が曖昧かつ複雑になってきたことがある。前述のように両者は不可分の関係にあるが、調査研究を進める上では、あえて意識的に分けて考えることが適切である。日本国内では、インターネット調査の分類を、「オー

ブン型」「クローズド型」といった調査実施上の形式で分けて考えることが多いようだが、既にこのことに疑問がある（参照：「実践ネット・リサーチ」、[71]）。ここには調査をどう考えるかの点で重要な調査対象者の選び方（抽出方法）についての含意が欠落している。筆者は、調査方式を現状のデータ取得環境の変化に合わせて、以下のように整理するのが適切と考えている。

- ・ データ取得機構が電子的か否か、つまりコンピュータ支援か（CA：computer-assisted）否か
- ・ 記録方式が自記式か（self-administered）、面接員などによる他記式か

こうして表1のように分類区分すると調査方式（調査モード）の理解が容易となる（米国の諸研究、とくにミシガン大学派の Groves, Couper, Vehovar 等の研究他をサーベイし、また我々の実験調査成果等に基づいて要約した、[3]、[11]、[14]、[44]）。

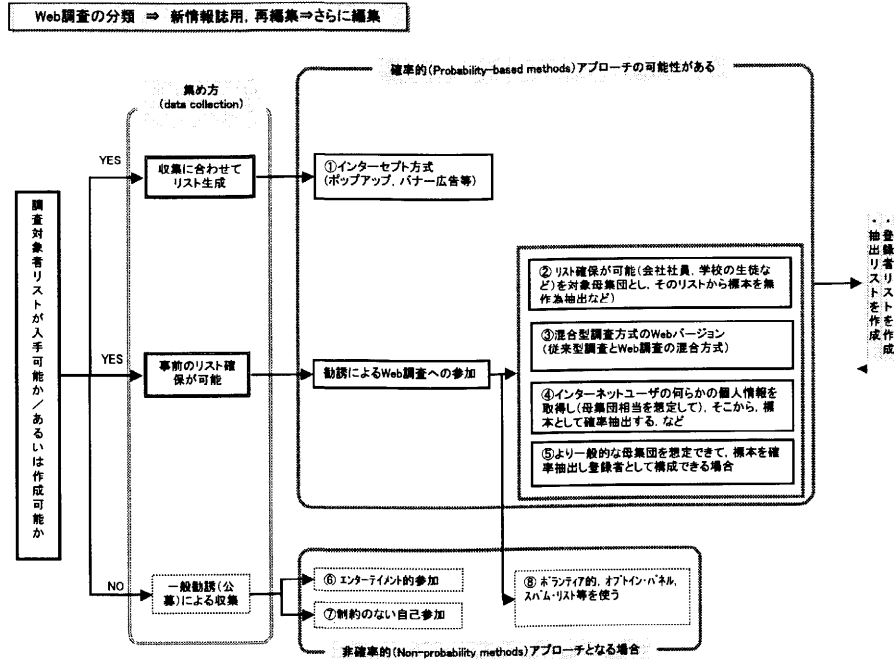
なおここで「面接」とは、訓練された面接員を介して記録（回答者からの回答取得）が行われることを言う。重要な事は、回答取得が回答者自身によるものか（自記式）、他の人（面接員）を介して行われるものか（面接員による記録、他記式）にある。

表1 調査方式（調査モード）の分類区分

		コンピュータ支援の有無（CA：computer assisted）			
		なし		CASIC, CADAC	
		面接員方式の有無		面接員方式の有無	
調査方式 (調査モード)	面接	あり 面接	自記式 訪問留置	あり (CAI) CAPI	自記式 CASI
	郵送	-	郵送	-	DBM (Disk by Mail)
	電話	電話	Facsimile	CATI	-
	インターネット	-	-	OFG	電子メール, Web調査

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

図1 データ収集方式と確率的・非確率的アプローチの関係



#### 4. 確率的アプローチと非確率的アプローチからみることの重要性

前述のように、調査法を考える一つの視点として、調査対象者の「選び方」と用いる「調査方式」とを分けて考えよう。ここで選び方を別の言い方でいうなら、調査対象者の確保の方法あるいはその具体的なリストが存在するか否かで分けられる。ここで重要なキーワードは、集める調査対象者について“統計的あるいは確率的な推論が可能か否か”である。Couper ([3] [10]) はこれに対して、あるヒントを提供している。彼の考え方を若干脚色して筆者なりに要約した(図1)。

そもそも、確率的アプローチと非確率的アプローチとを混乱したままでは議論できない。また仮に同じ調査対象を相手としても、調査方式(モード)が異なれば回収結果(内容)は異なると考えるのが妥当である。確率的アプローチと非確率的アプローチとは異なると考

えるならば、同じ質問項目を用いて、異なる調査方式で行った結果をそのまま比較することにはかなり無理があると考えべきである。要は、確率的アプローチの適用可能性の有無を知ることであり、それを意識して議論することでインターネット調査の内容はまったく異なるものとなるということである。

このことはインターネット調査における調査対象者(標本、サンプルと言いたいが誤解を招くのであえて「調査対象者」とする)の選び方あるいは抽出元となるリスト作成方法と、実際に用いる調査方式とをなるべく区分して考えるべきことを示唆している。これは後述するように、登録者集団の構築時の基礎情報となるものである。現状のインターネットを介して行う調査は、こうした枠組みから再考察することが求められているのである。

## 5. 現状をどう考えるか

筆者が見聞しました実験調査の体験から得た情報として、インターネット調査に参入している調査機関・企業は以下のように大別される。

ネット・リサーチ専業社グループ  
従来型調査の経験を持つ調査機関や企業で、インターネット調査業務に参入したグループ

に分かれるであろう。後者はさらに、

- 1 従来型調査の方法論の延長線上にまったく新たなインターネット調査システムを構築しようとするグループ
- 2 いわゆるネット・リサーチの仕組みをそのまま導入してきたグループ（にほとんど同じ方式を採用）

この両グループのインターネット調査への理解や調査実施の機構構築（調査システム構築）の指針・考え方には大きな差異があることは看過できない。つまり、インターネットを使った調査であれば何もかも同じとはならないところに現状の混沌の原因がある。

ここで - 1のグループに入る機関・企業は、積極的に基盤整備を進めており、すでにならかなり高い品質のインターネット調査を提供できる環境を整えつつある。しかし、こうした一歩先んじた環境であっても、さらなる研究的要素、未解決の課題が無数にあり、必ずしも完成度の高い調査方式が提供できるわけではない。同時に、新たな調査システム構築への投資効果の策定も看過できないわけで、インターネット調査の特徴とされる迅速性や効率性は確保できても、質の高い調査が廉価で提供できるかという微妙な問題である。これについては、12節で改めて述べる。

既に読者は気づかれただろうが「インターネット・リサーチ、ネット・リサーチ」と「イ

ンターネット調査」と言葉を微妙に使い分けたことには理由がある。我々の産学協同研究の体験から得た情報であるが、そもそもこの分野での言葉使いには「混乱がある」のである。典型例が、インターネット・リサーチ（ネット・リサーチ、ネット調査）でありインターネット調査、オンライン調査である。1996年頃に、この分野の研究を開始した頃は、我々も模索の中にあり、この種の電子的手段を介して（調査）データ取得を行う調査方式を一体、諸外国ではどう呼称し定義しているのだろうかというきわめて基本的なことに疑問があった。

一方、日本国内では「インターネット・リサーチ」あるいは「ネット・リサーチ、ネット調査」（正に和製英語で浅薄な表現）で呼ばれることが多い。しかもその定義も明らかではないのはいかにも落ち着かない。たかが言葉の問題と看過できないことでもある。これは、この分野の研究が未成熟で十分ではないことの証しでもある。

調査法研究と調査方式の考え方については既に若干述べたが、ここでさらに用語の混乱について触れておきたい。用語のことなどどうでもよい、適当に好きなように使えばよい、というのが世間の最近の風潮のように見えるし、確かにこれをどう考えるかは自由なことである。しかし言葉の意味の正確さや使い方に鈍感になってきていることも事実である。このようなことで、“インターネット環境下で行う調査”も、さまざまな言いようがある。ちなみに便利な世の中であるから、検索エンジンを使ってWeb上でこの言葉（Internet research）をサーチしてみるとよい。意外なことに日本国内は別として「調査法（研究）」とくに「統計的調査法研究」にあまり関係ないことが見えてくる。実はInternet research

で現れる多くのコンテンツは「研究」にインターネットをどう活用するか、といった内容が多い。この事情は、アマゾンドットコム等で専門書を検索してみても同様である。統計学、あるいはすこし間口を広げて、社会科学系、人文科学系などの専門書をもても、そこで使われる用語の多くは以下のようなものである。

電子調査 (Electronic survey)

インターネット調査、インターネット・サーベイ (Internet survey)

Web調査 (Web - based survey, Web survey)、電子メール調査 (E - mail survey)

オンライン調査 (Online survey)

オンライン定性調査 (Online qualitative research)

このようなことで、インターネット環境下で行う調査のことは、一般にはインターネット調査あるいはオンライン調査というのが適切である。「インターネット調査」と呼称する理由は、新たな調査方式 (調査モード) として、また調査法 (survey method) あるいは標本調査法 (sample survey) との関係で混乱なく議論したいことにある。また調査方式として統計的推論あるいは確率的アプローチの可能性がどこまであるのか、またそれが無理であって現状では非確率的なアプローチにとどまらざるをえないかを明らかにする意味もある。さらに重要なことは、調査対象者が誰であり、どう選ぶかという、基本的な問題である「登録者集団 (株母集団に代わるもの)」とを併せて考えることがあるからである (後述)。

筆者等がこうした指摘を繰り返し行ってきたことの影響が、最近ではインターネット・リサーチ、ネット・リサーチを用いず「Web調

査」と言ってみたり、いつの間にか「オンライン調査」(online survey)、「オンライン・リサーチ」などと言い換えるサイトも出てきた。

なお市場調査業界では、インターネット調査に関する調査綱領やガイドラインを策定する動きがある。例えば最近のESOMARのサイトを閲覧すると、「ESOMAR Guideline on Conducting Marketing and Opinion Research Using the Internet」(<http://www.esomar.org/esomar/show/id=104178>)として開示情報がある。これを通読すると終わりの方で「インターネット・リサーチの定義 (Definition of Internet Research)」とあるが、ここではオンラインを通じて行う調査行為 (オンライン経由の調査票を用い、オンライン上での面接方式を採る) ともしっかり書かれている。

要は言葉の表層的な意味合い、語呂合わせの問題ではなく本質はどこにあるかである。ともかくもネット・リサーチ、ネット調査とは言わないのである。

さて、インターネット調査とは、単純には「電子メール調査」(EMS: Electronic mail survey) と「Web調査」(Web - based survey, Web survey) に分けることで十分である。かつては様々な試みがあったが、現在はこの区分で説明できる。もちろん、今後の技術的な進歩改善につれて、新たな電子的な調査情報取得方式が登場するであろう。例えば、高機能の携帯電話、モバイル端末、双方向デジタルTV、IVR (Interactive Voice Response) 装置等、様々な調査方式や機器類が現れるであろうが、いずれも電子的調査情報取得の調査方式を提供するものという点では同じである。

[ インターネット調査とは ]

電子的データ取得法 ( CASIC , CADAC ) の一つの調査方式であって、ネットワークや WWW 環境を用いたインターネット調査システム下で実施される調査 ( Web 調査、電子メール調査 ) のこと

“ 主として ” インターネット・ユーザを調査対象者とする

コンピュータ支援の自記式調査 ( Computer - assisted self - administered ) である

とくに Web 調査では、ブラウザで閲覧する形式の電子調査票を用いること ( HTML , XML , Perl , C 言語などの言語で記述 )

必要に応じてマルチメディア対応の機能 ( 音声、動画・静止画などのイメージなど ) が利用した表現の豊かな調査

票設計が可能

回答取得過程がネットワーク上で双方向にリアルタイムに転送記録される

ここで第 2 項で “ 主として ” とした理由がある。実は、調査対象者をインターネット・ユーザに限定せずに、従来型調査と同じような標本抽出枠、たとえば住民基本台帳から選んだ対象者に対して、特別に開発した調査用電子デバイスを配布あるいは一定期間貸与して、このデバイスを介してインターネットで調査を行うという方式がある。筆者の知る限り、国内では既に数社がこのような方式を採用している ( 12 節参照 )。

インターネット調査とは、以上の仕組みを利用するための電子調査システムの基盤整備を始め、登録者集団 ( リソース、パネル ) 構築と運用管理、電子調査票や設問の設計、サンプリング操作、調査依頼から回収、データ処理・集計分析、調査報告など、調査の全過

程にわたって科学的な調査方法論を適用して行う総合的な活動をいう。

## 6. ビジネスとしてのインターネット・リサーチ 問題の本質からそれた議論や考察

科学的な調査法研究は関心のある研究者が行うべきもの、ビジネスとは別と決めつけることはあまりに乱暴である。なぜなら、調査法研究においてもビジネスとして成立するための位置づけ、その諸要件の考察が必要なこととは言うまでもないからである ( とくに、調査品質と調査経費・コストのトレードオフについての実証研究 )。しかも、ビジネスであるから利益優先であるべきを建前として、廉価、迅速のみをうたうことはいかにも稚拙な発想である。ビジネス・モデルとして利益につながればそれでよしと主張されても、研究者の立場からはその後続く言葉を失うだけであるが、そのような風潮も見えるのである。

インターネット調査についての日本国内の議論は、欧米とくに米国のように、研究と応用・実用の両面から産学の多くの人々が参加して研究や実用化の検討を進めてきたという過程がほとんどない。なぜか、ネット・リサーチの呼称のもとに、ある時期に唐突に登場したと思われることである。ベンチャービジネスの一つとして、調査 ( 市場調査、社会調査など ) とはあまり縁のない、新興のネット・リサーチ専門として登場した。

一方、欧米では、繰り返し述べるように電子的なデータ取得法、調査方式として、すでに 1990 年代半ばから、電子メールを用いた調査や DBM ( disk by mail ) 方式のようにフロッピーディスクに質問票を書き入れ郵送で授受を行う折衷方式などまで利用されていた ([ 14 ], [ 43 ])。併せて電話調査法とくに、1970 年代から始まったコンピュータ支援の電



話調査法やRDD方式の普及に連動した諸研究のノウハウの実績がある([23])。長い試行錯誤、実験研究の延長線上に、現在のインターネット調査の世界が拓かれているのである。盲目的な欧米迎合は歓迎すべきことではないが、また国内の従来の調査環境事情を十分に反映させたインターネット調査研究が望ましいのであるが、そのレベルにはとても至っていない。

では現状はどうか。承知のうえのことであるか否かに関わりなく情報が「不正確」であることには変わりはない。換言すると、そもそも従来型調査法との直接の比較には無理がある。その理由は、根本的に“調査方式の差異”があるからである。例えるなら、同じ調査票や質問を用いたとしても、多くの場合郵送調査と面接調査の結果が同じとなるとは限らない。しかし、ネット・リサーチの世界では、十分な論証もなく従来型調査に比して優れているとの主張が見られる。例えば、その簡便性や迅速性を主張する、高品質をうたう等のキャッチコピーが依然として氾濫し、以下のような用語・表現は雲霞のごとくみられる。最近では巧妙になって、あたかも統計的手法あるいは確率的アプローチが適用できるかのような表現を用い、その正当性や信頼性を主張しようとする危険な傾向さえ見られ、こうなると倫理観の存在さえ疑わしいきわどい行為にさえみえる。たとえば、

- ・ 母集団（あるいは枠母集団）を的確に作ったので問題がない
- ・ 登録者数がきわめて多く、また正確な登録情報を基に正しいサンプリングを行っている
- ・ 精度のよい調査である
- ・ 偏りのない、また補正を行った高品質の調査である

・ 自由回答の取得が容易である（回答者の本音が聞ける、情報量が多いなど）  
等々である。こうした実証のない乱雑な情報に目くじらをたてても、依拠するところも考え方も異なる中では議論は平行線のままである。実際に、ある企業団体系グループ主催のシンポジウムにおいて、Web調査と他の従来型調査の比較調査を行ったという結果を聞く機会を得たが、仰天の一語に尽きる内容に啞然とした。調査方式の異なる調査結果を単純に比較し、結果が合ったから類似している、合わなくても補正を行えば適正な結果となるというあまりに粗雑な論理展開に、幾つかの質問、疑問を呈示したが未だに回答を得ていない。

また最近の傾向として、自由回答設問を多用しこれとテキスト・マイニング技法を用いることで、回答者の「本音」がたちどころに分析できるといった意見も多い。しかし、これとて調査方式間の類似や差異を検証し、どのような場面でインターネットを使うことが優位であるのかの検証が必要である。筆者等の実験調査の結果では、確かにインターネット調査は見かけ上の書き込み量（率）が多く情報が多いように見えるが、その質的な内容や、回答行動にはいろいろな問題もある（回答所要時間、書き込み量の関係などの分析が必要）。当然のことながら、設問内容にも依存する。電子的に自由回答がテキスト情報として迅速に取得できるようになったという利点はあっても、自由回答の設計や分析法等の従来からの問題に加え、新たにWeb調査特有の測定誤差が発生し、それゆえ未解決の課題が多いのである。

## 7. インターネット調査の何が問題か？

インターネット環境を用いた調査に対して、

次のような立場があるようにみえる。

- ・ かなり問題がある調査方法であって、しかも良く分からない不透明な要素が多々あると考える人たち
- ・ 手軽に安い経費で調査らしい情報が簡単に得られるので便利と考える人たち

実際、最近の新聞紙面や週刊誌他の情報誌には、至る所にネット・リサーチの結果と称した記事が散見される。また「1000名に聞きました」とか「784名が回答しました」と見出しがあって、それがあたかも一般性・代表性をもった回答意見であるかのように解説・説明がなされた例が多数みられる。しかしここで注意してみると、

- ・ 誰を捕捉して、誰に対して行った調査であるのか（登録者集団の構成、特性等）
- ・ 計画標本をどのように確定し作ったのか
- ・ どのような回収標本であったのか
- ・ 発信数あるいは計画標本と回収標本・回収数のずれは
- ・ 有効回答数（有効回答率）の定義と算定方法は

など、調査としての基礎情報の記載が「まったくなくない」のが特徴である。とくに従来型調査では必須情報とされてきた「枠母集団あるいは標本抽出枠（サンプリング・フレーム）」や「調査不能理由」等の重要な情報の記載は一切ない。つまり、

- ・ 調査対象者リストをどう作ったか、あるいは標本抽出枠はどう作成したか？
- ・ その作り方が明示的で透明か、あるいは公募型のようにまったくの依存型であるのか
- ・ 標本抽出枠、計画標本、回収標本間の関連情報、等々

がまったくなくない。

では、これを人々はどのように見ているのだろうか。

か。「あやしい調査」と考える向きは、調査も質が落ちたものだ、何をいかがわしいことをやっているのかと苦々しく思い、片や楽観派あるいは無関心・無理解派は、データ取得の仕組みや質のことには目をつむり、ネット・リサーチは安かろう悪かろうかもしれないが、便利に使いこなせばよかろう、リスクの大きい意志決定などには用いず参考情報と思えばよかろう、といった軽い気持ちで重宝する。これら両派が混在し、しかも調査法としての科学性に立脚した十分な議論は避けたままに過ぎてきたのが現状である。こうした実状を見て「このままでは、今後の調査は一体どうなってしまうのか」と強い危機感を持って立ち向かうという風潮は、ごく一部をのぞいては一向にみえない。むしろ、ネット・リサーチをビジネス・モデルの成功例としてもてはやすというのが世の流れのようである。

しかし調査環境はこのまま、インターネットによる電子的調査の方向に無為無策、無節操のまま流されてしまってよいのだろうか。筆者がインターネット調査という「調査方式」があるということを初めて知った1990年代の半ば頃から、ずっとこうした疑問を持ち続けてきた。とくに1990年代半ば、今のWeb利用環境の普及のきっかけとなる様々な技術革新が登場した時期に、こうした疑問を持ったことが、1996年以降、「調査環境の変化に対応した新たな調査法の研究」の課題の下に継続的に実験調査を続けてきた理由である。

これらの経緯の詳細は、過去に行ったさまざまな報告（参考文献に挙げた多くの報告）を見ていただくこととし、ここでは現状のインターネット環境下での調査は「何が問題か」を要約してみたい。無数にある課題のうち、もっとも問題とされる項目を列記してみよう。

（1）回答者（回収標本）はインターネット・

ユーザを代表するのであろうか。

- (2) 回答者は、登録者集団（リソース、パネル）やそれに基づいて作られた計画標本を代表するのであろうか。あるいは、そもそも登録者をどう集めているのか。
- (3) インターネット・ユーザおよび回答者は一般の人（通常の従来型調査における調査対象者）を代表するのであろうか。
- (4) 今後、インターネット・ユーザが増えたとき、回答者は一般の人に近づくのであろうか。
- (5) 同様に、登録者集団の大きさ（登録者数）が大きいことが、その調査結果の品質が高いことを保証するのであろうか。
- (6) 回答者（回収標本）に特別な回答者、いわゆるプロ回答者集団（professional respondents）を含む恐れがないのだろうか。
- (7) 「世論調査」に用いることは可能であろうか。

ここに挙げた項目は、社会調査に関心がありそれに関わってきた人なら誰もが疑問に思う事柄である。しかし、現状のインターネットを介した調査では、これらに適切な解答を与えることはかなり難しいことでもある。結論から申せば、これらはいずれも「否あるいはほぼ否」と考えざるを得ないのだが、しかしそれでもなおインターネット調査を少しでも信頼に足る調査方式として確立するための「指針を得たい」という期待を持ちたい。これが産学協同の実験調査研究の目標でもあった。

これら各事項について、筆者等が行ってきた実験調査の結果から、これらの疑問への解釈・理解を述べてみたいが、それとて限られた実験結果であってすべての疑問が解消するわけではない。つまり、それぞれについて「ど

んな問題があるか」を指摘し実証することはできるが、では「具体的な解決策をどう得るのか」までには至らぬことが多い。しかし、少なくとも「何が問題であって、どこから解決すべきかの道筋」は見えたと考えている。問題解決への必要条件は指摘できるが、具体的かつ十分な解を得るにはさらに実証研究が必要ということである。

## 8. 調査の概要

筆者は1996年から2003年にかけて、複数調査機関の協力の下に産学協同研究として第1次～第4次実験調査を続けてきた。調査回数は、インターネット調査（主にWeb調査）が延べ36回、従来型調査（オムニバス、郵送、専用端末利用オンライン）が延べ12回、併せて48回を9機関の協力を得て行ってきた。これの詳細は、

- ・ CD - ROMとして配布・公開
- ・ 多くのシンポジウム、セミナーなどの開催
- ・ 日本マーケティング・リサーチ協会（JMRA）での情報の公開

（注：ここで、CD - ROM、資料類はJMRAのWebページから入手可能）

などで開示してきたので、詳細はこれらをご覧いただきたい（参考文献リストも参照）。ここでは一連の実験調査のうち、主に第4次実験調査の結果を引用するが、ここに示す検証結果の多くは、一連の実験調査の内容に共通することが多いのである。この第4次実験調査の概要は表2に挙げるような内容となっている。

なお表2に見るように、実験調査は“ほぼ同時期に、同じ調査票・設問内容を用いて”行った。つまり、調査実施機関（3機関、以下でA、B、C社と呼ぶ）と登録者集団の集め方（公募系、非公募系と呼んでおく）に違いがある

他は、ほとんど同一条件下で行われた調査である。インターネット調査については、3機関のうち、A社は、公募系、非公募系の2種類の登録者集団があり、B社は非公募系、C社は典型的な公募系である。

注1: ここで公募系(または公募型)、非公募系(または非公募型)とは、一般的な用語ではない。ここでは本来は前に述べた分類区分の表と図(表1、図3)に従えば、公募系=非確率的アプローチ、非公募系=確率的アプローチの可能性がある、といった対応と考えていただきたい。用語説明が煩雑になるので、本報告を通じてこうした言い方をしておく。

表2 実験調査の概要(第4次実験調査のみ)

インターネット調査(Web調査)	A社		B社	C社
	第1回調査テーマ:生活意識編			
登録者集団	A社(非公募型)	A社(公募型)	B社(非公募型)	C社(公募型)
調査方式	リソース内 サンプリング方式	リソース内 サンプリング方式	パネル方式	リソース内 サンプリング方式
調査時期	3/28/2002-4/4/2002	6/13/2002-6/20/2002	3/28/2002-4/8/2002	4/11/2002-4/25/2002
調査票のページ分割の有無	×	△(半数に実施)	×	○
設問レイアウト実験調査適用有無	○	△(半数に実施)	○	○
有効回収回答数(%)	939 (61.6)	3,392 (42.4)	931 (88.2)	725 (36.3)
	第2回調査テーマ:インターネット編			
調査時期	5/16/2002-5/23/2002	6/20/2002-6/27/2002	5/16/2002-5/27/2002	5/22/2002-6/7/2002
調査票のページ分割の有無	×	×	×	○
有効回収回答数(%)	894 (59.1)	2,587 (32.4)	896 (84.9)	691 (34.6)

オムニバス調査	A社	C社	郵送調査	B社
	第1回調査テーマ:生活意識編			生活意識+暮らし向き編
調査期間	3/22/2002-3/31/2002	5/7/2002/-5/15/2002	調査期間	7/4/2002-7/25/2002
方法	面接法	訪問留置自記式	方法	郵送
対象地域	東京30km圏	日本全国	対象地域	首都圏40km・近畿圏20km
対象者	15~59歳男女	15~79歳男女	対象者	エリアサンプリングによる自社保有パネル(15~69歳の男女)
サンプリング法	住民基本台帳からの二段無作為抽出	住民基本台帳からの層化多段無作為抽出		
計画サンプル数	1,236	2,200	計画サンプル数	1,000
回収数(%)	630 (51.0)	1,336 (60.7)	回収数(%)	946 (94.6) *
不能票数(%)	606 (49.0)	864 (39.3)	不能票数(%)	54 (5.4) *
	第2回調査テーマ:暮らし向き編			*2002/07/25現在
調査期間	5/17/2002-5/26/2002	5/7/2002-5/15/2002		
計画サンプル数	1,001	2,200		
回収数(%)	630 (62.9)	1,389 (63.1)		
不能票数(%)	371 (37.1)	811 (36.9)		

8. 1 回答者(回収標本)はインターネット・ユーザを代表するか

ここでまず問題となることは「インターネット・ユーザ」とは何かである。また、インターネット・ユーザの全体を知ることができるのかという問題もある。従来型調査であれば、センサス情報を参考とし、また住民基本台帳や選挙人名簿などの整った標本抽出枠の

利用が可能であったから、適切な確率的アプローチが可能であった。つまり、統計的な意味での母集団構成が明らかであった。しかし、インターネット・ユーザの総数も正確に把握できない中では、母集団と標本抽出の関係を知らずといった構図は期待できない。これをもって「であるからインターネット調査は怪しい」と切り捨てることは簡単である。しかし

インターネット調査の普及を考えると、従来型調査とはまったく異なる視点からこの問題にどう対処するかの具体的な方策を考えるべきである。

一方よく耳にする情報として、インターネット・ユーザ数が急増し、もはや国民の二人に一人はインターネット・ユーザであるという。情報通信白書（平成16年版、総務省）の推計では平成15年にはインターネット利用人口は7730万人を越えたとある。また、「インターネット白書2004」（インプレス）には、インターネット利用者数は6284万人強とある。またここには、インターネット利用世帯の約48%がブロードバンド対応になったともある。しかしこれらはいずれも「推計」であり、センサス情報とはまったく異なる次元の数値である（すでにこれら両者の推計には1500万人ものずれがある）。確かに、伸びが鈍化したとはいえ、今後も確実にインターネット・ユーザは増えるであろう。しかし、一部のネット・リサーチ専門グループが主張するように、インターネット・ユーザが増えたから調査結果の代表性が高まるという論理展開はあまりに乱暴であり、それを言うなら論証が必要である。

要するに、いま分かっていることは、仮にインターネット・ユーザが国民の大半を越えたとしても、依然として統計的な意味でのカバレッジ誤差は存在し、しかもその大きさは無視できないということであり、ましてや代表性を議論できるはずもないということである。今後、国民総背番号制あるいはそれに類似した制度が導入され、国民個々の同定が可能となり、しかもそれぞれのインターネット利用環境が捕捉できたとしても、おそらく正確な母集団情報は得られないと考えるべきである。

8.2 回答者は、登録者集団やそれに基づいて作られた計画標本を代表するか

これはインターネット調査を考えるうえで、もっとも重要なキーワードである。ここではまず「登録者集団」という概念を意識せねばならない。登録者集団とは、従来型の調査における「枠母集団にほぼ相当する」とことと考えればよい。しかし、そもそも目標母集団の想定が無理な中での議論となるので、筆者等はあえてこうした造語を充てているのである。つまり、目標母集団とした対象の中から実際に枠母集団として「誰を、どう選べるのか」「対象者リストが具体的にどう作れるか」に関わることである。なお、一部のネット・リサーチ専門家のサイトに、枠母集団なる言葉を使って説明を試みた例があるが、これは統計的な知識がまったくないことを自ら露呈しているようなものである。要点は明らかで、以下の課題を、いかにして、どこまで解決できるかにある。

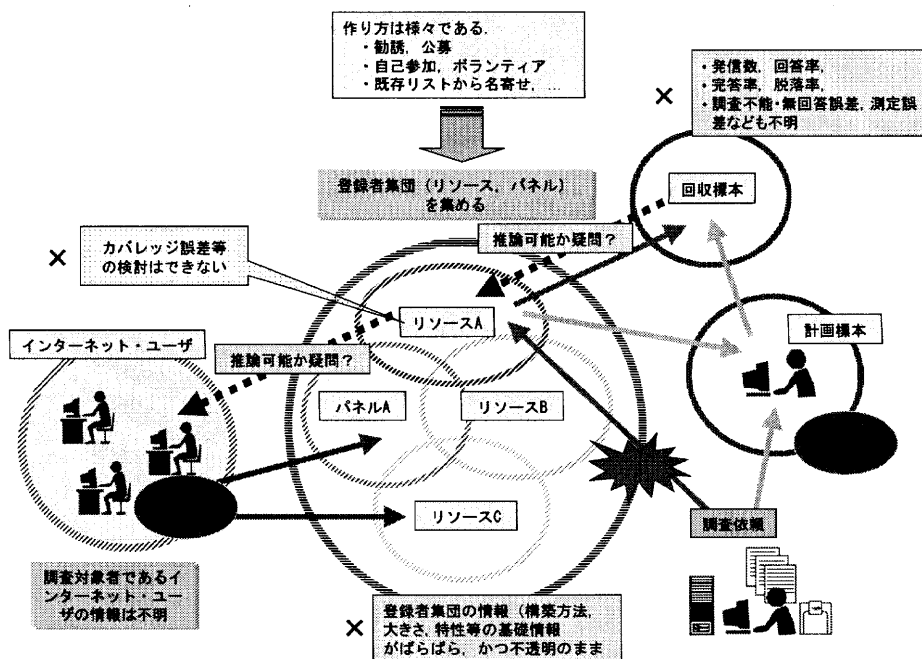
登録者集団の作り方は明示的か、つまり「どう集めたか」が明らかであるか。

計画標本の作り方は適切か、あるいは調査対象者の選び方は「登録者集団を代表」しているか、その縮図となっているか。回収標本は計画標本に合っているか、偏る場合はどのような傾向があるのか。

ともかく、現状のネット・リサーチでは、これらの情報は皆無といってよい。つまりどのような人たち（調査対象者）が、どのように登録されているかがまったく開示されることはないの、誰を調べたかが分からない。最近の一部のサイトで、集めた登録者集団や回収標本の属性情報の記述をみることがあるが、これだけでは十分ではないということである。

いま、図1に整理した内容に従って、現状のネット・リサーチを模式的に描いてみよう。図2は、いわゆるネット・リサーチにおける典型

図2 一般のインターネット・リサーチ



的な登録者集団の集め方を示す模式図である。この図にあるように、多くの場合は「この指とまれ」方式（完全な自己参加型あるいは公募型）であるのが特徴である。

では、こうした登録者の集め方であると、具体的にどのような現象が起こるのであろうか。例えば、実験調査や類似調査の結果から、以下のようなことが分かっている。

サイトによって、登録者集団の属性に大きな隔たりが生じる。例えば、性別や年齢分布に偏りが生じる。換言すると、調査運用サイトごとに登録者集団の属性が異なると考えるのが自然である。

回答者への謝礼などによる影響が出る、例えばこれが受けやすいサイトに謝礼を当てるとする主婦が多く集まるなど。

登録者集団への調査実施頻度や調査内容がよく見えないだけでなく、登録者の疲

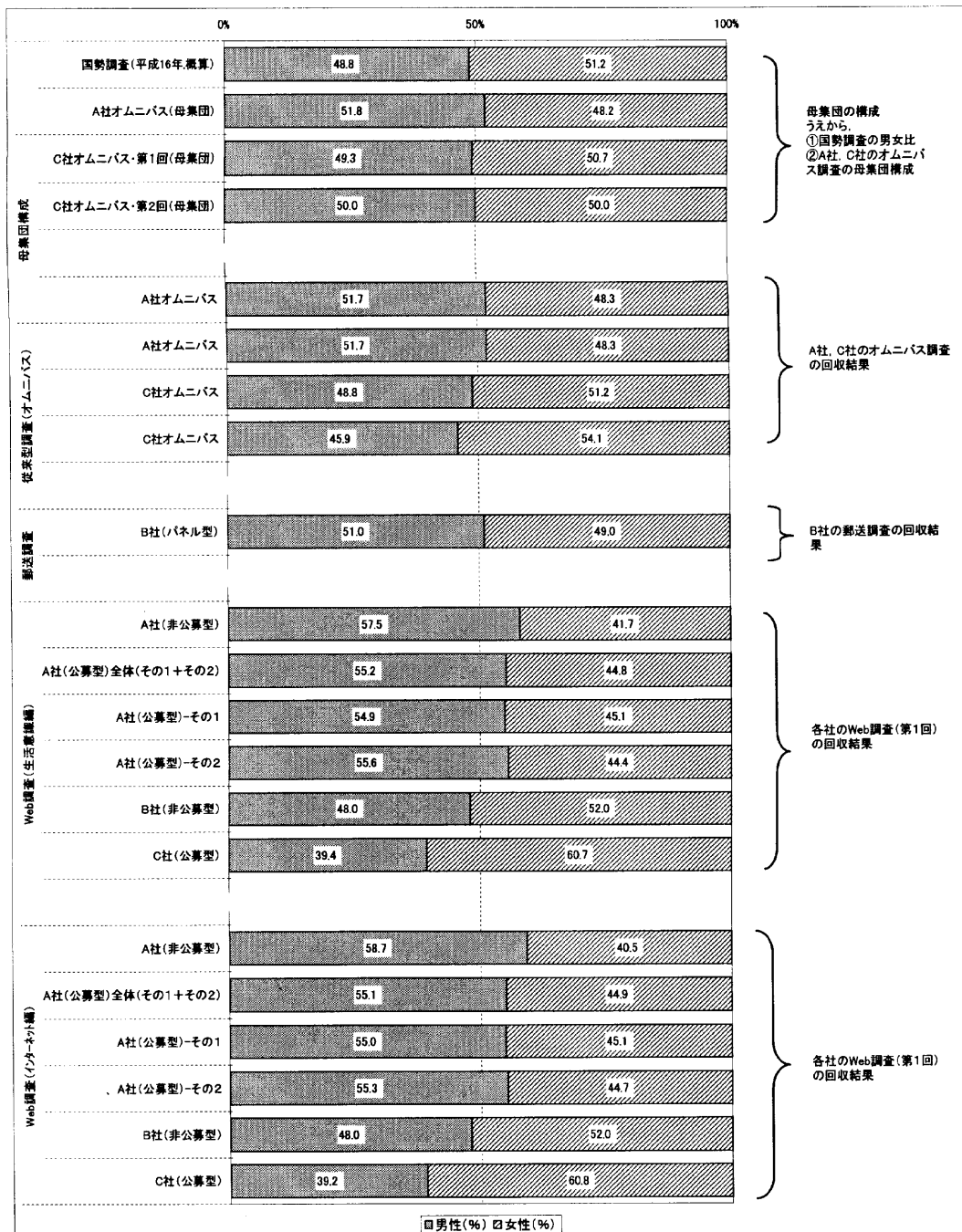
労・劣化をどう考えているか、調査のローテーション、運用管理などの情報はまったくないので闇の中であること。

こうした“偏りの例”を示すことはそう困難なことではない。一例として、実験調査に参加の各サイトの登録者集団の属性分布を比較した結果の一部をここに挙げておこう(図3)。

ここで、少なくとも従来型調査(オムニバスを2機関でそれぞれ2回、郵送を1機関で1回)の性別比率の分布はそれぞれが類似している。一方、インターネット調査については、実施サイトと公募型・非公募型とによってずいぶん差異があることが分かる。これは特別な例ではなく、実験調査すべてを通じて共通に見られた特徴である。

ちなみに、ある4つのサイトのWebページに掲載されていた類似情報を追加引用すると、表3にあるような性別比となっていた。実はこ

図3 登録者集団の属性の例（性別構成比率の比較）



ここでX社は、実験調査に協力いただいたC社であるが、図4内のC社と比較すると、調査時点が異なるのに性別比率はよく似ている。一方、Y社は性別比率がまったく逆の関係にある(男

性かなり多い)。いずれにしても、属性情報のなかでもっとも基本となる登録者集団の性別比率を取っても、既にこれほどの差異が生じるのである。

表3 4サイトからの引用情報(性別と年齢区分) (単位: %)

性別	W社	X社(C社)	Y社	Z社
男性	47.1	36.7	59.4	45.9
女性	52.9	63.0	40.6	34.1

年齢区分	W社	X社(C社)	Y社	Z社
10代	7.7	3.5	8.2	2.6
20代	30.0	30.6	27.6	28.7
30代	37.4	42.3	34.5	41.0
40代	17.4	17.6	19.3	19.8
50代	5.7	4.8	7.2	5.9
60代以上	1.8	0.9	3.2	1.9
登録者数(概数)	約25万7千人 2004年11月時点	約13万5千人 2004年10月時点	約65万3千人 2004年6月時点	約17万8千人 2004年2月時点

一方、図4の従来型調査(オムニバス、郵送)を見ると、これはその枠母集団とした標本抽出枠の情報もあるので、それを併せて記したが、これらとほぼ類似の傾向にある(当然のことであるが)。また、これらは国勢調査の結果(図の一番上)にも類似している。

### 8.3 計画標本と回収標本の関係

繰り返しになるが、実験調査は全調査回を通じて、ほぼ同時期に、同じ調査票・設問内容を用いて行った。これを念頭に計画標本と回収標本の関係を調べる。3機関のうち、A社は、公募系、非公募系の2種類の登録者集団があり、B社は非公募系、C社は典型的な公募系である。

通常の調査であれば、枠母集団に相当する標本抽出枠を決めて、適当な標本抽出法を決めれば計画標本が作られる。例えば、住民基本台帳や選挙人名簿を使ってその時点における調査対象者が計画標本として選ばれる。インターネット調査の問題は、枠母集団に相当

する抽出リストが時間軸の中で絶えず変化することにある。例えば、登録者の出入り(登録、登録抹消など)が頻繁に起こる。従って、どこかで時間を止めなければ「登録者集団の実数」は実は把握(確定)できない。実験調査では、これに対して以下のように対処している。

まず、ある時点で登録者集団を固定し、その総数を確定する。

それを標本抽出枠と考えて、必要数の計画標本を“無作為抽出”する。A社とC社にはこの方式を適用し、これを「リソース内サンプリング」と名付けた。これは登録者集団から無作為抽出を行う筆者ら独自の方式である。

B社については、登録者数が小規模であるので、全登録者を計画標本とした。

こうすると計画標本が確定できるので、結果として、登録者集団と計画標本、計画標本と回収標本との比較評価が可能となる。同時に、時間を一時的に固定したことで、



短期間とはいえ実査までの猶予期間が生じることから、登録抹消、電子メールアドレス変更等の理由で、調査依頼時点で未着が生じる（これは調査不能・無回答誤差の一種である）。

いわゆるネット・リサーチでは、ここの情報が開示されないだけでなく、そもそも登録者集団、計画標本、回収標本それぞれの関係を調べることや、その内容が開示されることはほとんどないので実態は不明である。加えて、多くの場合「30代の男性に聞きました」「20代の を使っている女性を選びました」さらに「 を所有する 名の方に聞きまし

た」といった「作為的サンプリング」や「レアサンプル対象」( 奇妙な用語である )が多いのであるが、これは筆者の考える確率的アプローチが可能な調査ではない。単なるケーススタディ( 事例検証 )あるいはインターセプト型に近い定性的調査と考えている。

さて、上の方式で得られた結果の一例を示そう。まず、表2に挙げた一連の実験調査のうち第1回実施分のインターネット調査についての計画標本と回収標本の比較を見る。次に、従来型調査( オムニバス、郵送 )について、同じくいくつかを比較する。

図4 計画標本と回収標本の関係 : A社公募系、第1回(生活意識編)

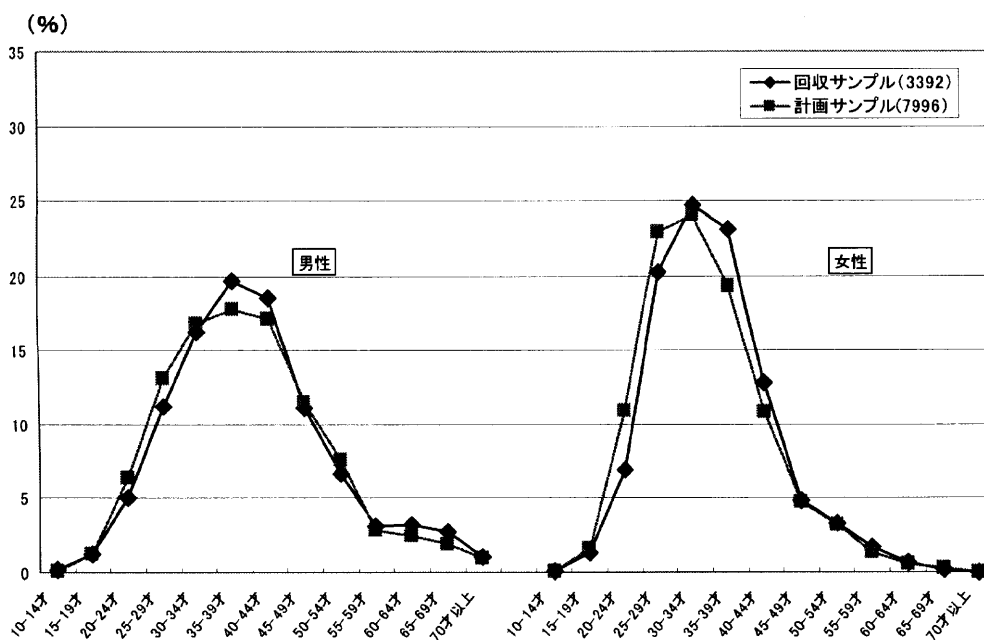


図5 計画標本と回収標本の関係：C社公募系、第1回（生活意識編）

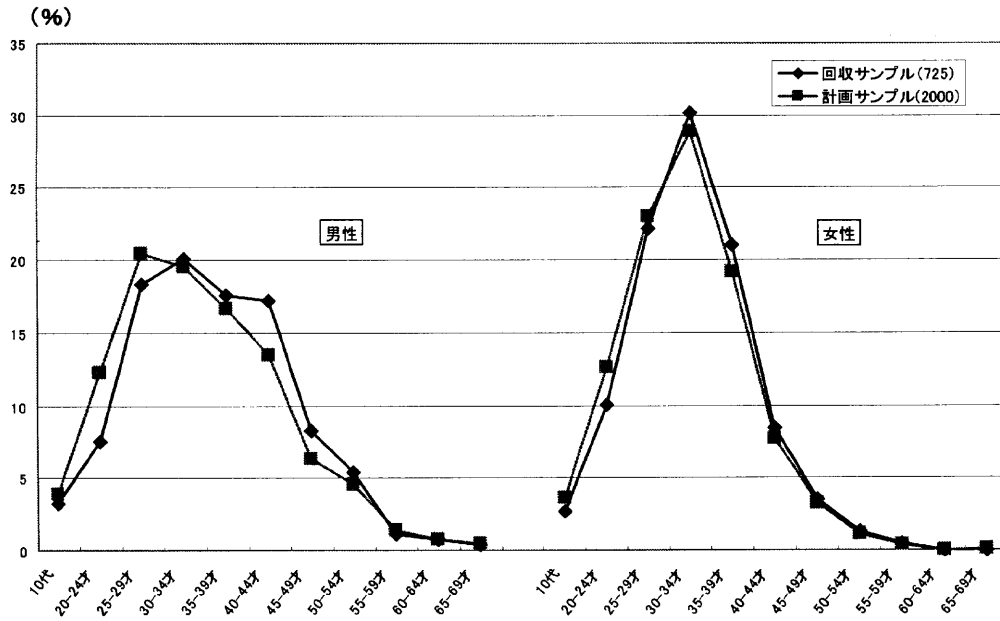


図6 計画標本と回収標本の関係：A社非公募系、第1回（生活意識編）

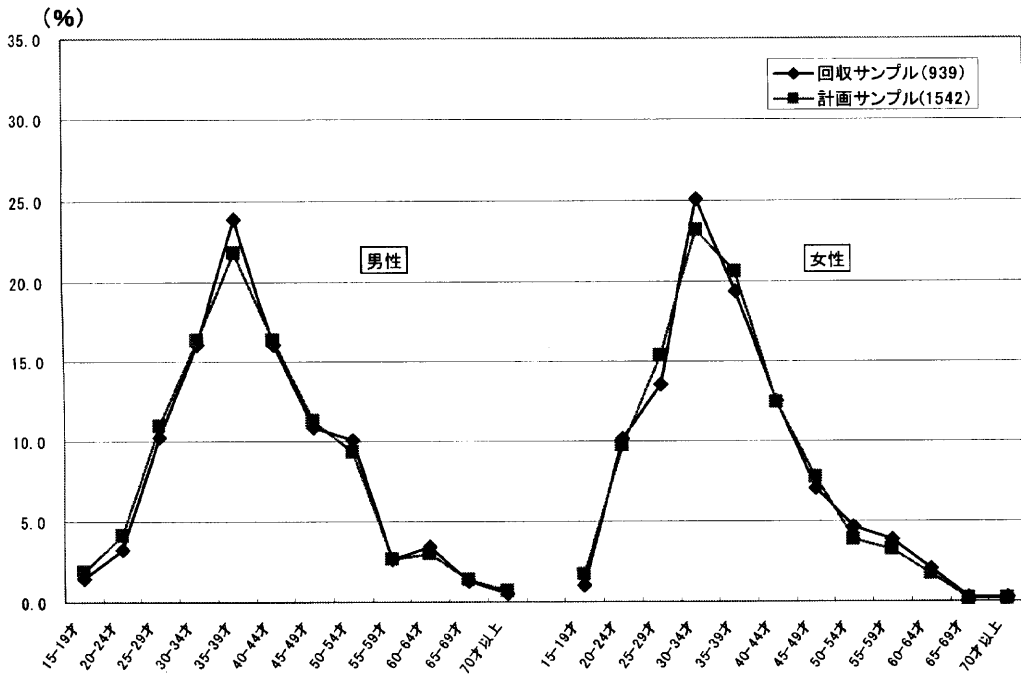


図7 B社非公募系、全パネル（計画標本）と第1回、第2回回収標本

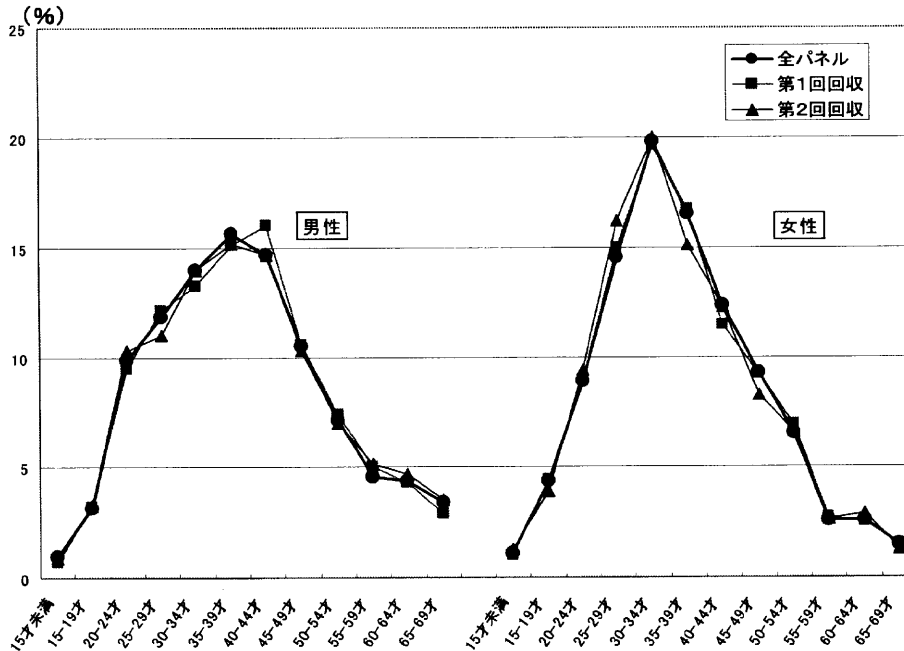


図4～7が計画標本と回収標本の属性分布（性年齢区分）の比較例である。説明するまでもなくサイトによって属性分布が異なることが分かる。つまりサイトにより登録者集団属性に差異があり、従って回答結果も変わるのである。

また、公募系の場合（図4、5）明らかに「計画標本と回収標本との間に系統的な偏り・ずれ」が生じる。実は、この現象は過去に行った公募系のすべてのサイトにおいて見られた現象である。第3次調査までの実験調査では、非公募系に相当する登録者集団が利用できなかったため、この現象が何によって生じるかが正確に見えなかった。しかし、第4次調査で複数の公募系、非公募系を比較する機会を得たことで、この現象が公募系特有のものであることが見えてきた。これを読み替えると、公募型の登録者集団では、既に登録者と回収者の間に不整合性が見られ、つまり“回収標本

は計画標本の縮図となっていない”らしいことが分かってくる。

一方、非公募系、つまり従来型の統計的サンプリングを行って選んだ登録者候補のうち「インターネット・ユーザであって、かつ調査への協力を合意した」登録者の集まりの場合（図6）は、当然のことではあるが調査協力度が高く、回答率も高い傾向になるので、計画標本と回収標本の結果にはほとんどずれを生じない。従って、回答率もかなり高いものとなる（後述）という、かなり当たり前の結果が現れる（図6、7参照）。

#### 8.4 従来型調査とインターネット調査の回収標本の比較

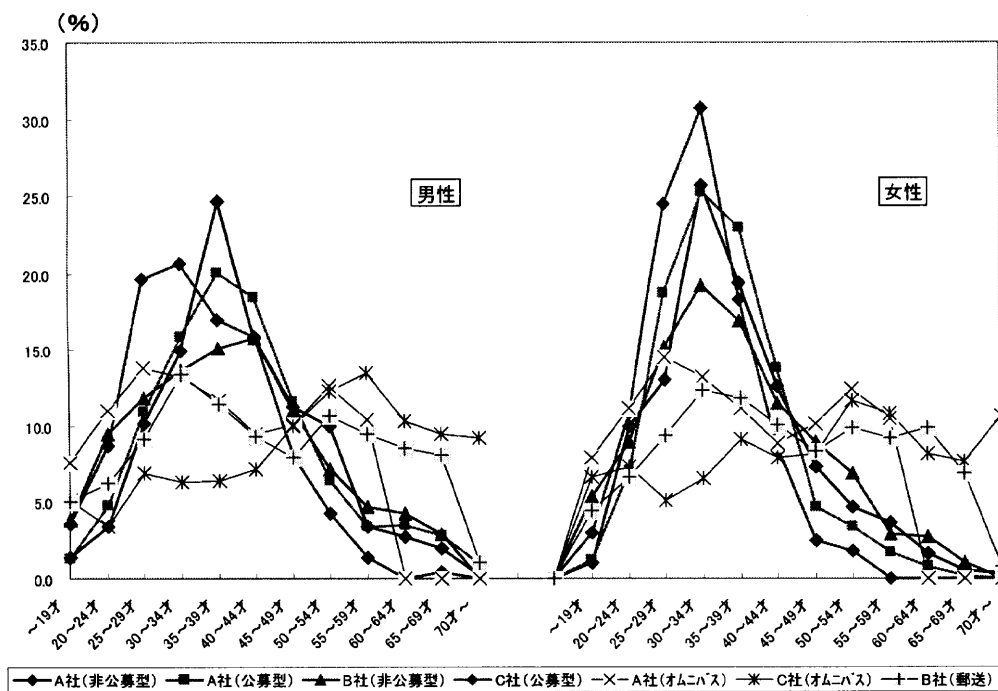
さらに、図4～7の回収標本の上に、従来型調査の回収標本の分布を重ねたものを作る（図8）。一般の母集団は団塊の世代（とその子供の世代）の双峰性の分布となる。従って従

来型調査の回収結果もこれに類似した傾向を示すはずで、確かに図8ではそのようになっている。これに対してインターネット調査における属性分布は単峰性の分布となり、しかも特定の年齢層に集中する。このように、「調査方式の差異が顕著」であり、調査方式間の優劣も簡単に議論できないことを示している。

このことは、仮に今後インターネット・ユーザ数や登録者数が増えたとしても、この歪みの解消が科学的に検証されない限りは代表性に問題が残るということである。非公募系のような登録者集団の作り方で、つまり、ま

ずインターネット・ユーザを選びさらに応諾者のみを登録者集団とする方法をとっても、やはり代表性に問題がある。登録者集団を集めたプロセスは確かに透明化されるものの、応諾率は一般に高いとはいえず（せいぜい10数%から20数%程度）、やはりこれでは一般の住民・社会人を代表するとはいえずもない。これらの特徴は、現在できることとして、登録者集団の集め方とその属性情報を具体的に開示せねば、その調査結果は信頼できないことを示している。（図2の集め方の模式図を参照）

図8 登録者集団の属性の例 - インターネット調査と従来型調査の比較（回収標本について）



さて以上の情報から、以下のような特徴が見えてくる。

調査サイトによって、明らかに属性分布の構成が異なる。

同じ調査機関（A社）であっても、公募系と非公募系で大きく結果が異なる。

公募系同士（A社、C社）では類似し、しかも計画標本と回収標本のずれが生じている。

非公募系同士（A社、B社）も類似し、こちらは計画標本と回収標本にはずれがほとんど見られない。結果として回答率は高い。

性別による年齢区分構成が微妙に異なる傾向にある。

つまり、インターネット調査では若年齢層（とくに、20代後半～40代後半）が多いが、サイトによってそのピークの特徴が異なる（つまり、年齢構成の分布が異なる）

一方、従来型調査では、いわゆる団塊の世代の2峰性の年齢構成の特徴が現れている。

ここで、前出のサイトの引用例から作成した年齢区分情報と比較する（表3）。実験調査では、A社、B社については登録者集団を厳密に確定に、そこから無作為抽出を行った。また、B社については無作為抽出を行わずに全登録者を対象に実施しているので、全パネルがすなわち計画標本そのものである。

一方、4社（W、X、Y、Zの各社）の年齢分布の傾向は、20代～40代に集中するという点では類似性があるものの、各年齢分布はかなり異なり、またX社（＝C社）を除いては、筆者等の結果とも異なっている。つまり、このように登録者集団、計画標本、回収標本の関係は「登録者の集め方が異なれば、当然回収標本、計画標本も異なる」というひどく当たり前のことが分かるのである。繰り返しになるが、この基礎情報がいまのネット・リサーチにはまったく闇の中にある。最近、回収標本の分析例をWeb上に開示するサイトがいくつか登場しているが、この情報だけを分析して見せても、実は問題の本質は何も解決しないことも分かる。また、偏りの原因は登録者の集め方が明らかにならねば分からぬことである。

しかし同時に、なるべく確率的アプローチに沿った登録者集団の捕捉方法と管理を徹底すれば、そして調査協力の応諾率を高め、回

答者の調査協力度を高めて回答率を上げることができれば、少なくとも計画標本と回収標本の一致性を高めることは可能であることも分かる（至極当たり前のこと）。ただこの場合、謝礼目当ての賞金稼ぎ、あるいはプロ化した回答者をいかに回避するかという問題がある。また従来のパネル型調査のように、期限を切った契約関係を結ぶ、調査回数を十分に管理するなど、登録者集団の劣化や疲労を防ぐための手段を必要とする。登録者集団のサイズを吹聴する前に、もっと基本的な行為があるといいたい。

では、インターネット調査の回答者は一般の人（通常の従来型調査における調査対象者）を代表するであろうか。また、今後、インターネット・ユーザが増えたとして、回答者は一般の人に近づくであろうか。これを見るには、インターネット調査と従来型調査との類似と差異はどこにあるかを知る必要がある。

しかし、調査方式の異なる調査結果を、単純に結果の類似・差異から比較してもなんら考察とはならない。しかも異なる調査方式で得た調査結果間に単純な補正法を適用して辻褃合わせを行うなどは暴挙としか言えない行為である。こうした浅薄な議論の背景には「インターネット・リサーチにいずれ移行するのだ」「どんなに比較検証を行っても所詮は曖昧模糊として実証ができようはずがない」という過信と誤信が見え隠れするように思えてならない。とくに、調査方式の異なる環境で得たデータセットに対して、加重補正法を適用すればよいとの意見や試みもある。我々もこのことには関心があって、実験的な試みを多数行ってきた。しかし、その内容から言えることは、やはり過大な期待に応えられるだけ有効な方法を見つけることはかなり大変との感触を得ている。さらなる慎重な研究が必要

ということである。自己体験では、理論を当てはめるに適した調査環境をいかに作れるかというはじめの段階（調査設計）ですでに大きな壁に立ち向かうこととなった。ささやかな実験例証を作ってはみたものの、実用化にはほど遠いのである。

既述のように、現時点ではインターネット・ユーザ全体を正確に知る情報はない。そこで、この質問の条件を少し緩めて、登録者集団となったインターネット・ユーザのうち、調査に回答してくれる（くれた）人達が従来型調査への回答者と比べてどうかを考える。同時に、従来型調査への回答者のうち、インターネット・ユーザである人とそうでない非ユーザの回答行動を比較することには意味がある。

ここで留意すべきことは、インターネット調査に限らず、多くの場合「調査方式が変われば、回答の結果も変わる」ということである。従って単純に従来型調査とインターネット調査との結果比較を行っても、あまり効果的ではない。そこで、実験調査では「調査票や設問を同じ内容に揃えて、実施時期を同時並行的に」行っているのである（そのような実験計画とした）。調査方式の違いだけがなるべく現れるような調査計画を設計しなければ、単なる比較では意味がないということである。こうした比較実験調査から得られた知見は膨大な量になり、それぞれが特徴的な結果を示している。ここでそのすべてを述べることはできないので、ある質問の例を挙げるにとどめるが、これは稀な例ではなく他の多くの設問でも類似した傾向が観察されている。

では「インターネット・ユーザと非ユーザの違い」は、どのように現れるのであろうか。用いた質問は、以下のようなものである。

Q：物事を決定するときに「一定の原則に従うこと」に重点をおく人と、「他人との調和をはかること」に重点をおく人では、どちらがあなたの好きな「ひとがら」ですか？

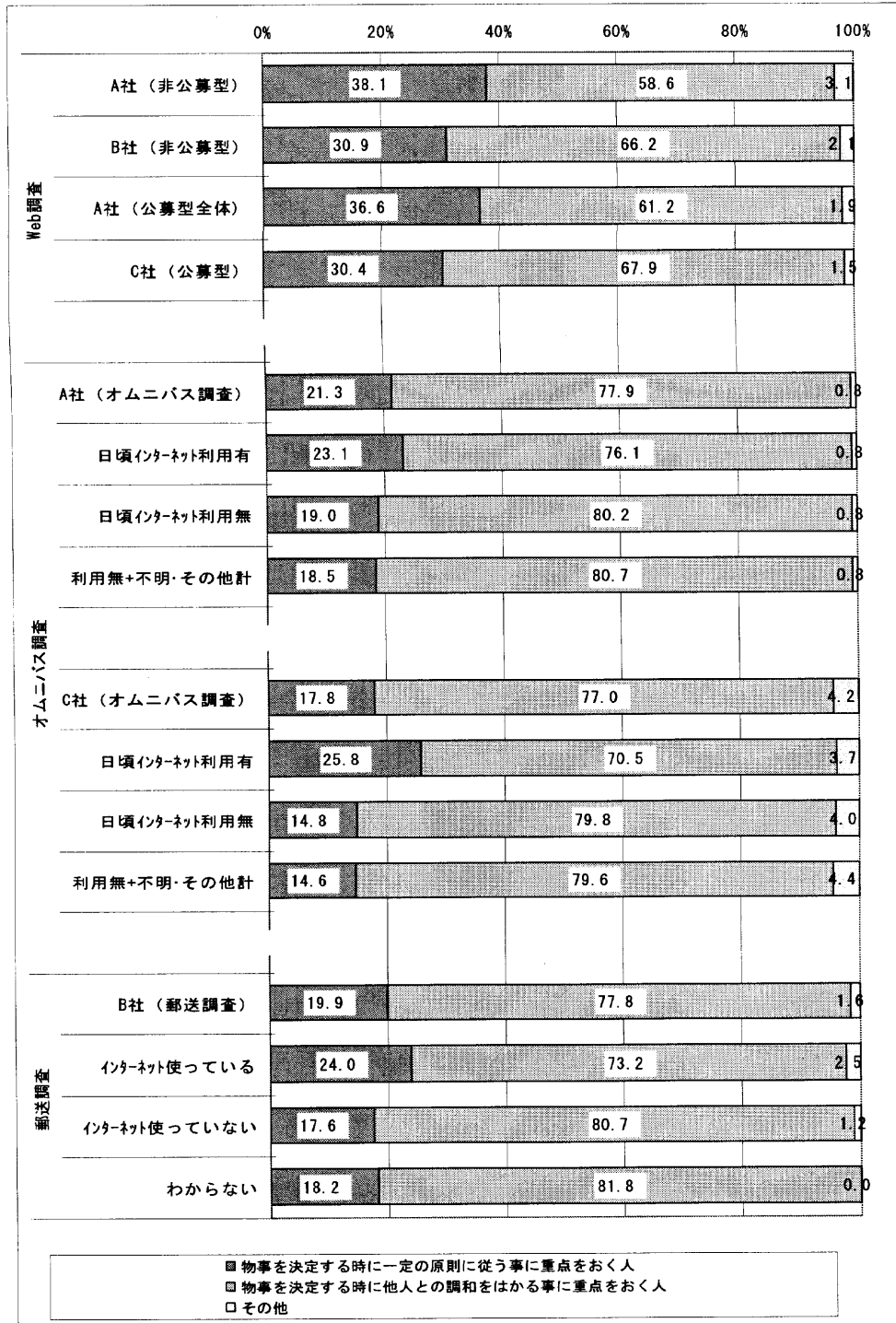
1. 物事を決定する時に一定の原則に従う事に重点をおく人
2. 物事を決定する時に他人との調和をはかる事に重点をおく人

これをインターネット調査と従来型調査で比較する。なお、従来型調査では、別の質問として「インターネットの利用の有無」を尋ねているので、これで層化し、インターネット・ユーザと非ユーザの回答の差異をみると同時に、ここでのユーザ層の回答がインターネット調査の回答とどう類似あるいは異なるかをみる。結果にはきわめて顕著な特徴がある（図9）。

- (1) Web調査と従来型調査の回答分布は大きく異なる。選択肢1（一定の原則に従う）がかなり多い。
- (2) 一方、従来型調査では選択肢2（他人との調和をはかる）が多い。
- (3) さらに、従来型調査のうち、インターネット・ユーザであるとした人は、非ユーザに比べ、いずれにおいても選択肢1が多い。しかし、そのポイントはWeb調査の回答者ほどは高くはない。

もちろん、Web調査の場合は、属性（とくに年齢構成）が従来型調査とかなり異なることは既に見たとおりである。であるからこの質問を見る前に、性別や年齢区分でブレイクダウンした情報の観察も必要である（実際に行っている）。しかしそれでもなお、このように、Web調査に自主的に参加した人、従来型

図9 従来型調査とインターネット調査におけるインターネット・ユーザ、非ユーザの比較例



調査におけるインターネット・ユーザ、そして非ユーザの間には、微妙な違いがあるということである。この例は稀なものではない。実験調査結果では、類似の結果が多数得られている。また、この設問に関しては一連の実験調査で繰り返し用いており、どの調査回でも同じ傾向を示している。もちろん、Web調査と従来型調査の結果がほとんど変わらない質問も多々ある。これはこれで、継続的に調査を行う中で、やはり類似した傾向を示すのである。確かにこのことは、そう単純には議論できない要素が多々ある。しかし、従来型調査と同様に、従来その他調査で実績を積んだ傾向がよく知られた質問を測定のインジケータとして組み入れて比較調査を行うことはきわめて重要である。 (つづく)

次号では以下の内容を掲載します。

- 9 . 登録者集団と有効回答率の関係をどう考えるか
- 10 . 特別な回答者、いわゆるプロ回答者集団の混入
- 11 . 調査品質と調査誤差の関係
- 12 . 新たなインターネット調査システムの構築に向けて
- 13 . 世論調査に用いることはできるのか
- 14 . 個人情報保護法、プライバシー、調査倫理の問題
- 15 . むすび