

ウェブ調査におけるパラデータの有効利用と今後の課題

— いくつかの実験調査にもとづくデータ解析による検討 —

Effective and Potential Use of Paradata in Web Surveys and Future Issues

— Discussion through Exploratory Data Analysis on the Experimental Survey Results —

大隅 昇 統計数理研究所
林 文 東洋英和女学院大学
矢口博之 東京電機大学理工学部
簗原勝史 (株) シー・エス・マーケティング・ジャパン

キーワード ウェブ調査, パラデータ (サーバー側パラデータ, クライアント側パラデータ), トラッキング・データ, 公募型・非公募型パネル, 確率的・非確率的パネル, 回答率, 参加率, 回答所要時間, 回答遷移と推移パターン, 総調査誤差, 調査の質

1. はじめに

最近の調査実施環境が直面する諸事象, たとえば, 回答率の低下, 無回答や調査不能の増加, そしてさまざまな調査誤差要因の影響により, 期待するような調査の質の確保が非常に難しくなっている。ウェブ調査も例外ではなく, 回答率や参加率は低下の傾向にあり, しかも検証すべき課題が無数にある。インターネットが普及したとはいえ, ウェブ・パネルあるいはそこから得られる標本は, 一般母集団はもとより, インターネット母集団をどう代表するのかという根本的な問題が依然としてある。とくに, 確率標本の利用が困難で, 多くは非確率標本となることで偏りが生じ, しかもさまざまな調査誤差 (とくに, カバレッジ誤差, 無回答誤差, 測定誤差) の影響が考えられ, それらがあまり明らかでないことがある (Tourangeau et al., 2013 ; Dillman et al., 2014)。

一方, 多くの研究で指摘されているように, ウェブ調査で得られる調査データは便利な商品となっており, 商用化が進んでいる (Couper, 2008 ; Tourangeau et al., 2013 ; 大隅, 2010)。大半のウェブ調査データの利用者にとっては, その質よりも, 価格や即時性, 利便性が優先されるようになった。しかし, 商用の調査は別として, 学術研究にウェブ調査を用いる場合は, かなり慎重な対応が必要である。

ウェブ調査に限らず調査実施過程のさまざまな場面で, コンピュータ支援による調査情報収集 (CASIC : Computer Assisted Survey Information Collection) が容易となった。情報量が増え貴重なデータが集められるようになったが, 問題が複雑になったともいえる。このことから, 欧米の調査方法論研究の根底には, 調査データ分析の共通した目標として, 可能な限り整った測定を行い, 総調査誤差 (TSE : Total Survey Error) の低減を考えることが調査の質の向上につながる, という考え方がある (Groves et al., 2004 ; 大隅・^{にお}鳩, 2012)。こうした研究の支流の1つに“パラデータの研究”がある。パラデータパラデータの概念, その誕生から現在に至る経緯や, その解釈, 適用場面などについては, Couper (2017), 松本 (2017) の報告に詳しいのでそれをお読みいただきたい。

本稿の目的は, “ウェブ調査”におけるパラデータの取得過程と利用場面について, 基本的な情報を紹介することにある。欧米の研究はかなり進んでおり, しかも多岐にわたる。一方, 日本国内ではほとんど知られていない分野でもある。ここではまず, 筆者らの過去に行った実験調査における基本的な指針を述べ, つぎにウェブ調査におけるパラデータとは何かを概観する。続いて, パラデータの

取得過程の考え方、具体的なパラデータの種類と例示、いくつかの調査で得たパラデータから得られる基本的な情報（回答遷移パターン、回答率、回答所要時間など）にもとづく簡単な分析例を述べる。おわりに、パラデータを巡る今後の課題、効用や問題点について触れる。

2. 実験調査の経緯 — パラデータ分析のきっかけ

欧米では、かなり早くからオンライン調査（インターネット調査、ウェブ調査）に関する多様な研究が行われていた（詳しくはこれらを参照；Bethlehem & Biffignandi, 2012；Couper, 2008；Dillman et al., 2014）。しかし日本国内では、マーケティング・リサーチなどの商用場面での急速な普及にくらべ、研究対象としてオンライン調査の特性そのものを体系的に詳しく検証することはほとんど行われてはいなかった。

そうした中で、大隅ほかは 1996 年末にインターネット調査と従来型調査法の比較研究を目的とした研究課題「調査環境の変化に対応した新たな調査法の研究」に着手し、1997 年にはじめて実験調査を実施した（第 1 次実験調査）。その後、2003 年までに、継続的に実験調査（第 2 次～第 4 次実験調査）を進めてきた（以上を [第 I 期]）。さらに、複数のウェブ・パネルを同時点で比べる共同実験調査（2006 年、[第 II 期]）、異なる課題の複数の実験調査（2010 年～2012 年、[第 III 期]）と進めてきた。

[第 I 期] の実験調査については、実験調査の目的、調査設計、調査内容、調査結果（分析概要や集計データ）ほか、個票を除くすべての情報を開示している（大隅ほか, 2000；大隅ほか, 2004）。他の実験調査の結果も機会をみて開示してきた。この一連の実験調査として、約 70 回（うち、ウェブ調査がおもなもので 52 回）の調査を行った。

実験調査では「実施要領」を設けて実施環境を標準化し調査を行った。重要なことは「ウェブ・パネル間に差異があること」で、適切な比較検証を行うには、調査実施条件をなるべくそろえることが必要である。このため、複数の調査機関への協力要請と合意形成を行った。また実験調査の開始時に、ウェブ調査の技術的要素や特性に注目し、いくつかの項目を要件とした。たとえば、「複数の公募型・非公募型パネルの比較」「同時点調査」「調査票の共通化」「回答制御の適用可能性」「ほかの調査方式（郵送、面接、訪問留置）との比較」「標本抽出枠の設定と標本抽出法の共通化」「トラッキング・データの記録分析」などを実施要件として設定した（詳細は大隅ほか, 2000；大隅ほか, 2004 参照）。

当時は回答者行動を電子的に追跡し統計分析を行うことを説明する用語がなかったため独自に「トラッキング・データ」と名付け、パラデータという用語ができる前から、これを用いてきたが、意味するところはこれにほぼ同じである。また、「公募型・非公募型パネル」とは、確率的パネルを用いるか否かに関連する。これも実験調査時の「登録者集団」（ウェブ・パネルであり標本抽出の対象とみなすことができる標本抽出枠相当の集団）の構築方法を説明するために用意した語句である（大隅ほか, 2000）。

ウェブ調査では、用いるウェブ・パネルの形態をどう考えるかは重要である。本稿では実験調査時に合わせて、“公募型と非公募型”に分けて考えるが、これは欧米研究で言う、非確率的パネルと確率的パネルに“ほぼ”対応すると考えてよい。たとえば、Couper (2000a) による 8 分類にしたがうと、「公募型」はボランティア・パネル（オプトイン・パネル、アクセス・パネル）であり（Couper 分類の「タイプ 3」）、「非公募型」は、確率的パネルに相当する（Couper 分類の「タイプ 7 あるいは 8」）。

非公募型パネルの利用が重要だが、過去の実験調査時点ですでに、これの利用は限定されていた。また現時点ではほとんど存在しない（パネル構築が面倒、調査システムの維持管理、調査費が高いなどの理由で、商用として成り立たない）。欧米もこの点で類似するのだが、それでもなお、米国の

KnowledgePanel® (GfK Knowledge Networks 社), オランダの LISS パネル (Longitudinal Internet Studies for the Social Sciences ; Das et al., 2011) などがある。また最近, NORC (シカゴ大学) が AmeriSpeak® という確率的パネルを用いたサービスを開始した。

重要なことは「確率的パネルと非確率的パネルによるウェブ調査」があり, 両者の特性はかなり異なること, 両者の違いを意識した調査や分析が必要ということであるが, これを留意して行った日本国内の調査研究はきわめて少ない (樋口ほか, 2012)。

3. ウェブ調査の特徴とパラデータの分類

過去の研究報告で, ウェブ調査のさまざまな特徴が指摘されてきた (Couper, 2008 ; 大隅, 2010)。とくに, パラデータ (トラッキング・データ) の分析から見えてくる特徴としては, 「回答率・参加率と調査内容の関係」「完答と未完答あるいは無回答, 中断の関係」「回答所要時間の傾向」「初頭効果の可能性」「質問項目の難易度と回答遷移パターン」「調査票形式の回答への影響 (グリッド, プルダウン・メニュー, スクロール形式とページング形式など)」「ストレートライニングを含む労働最小化行動・最小限回答行動の可能性 (satisficing behavior ; つまり, 調査質問を読むことや回答するための労力を割かないこと)」「自由回答データの取得過程の評価」など, さまざまな項目を調べてきた (大隅, 2010 ; 大隅, 2012)。こうした傾向分析を行ううえで注目される研究課題の1つが, パラデータとそれを用いた調査評価分析の方法である。

パラデータの特徴は, 調査過程における諸事象を, 調査実施者側 (調査主体) から回答者 (調査対象者) 側を観察・測定し, 調査の質の改善, 調査誤差の低減などに役立てることを意図して積極的に集めることにある。また, ウェブ調査に固有の考え方もない (Couper, 2017; Couper, 1998 ; Nicolaas, 2011; West, 2011 ; Kreuter (ed.), 2013)。

まず, “ウェブ調査におけるパラデータ”の要件を整理する (Couper, 2014 ; Kaczmirek, 2009 ; Heerwegh, 2011 ; Kreuter & Casas-Cordero, 2010)。

- ・ 調査対象者または回答者に対して, 明示的な応答 (回答記入, 口頭回答, 回答選択肢を選ぶなど) を求める必要がないこと。
- ・ つまり, 調査対象者の意識しないところで (観測測定が行われていることが知らされずに), 調査システム上で集められるデータであること。
- ・ データは, 調査対象者単位・個人単位で利用可能であること (実際には, 発生事象単位となる)。
- ・ データは, より一般的に意味のある調査情報として要約化されること。
- ・ 具体的には, 調査の質を評価する“指標”として利用可能であること。
- ・ 原則として, コンピュータ支援の環境下で集める。具体的には, 調査用具の情報機器類 (PC, タブレット端末, スマートフォンなど) を介して自動的に集めること。
- ・ 電子調査票の作成には (調査回答データの取得のために), 原則として HTML を用いること。
- ・ パラデータ収集のために, ソフトウェアあるいはスクリプト言語などを用いて作成したツールを必要とすること (JavaScript, PHP, Java, Perl, ASP.NET など)。

ウェブ調査におけるパラデータはその構造的な特性, とくに“情報の集約度”の程度に応じて4段階に分けて考える (Kaczmirek, 2009 ; Heerwegh, 2011)。「第1段階」は, 調査過程で生じる回答者の個々の回答行動・動作 (マウスのクリック, 移動, 画面内やブラウザ内での位置, キーボード打鍵などの

発生時点とその時間情報)をパラデータとしての測定記録することであり、通常はこの段階のデータは“非構造的な文字列情報”である。「第2段階」では、第1段階で取得のデータを、回答者単位で動作単位の回数を計数し各動作を説明する変数に変換する。「第3段階」では、回答者および変数の単位で情報を要約する。そして「第4段階」では、ここまでの要約情報から、集約度の高い具体的な諸指標(統計量)を誘導し統計分析を行う。たとえば、回答率や参加率、完了率、無回答率、中断率、ページあたり回答所要時間、総回答所要時間など、多数の指標を用いてパラデータ情報を総合的に要約分析する。

さらに、調査システム上で具体的にパラデータを集める際の“技術的側面”からこれを分ける(Couper, 2014; Heerwegh, 2011; Callegaro, 2015)。つまり、実際の測定メカニズムと具体的に測定される内容を「サーバー側パラデータ」(SSP: server-side paradata)と「クライアント側パラデータ」(CSP: client-side paradata)の2つに分けて考える。

SSP(サーバー側パラデータ)とは、ウェブ調査を運用する調査システムを実装したサーバー上で観測される事象のログ情報から得られるデータのことで、調査実施者あるいはシステム運用者が測定可能なデータである。CSP(クライアント側パラデータ)とは、調査対象者(クライアント側)のコンピュータ上で生起する事象を観測記録することで得られるログ情報とそれをサーバー側(調査実施者側)に送ることで得られるデータのことをいう(この処理は、回答者にはみえていない)。

調査票設計にはHTMLを用いるが、この他に、調査システムの管理運用やパラデータを取得するツールを用意するために、さまざまなスクリプト言語も用いる。また、利用目的の性質上、通常は、CSPとして取得のデータのほうがSSPよりも多くなる傾向にある。たとえば、SSPの代表的な情報は、IPアドレス、調査実施側(サーバー側)が配信した調査票ページの発信情報、クライアント側からデータを受信した際の時間情報(タイムスタンプ)、回答情報(フォームデータ)を含む、ごく限られたログ情報である。

一方、CSPは、調査票にアクセスした回答者が、個々のウェブ・ページ内でどのような回答行動を取るか、その事象を追跡し記録することで得られるデータである。HTMLや、それにJavaScriptで作成したパラデータ取得用のツールを組み込むことで、CSPとして、回答者の回答行動を時間軸に沿って非常に細かく追跡できる。

回答者の回答行動の記録は、動作(具体的にはマウス・クリック、キーストローク、タップなど)が生起した時点測定することで得られる。たとえば、調査票のあるページへのアクセス時点やページ閲覧時点、質問文の選択肢を選んだ時点、選択肢の選び方が一度でなく何回か選び直した時点、最後に選んだ選択肢のクリック時点、と回答行動履歴を詳細に記録する。

通常の調査測定では、回答者が質問のどの選択肢を選んだかという情報だけを“静的に取得”している(選んだ選択肢がわかるだけ)。しかし、パラデータとしては、最終の選択肢が決まるまでのすべての行動を、双方向的なやりとりを通じて“動的に記録”する。また、これが調査票の全ページ(ウェブ・ページ単位)ですべて記録される。

測定した原始データは、通常は“長い「文字列」”として記録される(集約度の「第1段階」)。いわゆる非構造的データであり、行列形式にもなっていない。なお、CSPの取得が可能なソフト(商用ソフト)やサービスが、すでに公開されている(Clickdensity, Clicktale, Crazy Egg, m-pathyなど)。また研究者が公開しているソフトもある。たとえば、Heerweghの作成したJavaScriptを用いたデモソフトや、Kaczmarek(2009)のUCSP(Universal Client Side Paradata)などがある。前者は、簡単に試すことができる。回答行動の分析を目的として、CSPの取得を意図したツールを調査システムに組み込

むことは、日本国内の商用ウェブ・パネルでは、おそらくほとんど用いられていない。

4. パラデータから得られる情報

文字列のままでは統計分析には適さないので、これを分析可能な形式に変換する操作やソフトウェアが必要となる。われわれの場合は、HTML とサーバー側スクリプト言語である PHP (PHP Hypertext Preprocessor) を用いて、おもに SSP として文字列を記録している。ここで、PHP の代わりに Perl や ASP.NET など利用可能である。調査票の質問量やレイアウトの複雑さにもよるが、多くの場合、こうした文字列情報は非常に長く量も多い。文字列がどのようなものか、例をみるのが早いので、簡単な例とそこで取得できる情報の一部を示そう。図 1 は、<調査 1> で得られたある回答者のやりとりの文字列情報の一部である。ここには、ある回答者について、8 ページある調査票のうち、はじめの 2 ページの追跡状態を示してある。

```
①(管理サーバー名),2012/02/24,10:41:46,(IP アドレス),"*GET_VARS*','hjn=','userid=','*POST_VARS*","Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.1; WOW64; Trident/5.0)
②(管理サーバー名),2012/02/24 10:41:51,(IP アドレス),"*GET_VARS*','*POST_VARS*','goNextPage=次へ','userid=','hjn=','jobno=','current_step=page01','step=page01','page=1','','Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.1; WOW64; Trident/5.0)
③(管理サーバー名),2012/02/24 10:43:29,(IP アドレス),"*GET_VARS*','*POST_VARS*','Q1=4','Q2=3','Q3=2','Q4=1','Q6=1','Q7_2=1','Q7_1=1','Q7_6=1','Q8_a=1','Q8_a_3_Other=','Q8_b=1','Q8_b_3_Other=','Q8_c=1','Q8_c_3_Other=','Q8_d=1','Q8_d_3_Other=','Q8_e=1','Q8_e_3_Other=','Q8_f=1','Q8_f_3_Other=','goNextPage=次へ','userid=','hjn=','jobno=','current_step=page02','step=page02','page=2','Q5_Q25_flg=2','Q7shf=2,4,1,3,5,6,7","Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.1; WOW64; Trident/5.0)
```

図 1 文字列情報の例 (おもに SSP)

注 1: 丸数字は説明のために付けたもので、実際の文字列には含まれない。

注 2: マスク部分は開示できないので伏せ字とした。また、(管理サーバー名) (IP アドレス) には実際の情報が入るが、ここは仮の文字列を入れた。

この例では、文字列は ASCII テキストファイルとなっており、「”」で括った部分が 1 つのセル (要素単位) であり、カンマで区切られたエリアが 1 つの情報要素となる。たとえば、①は、回答者の基本情報 (ID など) を確認し、アクセスがあったことを記録する。続く②では、時刻の他、ジョブ管理番号など、システム管理に必要な情報 (管理サーバー、IP アドレス、タイムスタンプなど) と、回答者がウェルカム・ページ (表紙ページの page01) に入り、「次へ」をクリックしたことが記録される。③では、同じく時刻と管理情報、それに、この回答者が選んだ質問 1 (Q1) から質問 8 (Q8) までの個々の選択肢が記録され (受信データ)、さらに「次へ」で次ページへの移動が確認される。以下同じように、8 ページまで続く。ここの例で示した SSP に加えて、CSP で得られる情報を含め、文字列として具体的に得られる“おもな情報”を表 1 に要約した。

各回答者の示す個々の行動が、時間軸に沿って発生事象単位で非構造的な文字列情報として記録されるが、前述のようにこのままでは統計分析を行うことができない。このとき、ある回答者の個々の事象数（発生件数）は、回答者によってさまざまであるから、調査回答そのものの回答者数や回答数（頻度）とは一致しない。パラデータの処理の面倒な理由の1つが、この回答者の調査質問への回答とパラデータは1対1に対応しないことにある（後述）。

これを統計分析可能な形式（たとえば行列形式）に変換するソフトを作成して使える形にパラデータを整える（集約度の「第2段階」）。また効率的なパラデータ分析を行うには、処理分析ツールを作ることが必要である。ここでは、おもに統計ソフトウェア JMP®（バージョン 12.2.0, SAS Institute Inc.）を用いて分析を行った。

表1 パラデータ（SSP, CSP）の文字列情報から得られるおもな内容

取得情報	取得内容の例
ユーザーID	パネル登録者に付与のID
タイムスタンプ（事象発生時の時間情報）	<ul style="list-style-type: none"> ミリ秒（ms）単位、マイクロ秒（μs）単位で測定が可能 回答所要時間の測定（選択肢単位、質問項目単位、ページ単位など） 調査票へのアクセスのみ・回答中断の情報 事象発生時点（例：クリック）とその時間間隔など ページ閲覧開始や質問文の選択肢の選択時のクリック時点の時間情報
回答データ	<ul style="list-style-type: none"> HTML（フォームデータ）の形で、回答選択肢別に回答者から受信
回答選択肢の選び方	<ul style="list-style-type: none"> ラジオ・ボタン、チェック・ボックスのオン/オフ、その回数 選択肢を選ぶ順序、回答選択肢変更の有無とその履歴など
自由回答	<ul style="list-style-type: none"> テキスト・フィールド、テキスト・ボックスの識別 書き込みの有無、書き入れた文字数やボリューム確認 キー操作情報や記入所要時間
ユーザー・エージェント文字列情報（UAS）	利用ブラウザ、OS、それらのバージョンなどの情報
クライアント側のPCの情報	<ul style="list-style-type: none"> 画面サイズと解像度；ブラウザの画面サイズ 搭載ソフト/プラグイン・ソフト（FlashPlayer, Java など）のインストールの有無と設定有効の確認

5. パラデータの分析例

分析例を示すために用いる調査データを、過去の実験調査から4つ選び、各調査の特徴を表2に要約した。4つの調査は、2節で述べた区分に従うと、[第Ⅱ期]（公募型パネル）、[第Ⅲ期]（非公募型パネル）に相当する。パラデータの特徴がよくみえるように、調査目的と内容、調査票設計などが異なる調査を意図的に選んだ。とくに、調査票の質問数（項目数と選択肢数）、レイアウト（グリッド形式の使い方、その個数の多少）、ページ数の多少、設定機能（回答制御の程度、分岐やパイピング処理）などに変化のある調査を選んだ。

表2 分析に用いる調査の概要

項目	略称	<調査1>	<調査2>	<調査3>	<調査4>
調査課題		日常生活と伝統的な価値観に関する調査	本と読書についてのおうかがい(電子書籍利用, 本と読書調査)	本と読書についてのおうかがい(電子書籍の利用に関する国内状況調査)	普段の生活やインターネットなどについてのおうかがい
調査名略称		基底意識調査	本と読書に関する調査	電子書籍の利用動向調査	ウェブ・サイト共同実験調査
調査実施年月日		2012年2月24日(10時)～3月1日(10時)	2010年1月8日(17時)～1月12日(9時)	2011年2月7日(12時)～2月8日(9時)	2006年3月23日(10時)～3月29日(10時)
調査対象地域		首都圏(埼玉県, 千葉県, 東京都, 神奈川県)		首都圏(埼玉県, 千葉県, 東京都, 神奈川県)および京阪神圏(京都府, 大阪府, 兵庫県)	
ウェブ・パネルの種類		非公募型	非公募型	非公募型	公募型
全質問数とその内訳		41	27	40	47
①最大クリック数 [ラジオ・ボタン, チェック・ボックスの総数]		292	206	396	536
②テキスト・フィールド数		20	15	39	20
③自由回答の質問数		4	1	2	0
④グリッド形式のみの質問数 [グリッド形式の最大クリック数]		ラジオ・ボタン(2問) [52]	ラジオ・ボタン(1問) [18]	ラジオ・ボタン(4問) / チェック・ボックス(1問) [87]	ラジオ・ボタン(5問) [117]
調査票のページ数 (ウェルカム・ページを含む)		8	16	25	7
質問項目のおもな特徴		<ul style="list-style-type: none"> 質問の意味がやや難しい テキスト・フィールドの「その他」書き込みが多い。 郵送調査とレイアウトをそろえたことでページ当たりの質問数が多い いくつかの質問で, 選択肢の並び順をランダム化した 	<ul style="list-style-type: none"> テキスト・フィールドへの数値記入が多い 分岐, パイピング処理がある ページあたりの質問数を少なくした いくつかの質問で, 選択肢の並び順をランダム化した 	<ul style="list-style-type: none"> テキスト・フィールドへの数値記入が多い 分岐, パイピング処理がやや多い いくつかの質問で, 選択肢の並び順をランダム化した <調査2>と同じ質問を含む 	<ul style="list-style-type: none"> 全パネルで共通の質問を用いた チェック・ボックス形式の複数選択肢が多い グリッド形式かつ複数選択の選択肢数が多い質問がある ブルダウンが1問のみある ページ当たりの質問数が多い 一部, プログレス・インジケータあり
回答制御の有無 (*いずれも「部分的に回答制御」)		<調査1>～<調査3>は以下を設定。 ・各設問で無回答を選んだ場合, 警告を出して回答要求を行う。 ・ページ間の往復は可としページ間の回答行動の履歴をすべて記録。		<調査4>は以下のようにした。 ・同一回答者の二重回答, 重複回答は許容しない。 ・各設問への無回答の選択は可とする。 ・ページ間の往復は可とし, ページ間の回答行動の履歴をすべて記録。	
期限内の回答時間の強制打ち切り		いずれも行わない(つまり, 調査期限内, 常時オープンとした)			
無回答時のアラート		いずれも「あり」			
主要参考文献ほか		林ほか, 2010	矢口・大隅, 2010	総務省, 2011	前田ほか, 2005; 前田ほか, 2007; 大隅, 2010
関連情報の所在 [URL]		http://wordminer.org/wp-content/uploads/2010/09/j00080a.pdf	http://wordminer.org/wp-content/uploads/2010/09/j00079a.pdf	http://www.soumu.go.jp/menu_seisaku/ictseisaku/ictriyoushinict.html	http://www.wordminer.org/wp-content/uploads/2013/04/244_0.pdf

5.1 調査情報の要約 — 調査の質を知ること

<調査 1> (非公募型パネル) を例として、基本の調査概要情報を表 3 に示した (集約度の「第 3 段階」「第 4 段階」の一部)。こうした項目を記録することは、ウェブ調査を用いる上で重要である。一般に回答率に注目するが、これだけでは十分とはいえない (Bethlehem & Biffignandi, 2012; Groves et al., 2004)。調査の質を考えるうえでは、無回答 (率) とその発生理由を知ることが重要である。とくにウェブ調査では、高い回答率は期待できないので、無回答情報の分析は重要である。

このとき、「回答制御」が関係する。かりに完全な回答制御を設定した場合、無回答は生じない。パラデータを測定しても、ここにもみるような調査依頼数と回収数の不一致は、原則として生じない (ここだけ見て、ウェブ調査はゴミが少なく、回収データが整っているとの意見があるが不適切である)。しかし、回答者の調査内容への関心度や理解度、PC 操作のスキルなどによっては、滞りなく完答してくれるとは限らない。実際、かならずある割合で「完答できない人」がある。また、調査システムの支障などで、回答者が回答する意思があるのに、回答できないこともある。こうした状態をパラデータ分析で把握し、その調査の質を測ることが重要である。ウェブ調査を学術研究の調査ツールとして用いる場合は、こうした情報を“調査の質の評価情報”の 1 つとして示すことが必要である (実際には、ほとんど行われることはないようだ)。

表 3 の「調査依頼発信数」が実際の調査依頼人数であり、「有効回収回答数」が得られた実質的な回収標本の大きさである。つまり「回答率」は約 58.1 (%) となる (ウェブ調査としては非常に高い)。「無回答」と「無回答の内訳」欄にある情報は、字句の通り、「未着」「アクセスなし (無接触)」、「アクセスのみ」 (調査対象者が調査票にアクセスしただけで回答を行わなかった)、「回答中断」などの情報で、これらはパラデータの分析から得られる。

無回答は面接調査や郵送調査でも見られる現象だが、ウェブ調査では、上のようにこれの内容を具体的に知ることができる。しかし、発生理由までは正確には分からない。ここらをどう調べられるかが課題の 1 つである。「アクセスログ数」の欄は、ウェブ調査に固有の情報で、パラデータの分析で得られる。アクセスログ数は、調査対象者の実際の回答行動と、実際に起きた個々の事象の発生件数 (延べ件数) として記録される情報である。

「調査依頼発信数」と「アクセスログ数 (延べ件数)」の関係がわかりにくいので別の情報を用意した。図 2、表 4 は、回答者が調査票の各ページをどのように閲覧したかを、「回答遷移情報」 (ページ遷移パターンと回答形態パターン) として要約した情報である。ページ遷移パターン (図 2) は、回答を最後まで終えることのできた人 (回答完了件数: 延べ件数) と、回答を完了できなかった人 (回答未完了件数: 延べ件数) に分けてある。回答完了 (583 名) には、各ページを問題なく回答し終えた人 (567 名) と、挨拶文や調査内容などの説明のある表紙 (ウェルカム・ページ) に何度もアクセスしたのちに回答を完了した人や同じページを複数回確認 (往き来) したあとで回答を完了した人 (16 人) が含まれる。また、調査の案内情報は確認したがまったく回答をせずに「アクセスのみ」となった人 (82 名)、途中で回答を行ったが「中断」した人 (25 名) があつたとなる。

表 4 の左欄が、1,002 名 (計画標本の大きさ) のうち、実際に配信した 1,001 名の調査対象者の「回答形態パターン別の集計」である。回答形態パターン数が 11 通りあり、各パターンの回答者数が要約されている。右側の列は、回答動作の発生事象別の情報、つまり延べ件数 (1,088 件) でみた各パターンの要約情報である。「アクセスのみ」は、もともと回答行動の記録が行えないので同数である。パターン 4 の「アクセスのみ⇒回答完了」は、一度調査票を眺めたあと、(何らかの理由で) 再度調査票に訪問し、完答できた人である。パターン 3 の「アクセスのみ⇒アクセスのみ⇒回答完了」の場合、2 名が (動作

事象として) アクセスを2回行い(4回のアクセス), そのあと回答を完了したことで, 延べ6動作発生したとなる。このように, 延べ件数と実際の回答者数が異なり, これがこの種の情報の特徴である。ここで, この回答遷移パターンをグラフで示すこともできる(後述)。郵送調査などで行う事後の回収数の履歴情報も, ウェブ調査では取得データから容易に結果が得られる(矢口ほか, 2010)。

表3 <調査1>の調査実施情報の概要

調査実施期間	2012年2月24日 10:00開始 ~ 2012年3月1日 10:00終了
調査課題	日常生活と伝統的な価値観に関する調査 [基底意識調査]
謝礼	250ポイント (有効回答者全員に250円相当)
標本抽出の方法	ウェブ・パネル内サンプリング方式 (層別無作為抽出) ^{a)}
標本抽出枠	非公募型のウェブ・パネル
調査対象地域	首都圏 (埼玉県, 千葉県, 東京都, 神奈川県)
登録者数 [パネルの大きさ]	18,875 (人) [登録者数確定日時: 2012年2月10日]
計画標本の大きさ	1,002 (人)
調査依頼発信数 [総配信数]	1,001 (人) (100.0) / 依頼発信前除外が1 (人) ^{b)}
有効回収回答数と回答率 (%)	582 (人) (58.1) [回答率] ^{c)}
無回答と無回答率 (%)	418 (人) (41.8)
無回答の内訳と割合 (%)	未着15 (人) (3.6) / アクセスなし383 (人) (91.6) / アクセスのみ6 (人) (1.4) / 回答中断14 (人) (3.3) / 回答送受信異常はなし
アクセスログ数 [延べ数]	1,088 (人) (100.0)
アクセスの内訳 [延べ数] と割合 (%)	アクセスなし398 (人) (36.6) / アクセスのみ82 (人) (7.5)
*) アクセス延べ数は「82+608=690 (件)」	回答送信数608 (人) (55.9) とその内訳 [回収回答数 (人) 583 (95.9) ^{c)} / 回答中断 (人) 25 (4.1)]
非登録者の回答およびID不明	いずれもこのパネルでは生起しない
督促回数	1回 (メールで2月28日; 男女20代, 女性30代に対して) 内訳 [発信数: 99 / 未着数: 0]

a) 総務省「住民基本台帳」に基づく20~69歳の人口(2011年3月31日現在)から割当法で配分(性別×年齢で割当)

b) 調査時点で, 転居による調査対象地域外の対象者を除外

c) 583 (人) の回答があったが, 回収後に1名の不適切回答があり集計から除外, よって582 (人) が有効集計対象者

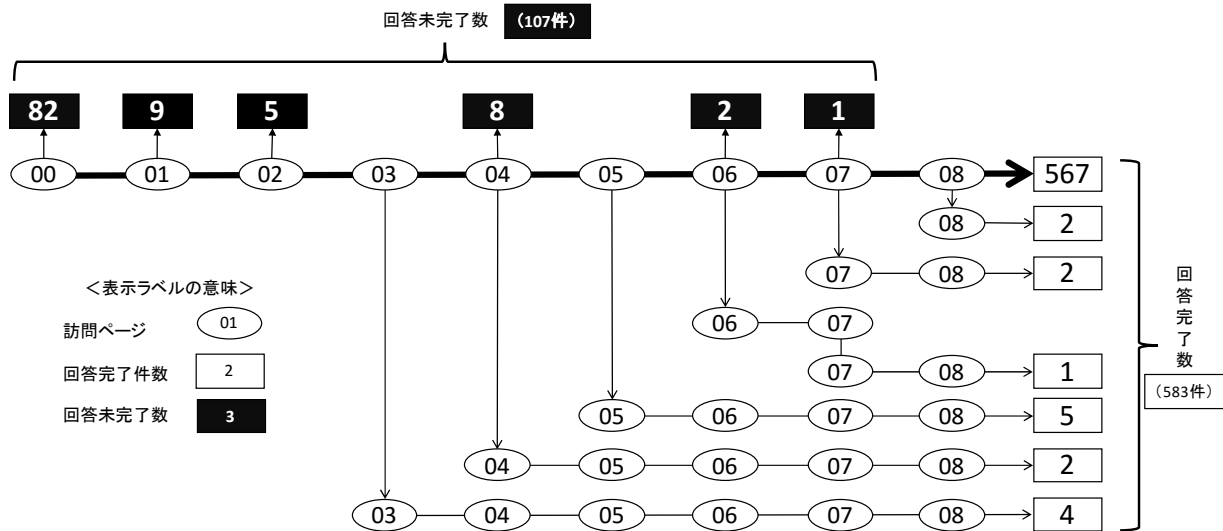


図2 ページ遷移パターンの要約

注：延べで併せて690件（583+107）のアクセスがあったことになる。

表4 回答過程（回答形態パターン）の要約

<単位：調査対象者>		<単位：延べ件数>		
パターン	対象者数	アクセスのみ	回答中断	回答完了
1 アクセスなし	398			
2 アクセスのみ	6	6		
3 アクセスのみ⇒アクセスのみ⇒回答完了	2	4		2
4 アクセスのみ⇒回答完了	7	7		7
5 回答完了	512			512
6 回答完了⇒アクセスのみ	46	46		46
7 回答完了⇒アクセスのみ⇒アクセスのみ	9	18		9
8 回答中断	14		14	
9 回答中断⇒回答完了	5		5	5
10 回答中断⇒回答完了⇒アクセスのみ	1	1	1	1
11 回答中断⇒回答中断⇒回答中断⇒回答中断⇒回答中断⇒回答完了	1		5	1
合計	1001	82	25	583

(1088)

5.2 回答率の傾向分析

回答率パラデータとすることかどうかについては、いろいろな議論がある。提唱者のCouper (2017)によると、これをパラデータと考えるかは慎重な議論が必要であるとしている（ここで始めてこの指摘を行った）。しかし、回答者が調査票を確認してから回答完了までの時間情報を追跡し、これを要約化することで回答率が得られるという点で、パラデータと無関係ではない。

回答率は、調査依頼者数（あるいは実際の配信数）のうち実質的に調査に協力し回答完了（完答、回収はできたが一部回答が欠損の項目無回答）となった回答者数の割合である（注：非確率的パネルの場合、“参加率”の利用が勧められているがここは回答率とした；AAPOR, 2016；ISO26362；Callegaro & DiSogra, 2008）。一方、回答を完了できない無回答や中断の傾向を「無回答率」「中断率」で測ることができる。回答率とあわせて、「無回答率」「中断率」を調べることは重要だが、ここでは回答率のみを取り上げる。実は、欧米の多くの研究では、回答率や中断率の大小が、ウェブ調査における無回答誤差や測定誤差に、どの時点でどのように影響するか具体的な証拠は、いまのところははっきりとは分かっていないという意見もある（Tourangeau et al., 2013）。

回答率の傾向を詳しく知るには、たとえば、異なる複数のウェブ・パネルに、共通の調査条件を適用し、回答率の変動と用いた質問項目の回答分布を比較することが考えられる。ウェブ調査では、こうした比較実験が他の調査方式に比べて容易に思えるが、複数のウェブ・パネルの同時実験調査は、実はそれほど多くはみられない。1つの共通した調査課題を、複数のウェブ・パネルでほぼ同時点で実施した実験調査例がいくつかある（Willems et al., 2006 ; The Advertising Research Foundation, 2009 ; Couper, 2010）。たとえば Willems 他 の分析では、パネル間の参加率の変動は大きいものの、各ウェブ・パネルの回答者の質問への回答傾向と参加率の大小との間には大きな変化は見られなかったという。

われわれの実験調査でも類似の調査を行ってきたので、ウェブ調査 52 件の回答率の分布を箱ひげ図で示した（図 3）。なおここで、[第 I 期] に行った調査のうち、調査時期と調査内容がほぼ同じ調査を 4 グループ（G01～G04）に分けて、回答率を示した（図中、G01～G04 の●印公募型と▲印非公募型）。さらに前述のように「G05 [第 II 期]」は 2006 年に行った複数の公募型パネルによる同一の調査内容、同時点調査である（図中◇印）。また「G06 [第 III 期]」は、調査内容に違いがあるが、多くは同一の非公募型パネルを利用している（図中 G04, G06 の▲印）。この過去に行った分析から、以下のような特徴が観察されている。

- ・ 公募型（●印、◇印）と非公募型（▲印）の回答率を比べると、非公募型がやや多めとなるようにみえるが、高い回答率の調査が少数あるため（G04）、そう大きな差異は認められない。
- ・ 回答率の大小により、異なるウェブ・パネルの同一質問項目への回答傾向が異なるかどうかは、あまり明らかではない（調査課題やパネル構成に依るようで、この点で欧米研究に類似）。
- ・ 欧米の研究では、回答率は次第に下降する傾向にあるといわれているが、実験調査の実施年次の範囲では、そうでもないが変動は大きい。むしろ、調査内容、質問量の多少、質問内容への関心度や難易度に影響されるようにみえる。
- ・ 用いるウェブ・パネルの、人口統計学的変数、とくに性別や年齢区分によって、回答率に違いがある（大隅・前田，2008 ; 大隅 2010）。

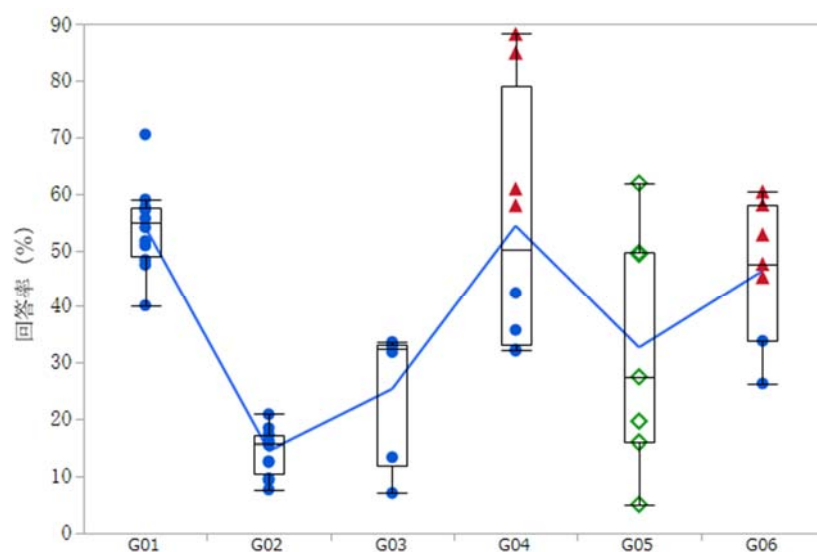


図 3 回答率の傾向（調査実施時期と調査内容でグループ化）[折れ線は平均]

5.3 回答所要時間の考え方とその分析

ウェブ調査に限らず「回答所要時間」は調査の質を測る重要な指標の1つである（例：態度強度の測定）。たとえばCATIであれば、調査員が回答者と応対する際の情報が記録されるので、質問文への回答時間、応対時間、対話情報などが分かる。また個別訪問面接聴取であれば、調査員が、調査開始から終了までの面接聴取の情報を記録し、のちにこれを電子ファイル化し、分析を行うであろう。これらはいずれもパラデータの種類である。

ウェブ調査の場合、回答所要時間という言い方だけではやや曖昧である。回答者は調査依頼（のメール）を確認し、つぎに調査票にアクセスし、内容を読み、指示に従って質問文を読み、回答を考えて選択肢を選び…と回答を進める。このとき、すでに述べたように、パラデータから得られる時間情報は多様であり、回答者に何らかの動作が発生すれば（例：マウスの移動やクリック、キーボードを打鍵し文字列を書き込む）、そのたびに各動作とその発生時点、各動作の所要時間（時間間隔）を求めることもできる。つまりクリックが動作発生時点の確認の有無を決める（注：スマートフォンやタブレット端末の場合、クリック動作をタップやスタイラス・ペンによるタッチと読み替えればよい）。

“調査票のあるページ内の動作発生事象”の時間情報を測るとき、いくつかの要素に分ける（Kaczmarek, 2009）。なおここで「ページ」とは、ブラウザ上に表示される区分単位のことをいう。たとえば、 t_A :

回答時間（クリック時点；answer time or crick time）、 t_P : ページ時間（データ送信時間；page time）、 t_M :

マウス移動時の運動量記録（位置、軌跡など）、 t_T : 思考時間（回答内容を考える時間）、そして t_{L1}, t_{L2} :

タイムラグ（サーバー側とクライアント側の情報授受で生じる時間差）、に分けて、これを元にページを配信してから、回答者（クライアント）から回答を受け取るまでのページあたりの「回答所要時間（サーバー時間；server time）」(t_S)とする($t_S = t_P + (t_{L1} + t_{L2})$), ここで、 $t_P = t_A + t_T + t_M$, また $t_A < t_P < t_S$)。

こうした“時間変数”を設定した上で、回答者が調査票の開始ページから終了ページまで進む過程の回答行動を追跡する（Kaczmarek, 2009；Heerwegh, 2011）。パラデータの取得・分析を重視する研究者たちの関心は、得られたパラデータ（SSP, CSP）の分析により、調査の質を測る指標となり得るのか、回答行動の何を測っているのか（例：質問文の内容の適否、質問数や調査票のボリュームや回答負荷など）、パラデータの種類で質の違いはあるのか、とくに、3つの時間変数 t_A, t_P, t_S の相互の関係が調査の改善にどう関与するのか、などにある。取得情報の分析は多様であるので、ここではSSPから得られるもっとも簡単な“ページあたりの「回答所要時間」”(t_S)を調べる。

5.4 ページあたりの回答所要時間の分析

取り上げた4つの調査には、表2に示したように調査票の質問数や選択肢数、質問文の種類（選択肢型質問、自由回答質問）やレイアウトなどによりかなり違いがある（調査票の構造の違いがみえるように意図的に選んだ）。

まず表5は、<調査1>の各ページの質問数とそこに含まれる回答選択肢を構成する部品とその個数ほかである。<調査1>は、1ページあたりの質問数が多い。実はこの調査は、郵送調査との比較を目的

とした2010年調査（林ほか，2011）の追加調査であり，その郵送調査で用いた質問紙型調査票に類似したレイアウトを用意した。

表5 調査票の各ページの構成：＜調査1＞の場合

ページ	質問数 ^{a)}	おもな特徴と構成部品（クリック数）	最大クリック数
page01	-	表紙（ウェルカム・ページ） [挨拶，調査の内容，回答指示，謝礼情報，回答見込み時間の通知など]	- ^{b)}
page02	8問（13問）	ラジオ・ボタン（44），チェック・ボックス（7） ／テキスト・フィールド（6）	51
page03	5問（12問）	ラジオ・ボタン（36），テキスト・フィールド（7）	43
page04	5問（13問）	グリッド形式ラジオ・ボタン（24），ラジオ・ボタン（20），自由回答（1問）	44 ^{c)}
page05	7問（13問）	グリッド形式ラジオ・ボタン（28），ラジオ・ボタン（8），チェック・ボックス（21），テキスト・フィールド（1），自由回答（2問）	57 ^{c)}
page06	6問（人口統計学的変数）	ラジオ・ボタン（38），テキスト・フィールド（3）	38
page07	9問（人口統計学的変数）	ラジオ・ボタン（44），チェック・ボックス（15）， テキスト・フィールド（3）	57
page08	1問	自由回答（1問）[調査への感想]	0 ^{c)}
合計	41問（67問）	-	最大292（クリック）

a) 括弧内はグリッド形式で表示の複数の質問項目を1問と計数したときの質問数；b) 「次へ」ボタンは除く；c) 「自由回答」は除く。

ここで＜調査1＞について，ページごとの基本統計量を求めた（表6）。また箱ひげ図でページごとの回答所要時間の分布を示した（図4）。これを見ると，回答所要時間の分布はかなり特徴的である。＜調査1＞に限らず，過去に行ったほぼすべての調査において，回答所要時間の分布には共通した傾向がある。

- ・ 裾が非常に長い大きく歪んだ単峰分布となる。
- ・ 分布の特徴は1つの統計量では測れない。平均値（図内折線）はほぼ過大評価となり中央値との乖離が大きい。中央値の利用を勧めることが多いが十分ではないことがわかる。
- ・ 変動が大きい。標準偏差のほか，歪度，尖度，四分位範囲などを総合的にみる必要がある。
- ・ 歪度が大きく非対称となり，尖度も極端に大きい。四分位範囲が小さく，限られた範囲に値が集中している。ちなみに＜調査1＞のプリテストによる予想総回答所要時間は15分であるが，実際の総回答所要時間の平均値は約14分，中央値は約11分であった。
- ・ 変動係数にみられるように，相対的な変動も非常に大きい。
- ・ 対数や逆数の利用も考えられるが，それでも歪みが大きい分布となる。少なくとも正規分布や対数正規分布のような分布にはならない。

- ・ 回答者がページ内情報を確認する負担が重いと回答者変動が大きくなりはずれ値が増える傾向にある。

＜調査1＞の回答所要時間は、「調査の前半から中間あたりの回答負荷が大きい」「page05は、質問数が多く選択肢数も多くクリック数が増える、自由回答質問があることで変動が大きく回答負荷が大きい」「全体に回答者間の回答行動の変動が大きい」などの傾向がみえる。

表6 <調査1>のページごとの基本統計量の要約

統計量	page01	page02	page03	page04	page05	page06	page07	page08
有効測定値数 (人) ^{a)}	573	570	570	570	569	569	567	566
欠測値の数 (人)	9	12	12	12	13	13	15	16
はずれ値の数 (人) ^{b)}	50	37	37	34	49	28	34	65
はずれ値の割合 (%)	8.7	6.5	6.5	6.0	8.6	4.9	6.0	11.5
平均 (分)	0.159	1.951	1.684	2.489	5.312	0.545	1.553	0.700
中央値 (分)	0.100	1.550	1.233	1.842	3.633	0.483	1.250	0.200
標準偏差 (分)	0.224	1.946	2.163	2.683	5.234	0.350	2.257	1.517
歪度 ^{c)}	9.175	8.273	8.732	7.038	3.146	5.437	16.874	5.185
尖度	116.428	87.431	99.382	68.614	13.780	45.304	342.958	34.231
変動係数 (%)	141.085	99.707	128.451	107.774	98.517	64.295	145.312	216.825
範囲	3.433	24.233	32.217	35.100	47.383	4.183	48.800	15.717
四分位範囲	0.100	0.925	0.817	1.538	3.883	0.250	0.667	0.554

a) 「有効測定値数」は、そのページ内の有効回答者数、「欠測値の数」は、そのページで測定されない人数（例：中断、スキップなど）、両者を加えると有効回収回答数（582人）となる。

b) はずれ値の数は、 $[1.5 \times \text{四分位範囲}]$ の外に出た測定値数を計数（JMPの「分位点範囲のはずれ値」を利用）。「はずれ値の割合」は、 $[\text{はずれ値の数} \div \text{有効測定値数}]$ の割合。

c) 歪度、尖度は単位なし、正規分布のときに0となるように調整。

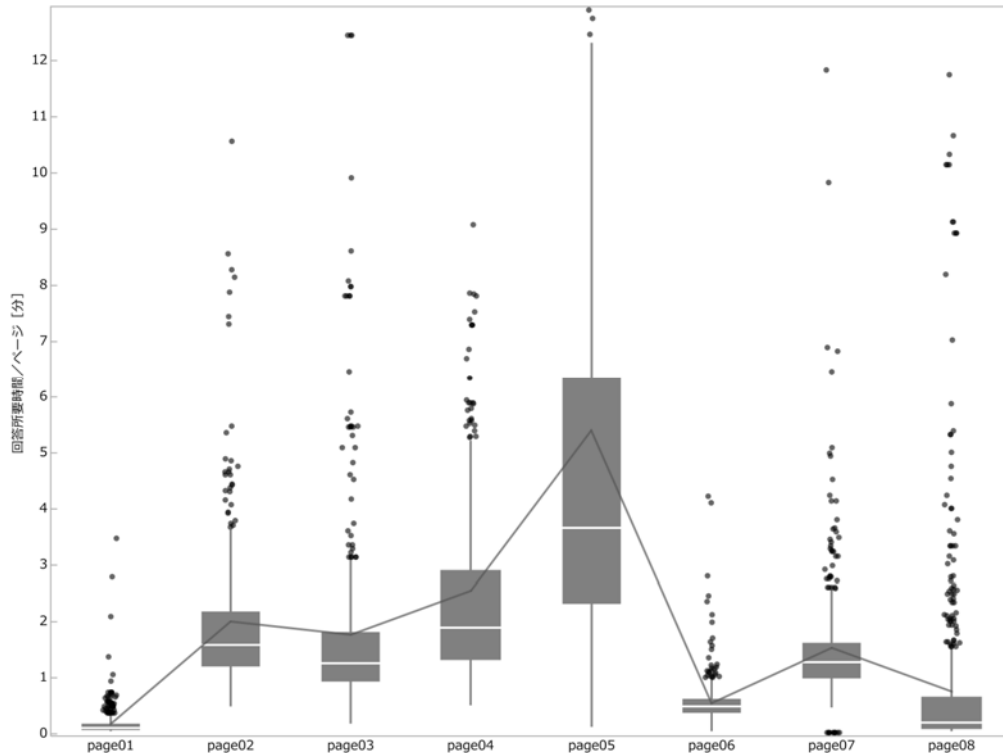


図4 <調査1>のページ別の回答所要時間の分布

[注：分布の裾が非常に長くなるので、値の大きい上の方をカットし縮めた]

以上から、個々の回答者が、各質問項目に対してどのような回答行動をとったか、たとえば、回答所要時間が非常に短いスピード者と、反対に回答に異常に時間を要した回答者とは、回答内容が異なるのかどうかなど、次の段階の分析が必要である。たとえば、回答者を回答パターンと回答所要時間で類型化し、質問項目の選択肢の選び方に初頭効果がみられるかを調べるなどが必要になる(前田ほか, 2007)。

「回答所要時間」を<調査2>についても調べよう(図5)。この調査は1ページに1問を配置したページが多い。回答所要時間の長いページの特徴を拾い出すと下のようになる。選択肢(ラジオ・ボタン、チェック・ボックス)の数が多い、自由回答の書き込みなどで回答所要時間が増え変動も大きい。ここでは、中央あたりと後半の自由回答、人口統計学的項目の質問で回答負荷が大きくなっている(調査票設計に若干無理があったかもしれない)。

- page07：選択肢数の多いグリッド形式のラジオ・ボタンとチェック・ボックス、それにテキスト・フィールドがある。
- page08：ここは「電子書籍と電子書籍端末」についての説明文があり、これを読んで、ラジオ・ボタンとチェック・ボックスから選択肢を選ばねばならない。
- page13：ここは、1つ前にある質問「今後の技術の進歩にともない、“紙の書籍”の時代が終わり、次第に“電子書籍”の時代になるだろう」という意見があります。あなたはこれについて、どのようにお考えですか。」(6つの選択肢)に対して、選んだ回答選択肢に対する意見を自由回答として記入する。よってかなり回答負荷が高くバラツキも大きい。
- page14~16：人口統計学的項目を設けたが、ページあたりの質問数が多くなったため、若干回答所要時間が長くなる。

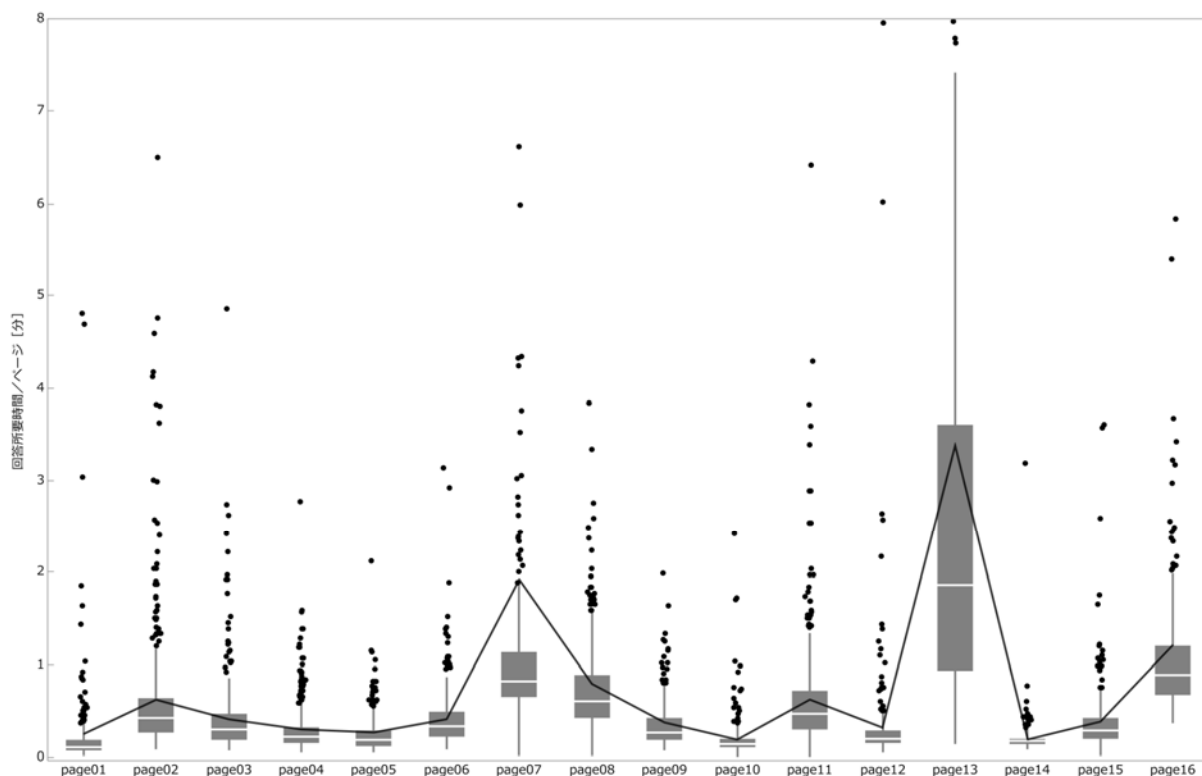


図5 <調査2>のページ別の回答所要時間の分布

<調査3>についても、ある傾向がみえる。これは<調査2>と類似の調査課題「本と読書、電子書籍」を扱っている。<調査2><調査3>は、調査時点は異なるが同じウェブ・パネル（非公募型）を用いた調査である。<調査2>に比べてページ数が多く、自由回答、分岐とパイピング処理、テキスト・フィールドへの入力などの煩雑な操作が多く、回答負荷が高い。

この<調査3>の page16 には、<調査2>で用いた自由回答質問（前述）とまったく同じ質問文を用いている。ここで、自由回答質問の回答所要時間の平均値と中央値を比べると、<調査2>（ $n=358$ （人））が、平均値=3.4（分）、中央値=1.9（分）、四分位範囲=2.7（分）、<調査3>（ $n=1,635$ （人））が、平均値=2.9（分）、中央値=1.7（分）、四分位範囲=約 2.5（分）であった。調査票の質問量や複雑さなどを考慮しても、両者は意外と類似した傾向にあるといえる。

5.5 回答所要時間の回答推移パターンの観察

同じ情報の見方を変えて、回答完了となった個々の回答者の各ページの回答所要時間の遷移の傾向をみる「回答推移パターン」として調べた。各ページについて全回答者について、回答者ごとに回答所要時間の回答推移パターンを描き（全体の傾向を観察）、その中で総回答所要時間が「10分未満の233名」（約40%）と「40分以上を要した15名」（約2.6%）の2グループの線を強調して描いた図が、図6である。

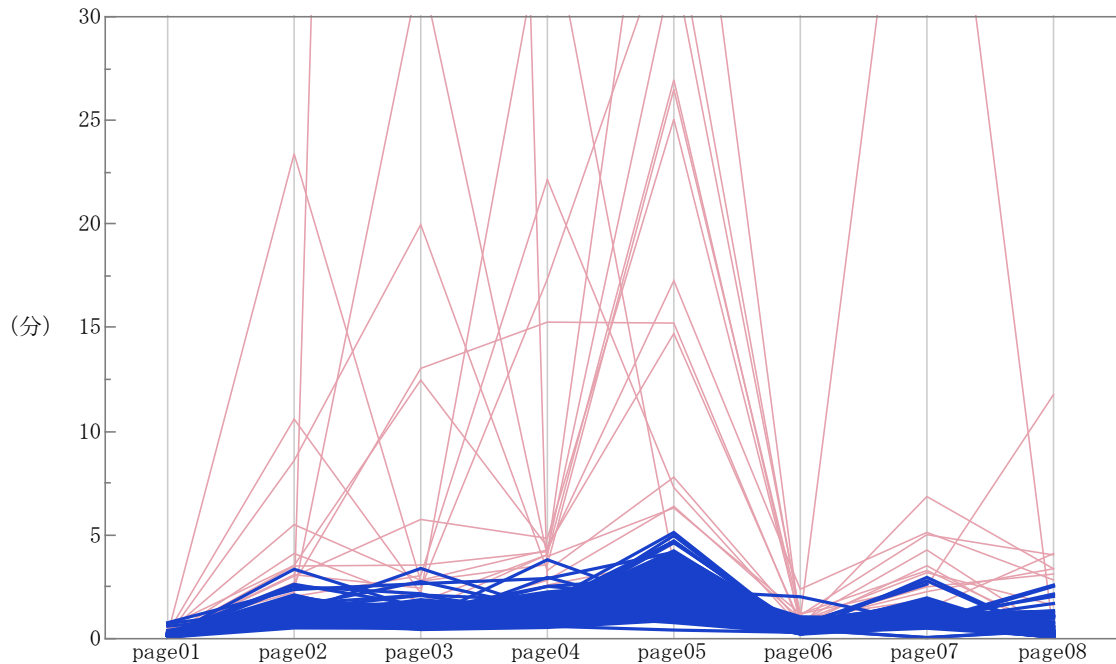


図6 回答を「10分未満で終えた233名」(下方濃い色)と「40分以上を要した15名」(上方薄い色)の回答推移パターン

回答所要時間が短い「10分未満」のグループ(濃い色)は、若干の例外はあるが回答パターンが総じて類似し全体に変動が小さい(回答所要時間が短い範囲にある)。つまり各ページをほぼためらいなく回答を終えている。一方、回答所要時間が長いグループ(薄い細い線)は、どこかのページで回答に戸惑った、たとえば「page05」でかなりの回答所要時間を要した人、他のページでも時間のかかった人など、さまざまなパターンが観察される。

同じ調査内容でも回答者の回答行動には違いがあるはずで、その違いをパラデータから得た回答所要時間のパターンで観察することもできる。しかし、こうした行動の違いが、回答者が実際に選んだ回答選択肢の分布にどう影響したのか(しなかったのか)、まではわからない。もう一步踏み込んで、たとえば、回答所要時間をいくつかの層に分け(たとえば、5分刻み、10分刻み)、これと各質問項目との関連分析を行うことが考えられる(前田ほか、2007)。

パラデータの利用価値は、回答者の回答行動と回答データ(調査データ)を併せた分析が可能となることで、従来の調査データだけの回答傾向の探査から一步踏み込んだ別の視点からの分析を可能とする。たとえば、次のような場面でパラデータの利用が考えられる(こうした分析も行ってきたが紙幅の都合で略す)。

- ・ 回答の手抜き現象(労働最小化行動など)の可能性、あるいは質問文をきちんと読み込んで、回答するか。
- ・ 回答拒否や無回答、回答中断はどう関係するのか。たとえば、自由回答質問があると回答中断が多めになるなどの現象はある。
- ・ 「微妙な質問」で、回答選択肢の選び方にためらいがあるか、あるいは簡単に選ぶのか。
- ・ プロービングや警告メッセージが必要かどうかの判断の目安。たとえば、スピーダーに近い回答者には、良く読んで回答するように促すメッセージを出す(筋金入りスピーダーには効果がないとい

う報告もある；Tourangeau et al., 2013）。

- ・ ストレートライニングや回答のスキップが生じたら、同じようにメッセージを出す。

このプロービングやプロンプトを画面に表示する手法は、労働最小化行動を低減するという報告もある（Tourangeau et al., 2013）。しかし、パラデータから得られる指標にもとづき、あらたな分析用変数をどのように作り、それをどのような質問文と組み合わせて分析すればよいか、あるいは適切なモデル化をどう行うかなどは模索の段階にあり今後の研究課題である。たとえば回答所要時間の区分化1つを考えても、上で「5分刻み、10分刻み」としたが、実際には難しい操作である。最近の研究に、回答所要時間の分布を5%刻みで区分し、「短い（上位5%）」「長い（下位5%）」「普通（中間90%）」と要約し、これと「回答不良」（スライダ形式、グリッド形式の質問で、ストレートライニングが観察された割合）との間に有意な関係がみられたという興味ある報告がある（埴淵ほか，2015）。一方、われわれの調査では、ストレートライニングが顕著に現れた例は皆無ではないが少ない（前田ほか，2007；大隅，2010）。とくに、調査票の質問量、質問の内容や難易度、予想回答所要時間との関係で、どのような区分がよいのかを決めることはそう簡単ではない。

6. おわりに — 今後の課題

ウェブ調査という調査方式の検証や取得調査データの質の評価や改善に、パラデータがどう利用できるのか、その考え方、利用可能性、具体的な取得方法、適用場面について、簡単な例を用いて紹介した。実際にはさらに踏み込んだ分析、とくにパラデータの利用が有効な、得られた調査データの回答傾向分析、人口統計学的変数との関係、回答者のPCスキルと回答との関係、そして総調査誤差の低減にどうつながられるのかなど、重要な課題に言及することが必要である。パラデータを巡る話題は、すでに調査方法論の重要な研究課題の1つとなっている。同時に、現時点では、以下のような問題点も指摘されている。

- ・ 調査の質の改善にどのように利用できるのか、パラデータから調査の質の改善に寄与するような“具体的な指標”を探ること。
- ・ パラデータの測定操作が、調査過程に支障をもたらす懸念。たとえば調査システムの負荷が増大し、システムの不具合や誤回答の増加につながるようなことがあっては利用する価値がない。そういう懸念はないのか。
- ・ つまり、パラデータの質と実際の調査データの質の確保との間のトレードオフの検証が必要であること。たとえば、パラデータ測定を行うことによる調査データの質に対する対投資効果の検証など。
- ・ パラデータの誤差特性の検証、とくに総調査誤差（TSE）との関連を調べること。
- ・ パラデータ取得から分析に至る一貫した処理に必要な“効率的なツールの構築”が必要であること（これが意外と面倒）。とくに、諸指標の算出・分析を容易にするツールの開発。

情報通信技術の著しい進歩により調査用具の利用環境の多様化が著しい。たとえば、従来はおもにPC対応でウェブ調査を行っていたが、これがスマートフォンやタブレット端末などに移行し、回答環境の変化が大きく、無回答誤差や測定誤差の観察が益々むずかしくなっている（Callegaro, 2015）。

さらに看過できない課題に「倫理上の問題」がある。回答者行動を電子的に追跡することは、技術面ではかなりのことが可能である。しかし、“どこまで測定してよいか”，“回答者に対するインフォームド・

「コンセンストはどうあるべきか」といった問題には慎重な対応が求められる（たとえば、Singer & Couper, 2011；ここに、LISS パネルを用いて、パラデータの取得を回答者がどう考えるかを調べた興味深い実験調査報告がある。この他、Couper & Singer, 2013）。

われわれも、当初からこの倫理問題にどう対処するかを念頭に、実験調査を慎重に進めてきた。通常は、調査協力者には調査実施時に、得られた回答データの扱いについては個人情報の守秘を含めて、説明を行っている。ウェブ調査の場合も同様に、調査依頼時に類似の対応を行っている（登録時や挨拶文の中で説明するなど）。しかし回答時の回答行動の記録を行っていることまで、説明してはいない。われわれの実験調査では、技術的な制約もあるが、収集データの内容を限定し、主に SSP（サーバー側パラデータ）について収集分析するように努めた。これをさらに拡大して CSP（クライアント側パラデータ）までも詳細に集める場合には、回答者に対して何らかの説明を行うことが必要になるであろう。Couper（2017）にあるように、回答者にどこまでを知らせるかの見極めが必要である。このことは、ウェブ調査に限らず、電子的データ収集方式を用いる他の調査方式（CATI、CAPI など）にも通底することである。

パラデータに関する国内の研究は活発とはいえない。とくにウェブ調査の場合は、つぎに挙げるような理由で、現実にはパラデータの取得、分析が次第に困難になっていると思われる。

- ・ 欧米と異なり、学術的調査に適したウェブ・パネル、とくに確率的パネルが少ない。しかも、実験調査の機会が得られにくい（利用者・委託者ではあっても、調査方式自体の評価分析は行いにくい）。
- ・ 数少ない非公募型ウェブ・パネル（部分的に確率的パネル）も、次第になくなる傾向にある。
- ・ 多くは商用の公募型パネルであり、実際に多くの研究報告でもこれを用いている。商用パネルでは、パラデータ取得のためのツールを用意することに、協力が得られにくいであろうこと（日常的にパラデータを利用していないし、調査システムに測定ツールを組み入れることに消極的であろう）。
- ・ 上述のように、倫理上の問題も考慮せねばならない。

パラデータに関する欧米の研究は非常に進んでいる。これに比べて、本稿はいかにもつたない内容となったことは否めない。また、ウェブ調査に関しては、日本国内では類似の研究報告をみる機会もほとんどない。しかし、情報量は少ないものの欧米の研究とは微妙に異なる知見も得られている。わずかな証拠しかないが、おそらく、国の違いによるパネル登録者の特性、回答者の性向の違いなども関連すると思われる（回答傾向が国によって異なる可能性）。このようなことから、パラデータの研究はもとより、公募型パネルの特性を、さらに丁寧に検証することが喫緊の課題である。もっとも懸念することは、学術研究で、ウェブ調査を単に便利な調査ツールとして、安易に使うことである。「易きにつく」は簡単だが危険も大きい。まずは調査方式としての特性を詳しく調べ、利点・欠点を知った上で慎重に用いることが肝要である。

謝辞

本稿の執筆にあたり、Mick P. Couper 氏（ミシガン大学）、小野裕亮氏（SAS Institute Japan, JMP ジャパン事業部）、樋口耕一氏（立命館大学）、山田一成氏（東洋大学）、前田忠彦氏（統計数理研究所）の各氏には、貴重なご助言、資料のご提供などをいただきましたこと、厚く御礼申し上げます。また、これまでの実験調査の実施にあたって、ご協力をいただいた調査機関、関係者の皆さまに、この場をお借りしてあらためて厚く謝意を表します。

【欧文文献】

AAPOR (2016), *AAPOR: Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys* (Revised 2016).

http://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf

Bethlehem, J. and Biffignandi, S. (2012), *Handbook of Web Surveys*, John Wiley & Sons.

Callegaro, M. (2015), Using Paradata to Better Interpret Online Experiments and Non-experiments Examples with Web Surveys, *Workshop on innovations in online experiments*, Nuffield College, Oxford March 13, 2015.

https://cess-web.nuff.ox.ac.uk/files/ioe_slides/Slides_Callegaro.pdf

Callegaro, M. and DiSogra, C. (2008), Computing Response Metrics for Online Panels, *Public Opinion Quarterly*, 72, 1008-1032.

Casper, A. and Couper, M.P. (1997), Using Keystroke files to Assess respondent Difficulties with an Audio-CASI Instrument, *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 239-244.

<http://www.amstat.org/sections/srms/Proceedings/>

Couper, M.P. (1998), Measuring Survey Quality in a CASI Environment. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 41-49. <http://www.amstat.org/sections/srms/Proceedings/>

Couper, M.P. (2000a), Web surveys: A Review of Issues and Approaches. *Public Opinion Quarterly*, 64, 464-494.

Couper, M.P. (2000b), Usability Evaluation of Computer-assisted Survey Instruments, *Social Science Computer Review*, vol.18, 4; 384-396.

Couper, M.P. (2005), Technology Trends in Survey Data Collection, *Social Science Computer Review* 23 (4), 486-501.

Couper, M.P. and Lyberg, L. (2005), The Use of Paradata in Survey Research, in *Proceedings of the 55th Session of the ISI, Sydney, Australia*.

Couper, M.P. (2008): *Designing Effective Web Surveys*, Cambridge University Press. [紹介記事：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, 50-60.]

http://www.wordminer.org/wp-content/uploads/2013/04/201_0.pdf

http://www.wordminer.org/wp-content/uploads/2013/04/237_2.pdf [抄訳]

Couper, M.P. (2010), “Online Research in the U.S. - Yesterday, Today, and Tomorrow-”, 第 41 回 JMRA トピックスセミナー資料：米国におけるオンライン・リサーチ —“きのう”, “きょう”そして“あした”—, 2010 年 9 月 21 日.

Couper, M.P. and Singer, E. (2013). Informed Consent for Web Paradata Use, *Survey Research Methods*, 7(1), 57-67.

Couper, M.P. (2014), *Introduction to Web Survey Paradata*, WebDataNet 2nd Training School: Paradata, Alexandroupolis, Greece, 1-3 October 2014.

Couper, M.P. (2017), Birth and Diffusion of the Concept of Paradata, *Advances in Social Research*, No. 18, *Advances in Social Research*, March, 2017. (*) 「パラデータ概念の誕生と普及」, 松本 渉 (訳), 「社会と調査」, 18号, 14-26.

Das, M., Ester, P., and Kaczmirek, L. (eds.), *Social and Behavioral Research and the Internet*. New York: Taylor and Francis.

Dillman, D.A., Smyth, J.D., and Christian, L.K. (2014), *Internet, Phone, Mail, and Mixed-Mode Surveys – The Tailored Design Method* (fourth edition), John Wiley & Sons.

Groves et al. (2004, 2009), “*Survey Methodology*”, John Wiley and Sons. (*) 2009年版は第2版。2004年版は大隅昇 (監訳) 「調査法ハンドブック」朝倉書店刊がある。

Heerwegh, D. (2004), Use of Client Side Paradata in Web Surveys, *Paper presented at the International symposium in honour of Paul Lazarsfeld* (Brussels, Belgium June 4-5 2004).
<https://perswww.kuleuven.be/~u0034437/public/Files/Heerwegh%20Uses%20of%20Client%20Side%20Paradata%20in%20Web%20Surveys.pdf>

Heerwegh, D. (2011), Internet Survey Paradata, in Das, M., Ester, P., and Kaczmirek, L. (eds.), *Social and Behavioral Research and the Internet*. New York: Taylor and Francis, 325-348.

Heerwegh, D., The CSP Project Webpage. *Technical Report October 20, 2011*.
<https://perswww.kuleuven.be/~u0034437/public/csp.htm> (*) CSP 試用プログラム, パラデータの誕生経緯などの説明があった (現在は, アクセスが制限されているようだ)。

Kaczmirek, L. (2009), *Human-Survey Interaction – Usability and Nonresponse in Online Surveys*, Herbert von Halem Verlag.
<http://kaczmirek.de/ucsp/ucsp.html> (*) ここで, Universal Client Side Paradata (UCSP)の内容が確認できる。

Kreuter, F. and Casas-Cordero, C. (2010), Paradata, Working Paper No. 136, RatSWD, *Working Paper Series of the Council for Social and Economic Data*.
http://www.ratswd.de/download/RatSWD_WP_2010/RatSWD_WP_136.pdf

Kreuter, F. (ed.) (2013), *Improving Surveys with Paradata -Analytic Uses of Process Information*, John Wiley & Sons.

Nicolaas, G. (2011), Survey Paradata: A review, *ESRC National Centre for Research Methods Review Paper*, National Centre for Research Methods. http://eprints.ncrm.ac.uk/1719/1/Nicolaas_review_paper_jan11.pdf

NORC (2015), AmeriSpeak[®], <http://www.norc.org/Research/Capabilities/Pages/amerispeak.aspx>

Schonlau, M. and Toepoel, V. (2015). Straightlining in Web Survey Panels Over Time, *Survey Research Methods*, 9(2), 125-137.

Singer, E. and Couper, M.P. (2011), Ethical Considerations in Internet Surveys, in Das, M., Ester, P., and Kaczmirek, L. (eds.), *Social and Behavioral Research and the Internet*. New York: Taylor and Francis, 133-162.

The Advertising Research Foundation (2009) , *The Online Research Quality Council Information and Data Results*,
(*) ARF (米国広告協会) が組織した“The Online Research Quality Council (ORQC)”が行った調査。 17 の
オンライン・パネルと、電話・郵送調査の比較実験調査。

Tourangeau, R., Conrad, F.G., and Couper, M.P. (2013), *The Science of Web Surveys*, Oxford University Press.
(*) 「ウェブ調査の科学」として翻訳版が刊行予定 (朝倉書店)。

Willems, P., Van Ossenbruggen, R., and Vonk, T. (2006), The Effects of Panel Recruitment and Management on Research Results; A Study Across 19 Panels. *Proceedings of ESOMAR World Research Conference, Panel Research 2006*, Barcelona, Spain, 79-99.

West, B.T. (2011), Paradata in Survey Research, *Survey Practice*, Vol. 4, No. 4.
<http://www.surveypractice.org/index.php/SurveyPractice/article/view/112/html>

【邦文文献】

Couper, M. P., 2017, "Birth and Diffusion of the Concept of Paradata," 「パラデータ概念の誕生と普及」, 松本 涉 (訳), 「社会と調査」, 18号 (2017年3月), 14-26.

埴淵知哉, 村中亮夫, 安藤雅登 (2015), インターネット調査によるデータ収集の課題—不良回答, 回答時間, および地理的特性に注目した分析—, E-journal GEO, Vol. 10(1) 81-98.

https://www.jstage.jst.go.jp/article/ejgeo/10/1/10_81/_article/-char/ja/

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—, 日本行動計量学会, 第38回大会, 予稿集, 30-33.

<http://wordminer.org/wp-content/uploads/2010/09/j00080a.pdf>

林文・吉野諒三編 (2011), 伝統的価値観と身近な生活意識に関する意識調査報告書—郵送調査と各調査期間のWEB調査の比較—, 統計数理研究所.

樋口耕一, 中井美樹, 湊邦生 (2012), Web調査における公募型モニターと非公募型モニターの回答傾向—変数間の関連に注目して—, 立命館産業社会論集, 2012年12月, 第48巻第3号, 95—103.

前田忠彦, 大隅昇 (2007), 自記式調査における実査方式間の比較研究—ウェブ調査の特徴を調べるための実験的検討—エストレーラ 2006年2月号, 12-19.

http://wordminer.org/wp-content/uploads/2013/04/207_10.pdf

前田忠彦, 中谷吉孝, 横田有一, 中田清, 中島一郎, 上嶋幸則, 大隅昇 (2007), Web調査方式による複数パネル間の比較実験, 日本行動計量学会第36回大会, 予稿集, 237-240.

<http://wordminer.org/wp-content/uploads/2013/04/65c356f4b3ad08490abe3e0cf5b33c91.pdf>

前田忠彦, 大隅昇, 中谷吉孝, 上嶋幸則, 渡會隆, 簗原勝史, 野田善治 (2005), 自記式調査における実査方式間の比較研究, 日本行動計量学会第33回大会, 予稿集, 256-259.

http://www.wordminer.org/wp-content/uploads/2013/04/207_11.pdf

松本 涉 (2017), データ取得プロセスの分析から調査を改善する, 「社会と調査」, 18号 (2017年3月), 5-11.

大隅昇ほか (2000), 「調査環境の変化に対応した新たな調査法の研究」報告書, 文部省科学研究費, 特定領域研究「統計情報活用のフロンティアの拡大」(略称: ミクロ統計データ), 研究計画 A02 班 (公募研究) 「ミクロデータ利用の社会的制度の問題点」(課題番号: 09206117, (平成9年度, 10年度).

<http://srdq.hus.osaka-u.ac.jp/ohsumiF/index.htm>

大隅昇ほか (2004), 「調査環境の変化に対応した新たな調査法の研究」報告書, 文部科学省リーダーシップ支援経費 (平成13年度, 平成14年度, 平成15年度).

<http://srdq.hus.osaka-u.ac.jp/ohsumi/index.htm>

大隅昇, 前田忠彦 (2008), インターネット調査の抱える課題 —実験調査から見てきたこと—, 「よろん」日本世論調査協会報, 第 101 号, P79-94.

http://www.wordminer.org/wp-content/uploads/2013/04/203_2.pdf

大隅昇, 前田忠彦 (2008), インターネット調査の役割と限界, 日本行動計量学会第 36 回大会, 35 周年記念シンポジウム「社会調査の現状と課題」, 予稿集, 197-200.

http://www.wordminer.org/wp-content/uploads/2013/04/203_1.pdf

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, 特集: 電子的調査情報収集法の動向—インターネット調査/オンライン調査, エストレーラ, No.143, 2-11.

http://www.wordminer.org/wp-content/uploads/2013/04/207_0.pdf

大隅昇 (2010), ウェブ調査とはなにか?—可能性, 限界そして課題—, 「市場調査」, 284 号 (2010 年 No.1), 4-19; 285 号 (2010 年 No.2), 2-27.

http://www.wordminer.org/wp-content/uploads/2013/04/193_0.pdf

大隅昇, 鳩真紀子 (2012), 「総調査誤差」をめぐって—ロバート M. グローヴス, ラース・ライパーグ論文「総調査誤差—過去、現在、未来—」を中心に—, 「よろん」, 2012 年, 110 号, 18-31.

http://wordminer.org/wp-content/uploads/2013/04/228_0.pdf

総務省 (2011), 「平成 22 年度 新 ICT 利活用サービス創出支援事業 電子書籍交換フォーマット標準化プロジェクト」調査報告書 (2011 年 3 月), (*) 第 6 章「電子書籍及び電子書籍端末の利用環境調査」としてウェブ調査と郵送調査の比較検証 (pp.475-586).

http://www.soumu.go.jp/menu_seisaku/ictseisaku/ictriyou/shinict.html,

http://ebformat.jp/dl/koukan_format_houkoku_2011_05.pdf

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, 日本行動計量学会, 第 38 回大会, 予稿集, 26-29.

<http://wordminer.org/wp-content/uploads/2010/09/j00079a.pdf>

<補足情報>

注1：回答率と参加率

回答率と参加率は異なる。回答率（RR：response rate）にはいくつかの種類がある（たとえば、AAPORはRR1からRR6までの6種の算出方法を挙げている）。回答率は算出に用いる項目の定義によっては完了率（completion rate）とほぼ同じ意味で用いることもある。ウェブ調査で非確率的パネル（ボランティア・パネル、公募型パネル）を用いるとき、回答率の算出がむずかしいので（例：インターネット母集団があまりであり、標本抽出枠が調査時に正確に確定できないなど）、参加率（participation rate）を用いることが勧められている（AAPOR2016；ISO20362；Bethlehem, 2012；Callegaro and DiSogra, 2008）。しかし日本国内の多くの商用サイト（公募型パネル）で、この参加率が開示されることはほとんどない。

われわれの実験調査では、当然インターネット母集団は明確ではないが、標本抽出枠に相当する登録者集団全体（ウェブ・パネル）の大きさ（パネル・サイズ）と電子メール・リストの情報を調査実施時点で把握していることを前提に、そのウェブ・パネルから無作為抽出により計画標本を用意したので“回答率”を用いることにした。

注2：ウェブ調査実施時の回答制御

「回答制御」とは、回答時に調査システム側から回答者に対して、さまざまな制約を設けることをいう。これを行えることがウェブ調査の特徴の1つである。「回答制御なし」は、まったく制限を設けない場合で、（回答期限内に）いつでも自由に回答できる、ページ間の遷移（行き来）は自由、回答選択肢の選び方（選ぶ、選ばない）は自由。つまり、ラジオ・ボタンやチェック・ボックスなどに応答（クリック）があってもなくてもよい。これは、郵送調査や留置自記式の回答に近い。こうした条件から、以下のように段階をふんで、制約を付け加えることができる。

- ・ 回答選択肢の無回答の選択は可とする、あるいは無回答となってもよい（選択肢をクリックしなくてもよい）。
- ・ ページ間の往復（行き来）は可とする場合と、これを抑制する場合、つまりページ間の行き来は許さないとする場合。
- ・ 回答選択肢には、かならずクリック／チェックを入れないと、先には進めない。
- ・ 同一回答者の二重の回答・重複回答は許容しない（つまり、質問を見直して再回答することを抑制する）。
- ・ ページは、一方向にのみ進行可能で（前進のみで戻りを許さず）、また、各質問の選択肢には、かならずクリック／チェックを入れないと、先には進めない、分岐質問なども同じように制御する。つまり回答制御は、すべてのことに制約を付ける、とするとき。

回答制御の有無は、無回答（調査不能、項目無回答）や回答中断などに影響するので、調査実施時にどのような回答制御条件を設定したかを把握することは、のちの分析に影響する。当然、他の調査方式との比較実験調査でも注意が必要である。多くの商用ウェブ・パネルは完全な回答制御を行っているものと考えられる。

注3：国内におけるウェブ調査の参加率

国内の商用ウェブ・パネルでは、回答率あるいは参加率をほとんど開示していないので、正確なことはわからない。最近では参加率の低下が著しいという情報もある。ある非公式情報では、参加率の算出方法にもよるが、数%～十数%程度とある（1桁台が多い）。とくに、若年層の参加率が低下しているともいわれ

る（スマートフォンの普及、PC の利用頻度の低下等が理由のようだが、この点では欧米の研究の指摘にも合っている）。

注 4：「労働最小化行動」（“satisfice”, “satisficing”あるいは“satisficing behavior”）

ここでは“satisficing behavior”に、「労働最小化行動」の訳をあてた。「最小限化行動」「最小限回答行動」などとする例もある。Couper（2017）では、この労働最小化行動をとる人を“satisfier”と記している。

回答者が、正しい回答を提供し得る努力を怠る、あるいは正しい回答を用意できないとき、結果として、回答者は、最小努力で自分の気に入った回答を提供しようと努めること。「ストレートライニング」は、この労働最小化行動の 1 つの例である。

注 5：調査方式について

調査方式（survey mode；調査モード）とは、調査データを収集するための具体的な手段のことをいう。単にモード（mode）と記すこともある。ここで、調査方式をいくつかの要素をもとに分類する。たとえば、つぎの観点から整理すると分かりやすい。これを目安に要約分類を行うと下表のように要約される。

- ・ 「情報伝達経路」が、視覚的か聴覚的か
- ・ 「コンピュータ支援」があるか、ないか
- ・ 「調査員の関与」があるか、ないか
- ・ 「回答記入の方法」が、非自記式（他記式）か、自記式か

表 調査方式の分類例

情報伝達経路	調査方式 (調査モード)	コンピュータ支援(CA: computer-assisted)の有無			
		CAなし		CA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
主に聴覚的な 情報伝達	面接	面接 (調査員による対面 面接)	…	CAPI	CASI Text CASI Audio CASI Video CASI
	電話	電話 (コンピュータによら ない電話調査)	ファクシミリ	CATI TDE IVR/T-ACASI	…
主に視覚的な 情報伝達	郵便	…	郵送(調査票) 訪問留置 ^(†)	…	DBM (Disk by Mail)
	インターネット (オンライン)	…	…	OFG	ウェブ調査 電子メール調査

(†) 「訪問留置」(drop-off mode, drop-off/pick-up survey)を入れるべき適切な位置が見つからないので、ここに置いた。まず、調査員が調査対象を訪問し調査票を“配布”する。続いて、一定期間を経過後に、ふたたび調査員がその調査対象を訪問し、回答済みの調査票を“回収”する、あるいは“郵送で返送”する、そのような含意でここに置いた。

なお、表内にある英文字略記の綴りとその訳を下に記した。

CASIC : Computer Assisted Survey Information Collection コンピュータ支援による調査情報伝達情報収集
CADAC : Computer Assisted Data Collection コンピュータ支援によるデータ収集
CAI: Computer-assisted Interviewing コンピュータ支援の面接調査 (面接聴取による場合)
CATI: computer-assisted telephone interviewing 通常, RDD 方式のサンプリングを利用した電話調査 (電話聴取) のこと。CAPI の電話調査版。
CAPI : computer-assisted personal interviewing コンピュータ支援の個人面接方式
CASI : computer-assisted self-interviewing コンピュータ支援の自答式
TCASI : Text CASI (T-ACASI)
ACASI : Audio CASI (聴覚的な CASI)
VCASI : Video CASI (視覚的な CASI)
IVR : Interactive voice response 音声自動応答装置による調査, ACASI の電話版, TCASI に同じ。
OFG : online focus group オンライン・フォーカス・グループ
DBM : Disk by Mail ディスク・バイ・メール (フロッピー・ディスクに調査票を書き込み, 調査依頼を郵送, 回答を回収する方式)
TDE: : touchtone data entry タッチトーン式データ入力

参考 :

Macer, T. (2003). We Seek Them Here, We Seek Them There - How technical innovation in mixed mode survey software is responding to the challenge of finding elusive respondents -, Banks, R. (ed.), *Survey and Statistical Computing IV, The Impact of Technology on the Survey Process*.

Smith, T.W. (2011). Refining the Total Survey Error Perspective, *International Journal of Public Opinion Research*, 23 (4), 464-484.

Groves et al. (2004, 2009), “*Survey Methodology*”, John Wiley and Sons. (*) 第 5 章を参照。

注 : 2004 年版は, 大隅昇 (監訳) 「調査法ハンドブック」朝倉書店刊がある。これの第 5 章。