



特集

パラデータの 活用に向けて

調査データを取得するプロセスのデータのことをパラデータ (paradata) と呼ぶ。欧米では、このパラデータを利用して調査を改善するための研究が盛んになってきたが、日本でもパラデータを活用できないのだろうか。

本特集では、パラデータの命名者であるミック・P・クーパー氏に、パラデータ概念の誕生から今日までの発展についてご寄稿頂いた。その上でパラデータに以前から関心をお持ちだった識者の方々に、日本の調査でのパラデータの活用を述べて頂いた。

前田忠彦氏は、調査員がデータを取得するプロセスを記録したデータである訪問記録の分析例を示す。保田時男氏は、自らの分析経験を振り返りつつ、訪問記録分析の意義を自省的に検証する。城川美佳氏は、電話調査におけるパラデータ分析の経験からその利用の意義を述べる。大隅昇他4名の諸氏は、ウェブ調査におけるパラデータの活用方法と課題を論じる。

冒頭の松本論文では、これら5本の論文の理解を促せるように、パラデータの基本概念の概説と日本国内の研究状況についての展望を示している。できれば前から順に6本全ての特集論文をお読み頂き、パラデータの活用についてご考察頂きたい。

(松本 渉)



特集論文

1

データ取得プロセスの 分析から調査を改善する

松本 渉

関西大学総合情報学部 准教授

1 はじめに

回収率の低下が叫ばれて久しい。『社会と調査』第5号では特集「回収率を考える」が組まれ、篠木(2010)の基調論文で50年間の低下傾向が示されていた。1950年代には8割から9割を維持していた日本の代表的な社会調査の回収率は、2008年には5割から6割に低下したのである(篠木, 2010)。統計数理研究所が実施している日本人の国民性調査の回収率は、2008年で52%(中村ほか, 2009)、2013年で50%(中村ほか, 2015)であるから、回収率が低迷しているということは、現在でも日本の社会調査の特徴であることは確かである。

ところで、半世紀前の社会調査との違いを考えた時、現在の社会調査のもう1つの特徴として、CATI (computer-assisted telephone interviewing) やウェブ調査にみられる電子的なデータ収集方法の普及があげられる。電話調査のCATIは、1970年代の終わりに米国で普及し(鈴木, 2003; 大隅, 2004)、日本では1990年代に導入された(鈴木, 2014)。ウェブ調査は、市場調査領域でめざましく普及したが、公的統計の領域でも利用が進んだ。2010年の調査では東京都のみでウェブ上での回答が可能だった国勢調査は、2015年の調査では全国的にウェブ上での回答が可能となった。コンピュータ支援型の調査方式は、電話調査以外でも徐々に浸透し、2013年に『世論

調査の新しい地平——CASI方式世論調査』(日野・田中, 2013)が刊行され、2016年11月には社会調査協会においてもシンポジウム「コンピュータ支援型調査の課題と可能性」が開催された。電子的なデータ収集方法の普及は続いている。

回収率の低下と電子的なデータ収集方法の普及という2つを考えた時、回収率の低下、ひいては調査の質の低下の問題を電子的に取得されたデータの特徴を生かして解決できないかという思いに至るのも不思議ではない。電子的なデータ収集方法においては、データ取得プロセスがより正確に記録されるためである。実はこのような着想はすでに書籍化されている。2013年に刊行された書物 *Improving Surveys with Paradata: Analytic Uses of Process Information* がそれであり、同書ではパラデータと呼ばれる調査データ取得プロセスのデータを利用して調査を改善することを正面から扱っている。

本論文では、まずこのパラデータに関する研究を概観する。その際、日本での調査の状況を意識しつつ同書を読み解いていく。その上で、日本国内での研究実践例にふれつつ、日本国内のパラデータ研究状況を総括する。最後に、日本のパラデータ研究の発展のための課題を示す。

2 パラデータとは何か

第63回の米国世論調査学会の年次大会(2008年開催)は、パラデータ概念への注目を高めた

であろう。この大会において、Warren J. Mitofsky Innovators Awardの授賞対象とされたのは、Mick P. Couperの功績“Introducing Paradata as a Tool to Understand Respondent Behavior”（回答者行動の理解のためのツールとしてパラデータを導入したこと）であったからである。

ではパラデータとはそもそも何であろうか。パラデータとは、もともとコンピュータ支援型のデータ収集システムによって副次的(para-)に生成される調査過程についてのデータの一般的な概念を表すものとしてCouper (1998) が取り入れた用語である。しかし、パラデータ概念は、調査方法論の研究(例えば、Couper and Lyberg, 2005; Groves *et al.*, 2005; O'Reilly, 2009) で活用されていく過程で定義が拡張され、調査過程の諸側面や他の調査モードにおけるデータ収集場面にも適用され始めた(Kreuter *et al.*, 2010)。パラデータ概念はこのように変化しており、標準的な定義が固まっているわけではない。しかし、2013年に至って調査の改善のためにパラデータの分析を利用することを正面から扱った書物が発刊された。冒頭でとりあげた*Improving Surveys with Paradata: Analytic Uses of Process Information*である。同書の刊行は、パラデータ概念の明確化と体系化について一定の進展を反映している。

例えば、Kreuter (2013) は、同書の中でパラデータ概念と類似概念との違いを明確に指摘している。1つにはメタデータとパラデータの関係である。Kreuter (2013) は、メタデータは最終生成物についてのデータであるが、パラデータは最終生成物を生み出すプロセスについてのデータとしている。日本の調査関係者がいうところの調査不能(拒否や不在など)は、最終的な調査の状況として判明するものであり、回収標本サイズとともにメタデータに該当するが、これらが確定するまでに何度か確認する「一時的な不在」といった調査途上の情報はパラデータに該当することになる。そしてこの考え方では、パラデー

タはメタデータより広い範囲をカバーすることになる。またKreuter (2013) は補助データとパラデータの関係も述べている。Kreuter (2013) によれば、補助データとは調査データ収集の前に提供されるデータであるのに対し、パラデータは調査データの収集中に生成されるデータである。日本の調査研究者も地域情報等に基づいて層化抽出を設計することがあるが、地域の都市化の程度や人口規模といった情報は調査データ収集前に把握されている。これらは補助データということになる。

上記の記述は類似概念との差異を明らかにしているが、もう少し具体的に「パラデータとは何か」に踏み込んでみよう。Olson and Parkhurst (2013) は、測定誤差の評価に有益と考えられるパラデータの例として、①タイムスタンプ、②キーストローク(キーボードを押して離すこと)、③マウスクリック、④行動コード、⑤声の特徴、⑥調査員による評価の6つをあげている。ここでいうタイムスタンプとは個々の質問に対する回答など調査における何らかの動作の日時である。キーストロークやマウスクリックなど他の行動に付随して記録され、回答時間の分析に生かされる。行動コードは、調査員と調査対象者の双方の行動を記号化したものであり、調査員が正確に質問文を読み上げたかどうか、対象者が回答の前に質問を投げかけたりしていないかなどである。行動コードも声の特徴も録音に基づきデータ化される。最後の調査員による評価は、日本の調査でも少なからず実施されていると思われる。2008年に筆者が中心となって実施した面接調査「市民社会調査」(詳細は、松本・前田, 2008) でも回答者の「反応の変化はどうでしたか」を調査員に尋ねている。その結果は、545人のうち、「最初から最後まで好意的だった」355人(65%)、「だんだん好意的になった」137人(25%)、「だんだん迷惑そうになった」40人(7%)、「最初から最後まで迷惑そうだった」11人(2%)、「無回答」2人(0%)であった。実際に実施を終え



た面接についての各質問における無回答との関連性の分析は、質問文のワーディングや順序の改善方法を考える材料になりうる¹⁾。

Olson and Parkhurst (2013) が列挙した 6 種類のパラデータのうち⑥の「調査員による評価」以外は、CAI (computer-assisted interviewing) や CATI あるいはウェブ調査の実施が前提であり、データ収集プロセスがコンピュータ化されていないと取得しにくい。日本では CATI やウェブ調査が増えつつあるといっても、まだ PAPI (paper and pencil interviewing) が主流であり、イメージをもちにくいかもしれない。いずれにしても 6 種類とも具体的に聴き取りをしている際に得られるパラデータである。

これに対して、個々の聴き取りや回答が開始するという意味で調査が開始していないが、調査プロジェクト自体は開始しているという意味では調査のプロセスデータとして得られるパラデータもある。Kreuter and Olson (2013) が、調査不能による誤差を調べるのに有益なパラデータとしてあげたものであり、①コールレコード、②調査員による観察、③調査員と世帯構成員とのやりとりに関する測定 の 3 つが該当する。コールレコードは、個別訪問面接調査における調査員の訪問記録 (visit record / call record) や電話調査におけるオペレーターの架電記録 (call record) のことであり、日本の PAPI による社会調査でもなじみがある。調査員による観察は、前出の調査員による評価と紛らわしいが、調査対象者の住居構造や付近の状況などの情報のことである。対象者から回答が得られなくても入手できる点で、調査員による評価とは異なる²⁾。世帯構成員とのやりとりとは、個別訪問面接調査・留置調査・電話調査において調査対象者本人以外と接触した場合に生じる現象である。これら 3 つは、日本でもイメージが持ちやすいと考えられるパラデータといえよう。

3 日本におけるパラデータ研究の展開

手前味噌な話になるが、日本で調査用語としてのパラデータの利用が確認できるのは、2011 年 9 月の統計関連学会連合大会における「パラデータにみられる不整合について」(松本, 2011) が最初である。その後もパラデータという用語が紹介されることはあったが(例えば、松本, 2014; 大隅, 2014)、それほど知られた表現とはなっていない。またコンピュータ支援型のデータ収集法の実践例自体が少ないので、パラデータの研究の蓄積も乏しいのが現状である。しかし、前述したようにパラデータ の概念は拡張してきている³⁾。当初はコンピュータ支援型のデータ収集方法での生成が想定されていたが、今では必ずしもコンピュータ化された調査を前提としなくなってきた。個別訪問面接調査における調査員の (PAPI での) 訪問記録や電話調査における架電記録もパラデータとして位置づけられるだけでなく (Kreuter and Olson, 2013 参照)、調査員による訪問記録や調査員の観察した記録もパラデータとして扱われている (Kreuter and Olson, 2013; West and Sinibaldi, 2013)。

パラデータという用語の利用の有無はともかく、PAPI を前提とした調査員の訪問記録や電話調査におけるオペレーターの架電記録の研究実践例であれば以前から日本でも存在する。もう少し具体的なイメージを述べよう。通常、調査票調査で入手するのは回収標本の回答結果 (表 1) である。しかし、実際 1,000 程度の回収標本であっても、電話調査であれば 5,000 近くの電話番号に電話をかけていることがある。5,000 の電話番号別に回答が完了したのか拒否されたのかといった最終的な状況を示したのが表 2 である。さらに 5,000 の電話番号のうち 1 度の架電で拒否や回答完了が確定することはまれで、通常は複数回架電する。そのことを踏まえて全ての架電状況を示したのが表 3 である。面接や留

表1 回収標本の回答データ

LISTID	Q1	Q2	Q3	Q4	Q5	Q6	……
7	1	2	1	2	1	1	……
13	2	2	2	2	2	2	……
25	1	2	1	1	3	2	……
33	1	1	2	1	1	3	……
4990	1	2	1	2	1	3	……

表2 架電番号の最終的な調査状況

LISTID	年月日	時刻	回数	最終的な状況
1	2017/2/20	14:01:11	4	拒否
2	2017/2/17	14:01:12	1	拒否
3	2017/2/18	19:01:21	2	FAX
4	……	……	5	対象者不在
5	……	……	2	事業用番号
6	……	……	3	拒否
7	2017/2/17	14:02:21	1	回答完了
4990	2017/2/21	18:05:30	5	回答完了
5000	2017/2/21	……	5	コールのみ

表3 全ての架電記録

No.	年月日	時刻	LISTID	回数	架電状況
1	2017/2/17	14:01:11	1	1	コールのみ
2	2017/2/17	14:01:12	2	1	拒否(確定)
3	2017/2/17	14:01:20	3	1	FAX音
4	2017/2/17	14:01:21	4	1	留守電
5	2017/2/17	14:01:31	5	1	話し中
6	2017/2/17	14:01:40	6	1	コールのみ
7	2017/2/17	14:02:21	7	1	回答完了(確定)

.....		4990	1	コールのみ	
.....		4990	4	コールのみ	
15300	2017/2/21	18:05:30	4990	5	回答完了(確定)
.....		XXXX			
15333			5000	5	コールのみ(確定)

※筆者が模式的に表現したもの。実際の書き方は調査会社によって異なる。

置といった調査員が訪問する調査でも同様で、回収や拒否、不在などの調査状況が最終的に確定する前に、何回か不在等を繰り返すので複数回の訪問の記録が残される(表4)。日本国内でも、これらを分析して調査の質や効率を改善するための示唆をうる研究が試みられてきた。

例えば、城川(2003)は、RDDによる電話調査での架電記録から、日曜日、水曜日、土曜日の順で接触成功数が多いことを明らかにしている。保田(2005)は、訪問記録自体を扱ったものではないが、訪問回数、非回収の決定日時と理由といった訪問記録から導かれるのと同等の情報を調査員に記録させて回収率改善の方法を論じている。さらに保田(2009)は、JGSSの訪問記録から、接触成功率は日曜日だけが高く、協力獲得率は必ずしも休日が高いとはいえないことを明らかにしている。松本(2009, 2010)は、CATIによる米国調査の架電記録の分析から、日曜日にコール数を過大に割り当てたために日曜日の回

収効率の低下を招く可能性を示している。有坂(2010)は、40・50代の男性と20代の女性については日曜日が他の曜日よりも明確に接触率が高いが、他の性別×年齢の層については必ずしもあてはまらないことを指摘している。上記の研究は全て曜日・時間帯等が調査過程に与える影響を中心に扱っているが、保田ほか(2008)のように訪問記録の分析から謝礼を折半して前渡しと後渡しを両方することの効果を具体化したものもある。改善の対象は質であったり効率であったりと色々であるが、いずれも訪問記録や架電記録により調査の何らかの改善を目的となされたものである。まさしくパラデータの分析は、調査における誤差と費用を研究する手段として追加されるのである(Kruter, 2013)。

訪問記録や架電記録以外のパラデータの研究としては、近年に入って杉野・俵・轟(2015)のように調査への協力度や質問への理解度といった調査員による評価⁴⁾をパラデータとして意識



表4 訪問記録の概略

ID	1回目訪問	2回目訪問	3回目訪問	4回目訪問	5回目訪問	最終結果
1	9月4日	9月5日				9月5日
	不在	拒否				拒否
2	9月4日					9月4日
	回収					回収
3	9月4日	9月5日	9月6日	9月7日	9月8日	9月8日
	不在	不在	不在	不在	回収	回収
4	9月4日	9月5日	9月6日	9月7日		9月7日
	不在	不在	本人不在	回収		回収
5	9月5日	9月5日				9月5日
	不在	拒否				拒否
.....						

表5 日本での研究事例

	パラデータの種類	研究事例
回答を開始する前に入手するパラデータ	①コールレコード(訪問記録・架電記録)	城川(2003), 保田・宍戸・岩井(2008), 保田(2005, 2009), 松本(2009, 2010), 有坂(2010)
	②調査員による観察	松岡・前田(2015)
	③調査員と世帯構成員とのやりとりの観察	
回答を開始した後に入手するパラデータ(回答を開始してから全ての回答を終えるまでに入手するパラデータ)	①タイムスタンプ(⇒調査完了時間・回答時間などへ)	埴淵・村中・安藤(2015)
	②キーストローク	
	③マウスクリック	
	④行動コード	
	⑤声の特徴	
	⑥調査員による評価	杉野・俵・轟(2015)

的に分析した研究や、松岡・前田(2015)のように住居の特徴等の調査員による観察を拒否のパターンとの関連から分析した研究ができた。埴淵ほか(2015)によるウェブ調査における個々の回答者の調査完了時間を分析したものもある。ただし、調査完了時間は、Olson and Parkhurst(2013)が分類した4水準のタイムスタンプ(①調査の開始時刻・終了時刻, ②セクション単位のタイムスタンプ, ③質問単位のタイムスタンプ, ④キーストロークやマウスクリックごとの各時点)のうち①調査の開始時刻・終了時刻のタイムスタンプから導かれるものであるが、個々の質問項目ごとの回答時間のように文字通りの調査プロセスのデータではない。調査ごとに得られる集計データであることを考えると、メタデータとも解釈することができるので注意が必要である。

4 日本のパラデータ研究の課題

Olson and Parkhurst(2013)およびKreuter and Olson(2013)の分類を参考に、前節での筆者の記述を整理し、日本におけるパラデータ研究の状況を示したのが表5である。ただし、Kreuter and Olson(2013), Olson and Parkhurst(2013)では、調査不能誤差を調べるためのパラデータ、測定誤差の評価のためのパラデータと表現しているが、それぞれの実質的な状況を考慮して、ここでは前者を「回答を開始する前に入手するパラデータ」、後者を「回答を開始した後に入手するパラデータ」(回答を開始してから全ての回答を終えるまでに入手するパラデータ)としてみた。

このように整理すると、日本国内では「回答を開始した後に入手するパラデータ」の研究が手薄であることは明らかである。「調査員によ

る評価」を除けば、「調査対象者が回答を開始した後に入手するパラデータ」の多くは、電子的なデータの取得方法が前提となっているからである。パラデータを活用して測定誤差を改善するためには、コンピュータ支援型調査の本格的な普及が必要になる。

埴淵ほか(2015)のようにウェブ調査における回答時間を分析する研究もあったが、回答者ごとの調査全体の完了時間はメタデータとも解釈できるものであり、より本格的なパラデータの分析を行うには、質問ごとの回答時間のように、よりミクロなレベルでのタイムスタンプを取得する必要がある。実は筆者も2015年に詳細なタイムスタンプの情報を伴うウェブ調査の実施可能性をとある調査会社に打診したことがある。しかし、当該調査会社にこのようなデータを取得する備えがないばかりか、取得自体に積極的姿勢がなかったために、その時は詳細なタイムスタンプの入手を断念せざるをえなかった。調査会社が本格的なパラデータの取得に備えるようになるには、多くの研究者がパラデータに関心を持ち、調査会社がパラデータの取得に備えたいかなるような社会状況ができあがる必要があるであろう。

実は、コンピュータ化の重要性は、「調査対象者が回答を開始する前に入手するパラデータ」についてもあてはまる。回収率の改善には調査員の行動把握が不可欠であり、そのためには訪問記録の分析が有効な手段の1つになるが(保田ほか, 2008), PAPIにおける訪問記録はそれほど正確ではないからである(松本, 2011)。このことは、調査票にも調査日時を記録し、不能票であっても調査不能を確定した日時を記録するようにした上で、訪問記録の最終訪問日時と調査票は記載の調査日時を照合すれば、通常両者が完全に一致しないことから確認できる(松本, 2011)。筆者が過去に関わってきた面接調査(松本・前田, 2008)においても±1日のずれはよく見られたし、最大で2週間程度の日付の食い違

いが見られたこともある。

このような日時記録の不整合の問題が生じるのは、単なる書き間違いもありうるが、訪問記録票が単なる業務管理ツールとして位置づけられており、調査結果のデータと同様に重要なデータと位置づけられていなかったこと、また紙である以上、記入スペースに上限があったことに起因すると思われる。訪問記録を正確にするためには、訪問記録を分析対象として明確に位置づけると同時に、記録の物理的な制約をなくするためにコンピュータ化を進める必要がある。

5 結びに代えて

本論文では、パラデータの概念が調査データ取得プロセスで入手されるデータ全般という広い意味で理解されるようになってきたことを確認し、この意味でのパラデータの研究についての日本の状況を総括することを試みた。その結果、調査プロセスのデータといっても調査対象者が回答を開始するに至るまでに得られるパラデータの研究に比べ、調査対象者が回答を開始した後に入手するパラデータの研究が少ない日本の状況がはっきりした。そして、その不足を解消するためには、コンピュータを利用した調査が必要であることを確認した。

冒頭では、日本でも電子的なデータ取得方法による調査が普及し始めているようなことを述べたが、実のところ発展途上である。例えば、CASI (computer-assisted self-interviewing), CAPI, CAIなどに決まった訳があるわけではない。computer-assistedの部分については、「コンピュータ支援付」「コンピュータ支援の」「コンピュータ支援型」といった日本語が考えられるが、英語表記のまま利用している日本語文献が多い。

パラデータに関わる英語も和訳するには厄介な概念が多い。例えば、visit recordは、面接調査や留置調査など調査員が介在する調査における



訪問記録を意味する。しかし、call recordは、電話調査における架電の状況を記録したものでもあると同時に訪問記録(visit record)の意味も兼ねる。そのため文脈によって両者を使い分ける必要がある。両方を意味する場合はコールレコードとカタカナ表記せざるをえない。さらに電話調査においては、call timeの和訳は架電時刻となるが、発信者と受信者で時差がある場合は、発信時刻と受信時刻と異なるので、両者を区別したい場合は、そのように用語を使い分ける必要がある。また、英語のfinal dispositionは、日本語でいう拒否や不在などの「調査不能理由」を含む表現であるが、調査完了も含むので「調査不能理由」では全てを表したことはない。また、ひとつひとつの訪問や架電を意味する単語としてattemptがあるがこれも定訳がない。

日本語と英語の間での調査用語のミスマッチは、実はもっと基本的な概念にもあてはまる。例えば、インタビューである。カタカナ語のインタビューは取材のための面会として理解される可能性をもっている一方で、英語のinterviewは訪問面接調査における個々の面接にも用いられる。さらに電話調査における聴き取りにも用いられるため、面接調査であることが明らかであれば「面接」と訳せばよいが、そうでない場合はどちらでも通じるように訳す必要がある。電話面接という表現が通用するかどうかは微妙なところであるから、文脈によって聴き取りといたり面接といたりすることになる⁵⁾。また同

様に、interviewerも面接調査と電話調査の双方における調査員を指す。電話調査の場合は、日本ではオペレーターと呼ばれることが多いが、これは調査員と呼んでも意味は通じる。

筆者は、過去に「社会調査」自体が日本独自の概念として形成されたため、日本の定量的な社会調査が、欧米のsurveyの研究と、学問領域として必ずしも重なりあっていない可能性を指摘した(Matsumoto, 2007)。パラデータの研究においても、ガラパゴス化しないように国際的な研究動向に対する目配りを失ってはならないと思われる。

【謝辞】本論文中に登場する「市民社会調査」は、科学研究費補助金若手研究(A)(2007～2010年度、研究代表者：松本渉)「非営利セクターの展開に関する日米韓国際比較」(19683004)の一環として実施されたものである。また、本論文をまとめるにあたっては、科学研究費補助金基盤研究(B)(2015年度～継続中、研究代表者：前田忠彦)「パラデータを活用した訪問調査法の精度管理と不能バイアス補正に関する研究」(15H03424)を活用させて頂いた。ミシガン大学のミック・クーバー先生には、不明な点についての質問に対して丁寧なご回答を頂いただけでなく、入手困難なペーパーまでご提供頂いた。一連のご支援に謝意を表したい。

注

1) 前出の Kreuter (2013) の分類に照らせば、最終生成物に付随する「調査員による評価」はメタデータとも解釈できそうである。メタデータをパラデータに含める立場かどうかによってパラデータと呼ぶかどうか分かれるのではないか。

2) 前出の Kreuter (2013) の分類に照らせば、地点情報と似たところがあり、補助データとも解釈できそうである。ここも定義があいまいな箇所と思われる。

3) パラデータの網羅範囲は広がってきたが、実は

アイトラッキングデータは含まない。アイトラッキングは日本国内でもよく見かけるようになってきたが、そのデータは実験的状況でしか入手できないためパラデータには含まないのである (Olson and Parkhurst, 2013)。

4) Olson and Parkhurst (2013) の分類に基づく表現。杉野・俵・轟 (2015) の論文では、調査員観察と表現している。

5) self-interview という英語もあることにも注意が要る。

文献

有坂路子, 2010, 「面接調査の訪問状況記録の検証」『新情報』98: 23-37。

Bethlehem, J., Cobben, F. and Schouten B., 2011, *Handbook of Nonresponse in Household Surveys*, Hoboken, New Jersey: John Wiley & Sons.

Couper, M. P., 1998, "Measuring Survey Quality in a CASIC Environment", In Proceedings of the Section on Survey Research Methods of the American Statistical Association.

Couper, M. P. and Lyberg, L., 2005, "The Use of Paradata in Survey Research", In Proceedings of the 55th Session of the International Statistical Institute, Sydney, Australia, April.

Couper, M. P. and Kreuter, F., 2013, "Using Paradata to Explore Item Level Response Times in Surveys", *Journal of the Royal Statistical Society: Series A*, 176 (1): 271-286.

Groves, R. M., Benson, G., Mosher, W. D. and Chandra, A., 2005, "Plan and Operation of Cycle 6 of the National Survey of Family Growth", National Center for Health Statistics. *Vital and Health Statistics* 1 (42), available at: http://www.cdc.gov/nchs/data/series/sr_01/sr01_042.pdf.

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』10 (1): 81-98。

日野愛郎・田中愛治編, 2013, 『世論調査の新しい地平——CASI方式世論調査』勁草書房。

城川美佳, 2003, 「医療・保健領域の電話調査, Waksberg変法」『行動計量学』30 (1): 111-119。

Kreuter, F., 2013, "Improving Surveys with Paradata: Introduction", Kreuter, F. (ed.) *Improving Surveys with Paradata: Analytic Uses of Process Information*, Hoboken, New Jersey: John Wiley & Sons: 1-9.

Kreuter, F., Couper, M. P. and Lyberg, L., 2010, "The Use of Paradata to Monitor and Manage Survey Data Collection", In Proceedings of the Section on Survey Research Methods of the American Statistical Association.

Kreuter, F. and Olson, K., 2013, "Paradata for Non-response Error Investigation", Kreuter, F. (ed.) *Improving Surveys with Paradata: Analytic Uses of Process Information*, Hoboken, New Jersey: John Wiley & Sons: 13-42.

Matsumoto, W., 2007, "A Study on Technical and Cultural Differences in Survey Methodology between Japan and the United States", *International Journal of Japanese Sociology*, 16 (1): 23-34.

松本 渉, 2009, 「RDD電話調査におけるコールレコードの分析——市民社会調査 (CATI) データを用いた日米調査法比較」『2009年度統計関連学会連合大会講演報告集』, 133。

———, 2010, 「米国民社会調査 (CATI - RDD) の調査設計と発信結果記録」『よろん』105: 22-27。

———, 2011, 「パラデータにみられる不整合について」『2011年度統計関連学会連合大会講演報告集』, 93。

———, 2014, 「アメリカの量的社会調査」社会調査協会編『社会調査事典』丸善出版, 710-713。

松本 渉・前田忠彦, 2008, 『市民の政治参加と社会



- 貢献の国際比較—日本調査報告書』(統計数理研究所研究レポート97)。
- 松岡亮二・前田忠彦, 2015, 『日本人の国民性第13次全国調査』の欠票分析—個人・地点・調査員の特性と調査回収状況の関連』(統計数理) 63(2): 229-242。
- 中村 隆・前田忠彦・土屋隆裕・松本 渉, 2009, 『国民性の研究第12次全国調査—2008年全国調査』(統計数理研究所研究レポート99)。
- 中村 隆・前田忠彦・土屋隆裕, 2015, 『国民性の研究第13次全国調査—2013年全国調査』(統計数理研究所研究レポート116)。
- 大隅 昇, 2004, 「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」『新情報』91: 1-24。
- , 2014, 「ウェブ調査」一般社団法人社会調査協会編『社会調査事典』丸善, 106-113。
- Olson, K. and Parkhurst B., 2013, “Collecting Paradata for Measurement Error Evaluations”, Kreuter, F. (ed.) Improving Surveys with Paradata: Analytic Uses of Process Information, Hoboken, New Jersey: John Wiley & Sons: 43-72.
- O’Reilly, J., 2009, “Paradara and Blaise: A Review of Recent Applications and Research”, Paper presented at the International Blaise Users Conference (IBUC).
- 篠木幹子, 2010, 「社会調査の回収率の変化」『社会と調査』5: 5-15。
- 杉野 勇・俵 希實・轟 亮, 2015, 「モード比較研究の解くべき課題」『理論と方法』30(2): 253-272。
- 鈴木裕久, 2014, 「世論調査」社会調査協会編『社会調査事典』丸善, 334-337。
- 鈴木達三, 2003, 「電話調査法についての一考察」『行動計量学』30(1): 73-91。
- West, B. and Sinibaldi, J., 2013, “The Quality of Paradata: A Literature Review”, Kreuter, F. (ed.) Improving Surveys with Paradata: Analytic Uses of Process Information, Hoboken, New Jersey: John Wiley & Sons: 339-359.
- 保田時男, 2005, 「回収率の改善のための調査員の制御についての考察」尾嶋史章編『現代日本におけるジェンダーと社会階層に関する総合的研究』(科研費成果報告書): 319-327。
- , 2009, 「JGSSにおける調査員の訪問記録の分析」大阪商業大学 JGSS研究センター・東京大学社会科学研究所編『JGSSで見た日本人の意識と行動—日本版 General Social Surveys 研究論文集[8]』79-90。
- 保田時男・宍戸邦章・岩井紀子, 2008, 「大規模調査の回収率改善のための調査員の行動把握—JGSSにおける訪問記録の分析から」『理論と方法』23(2): 129-136。

2

パラデータ概念の誕生と普及

ミック・P・クーパー

ミシガン大学調査研究センター 研究教授

松本 渉 訳

1 はじめに

本論文では、パラデータ (paradata) の概念の歴史と発展についての短い展望を述べる。まずこの用語の当初の開発と多種多様な異なる種類のパラデータへの拡張について取り扱い、その上でパラデータの例とパラデータの用途を概観する。本論文は、1998年に私が初めてパラデータの用語を取り入れて以降、パラデータがどのように進化してきたかについての個人的な所感である。『社会と調査』の本特集がパラデータに焦点をあてたことについては、とりわけ喜ばしい。本論文は、パラデータ概念の一般的な使用方法を説明するだけでなく、日本の研究者におけるパラデータの利用をより一層促すことに役立つはずである。

2 パラデータ (Paradata) の起源

パラデータまたはプロセスデータの考えは、以前から存在していたが、この用語の起源は、1998年にまで遡る。この用語は特定の状況に応じて開発されたものであるが、その後様々なデータの種類の網羅するように拡張されてきた。具体的に言うと、私はCAI (computer-assisted interviewing)¹⁾のソフトウェアによって自動的に生成されるデータの種類の指す用語を開発した (Couper, 1998)。その当時 (そして今もな

お)、これらをトレースファイル、監査証跡 (audit trail)、ログファイルと呼ぶ人たちがいる一方で、キーストロークファイルと呼ぶソフトウェア製品もあった。それらに共通していたのは、CAIシステムを使用して調査を管理した際に、調査員によって入力されたすべてのキーストロークを保存することであった。すべて似たようなものを測定しているので、私は異なるシステムのユーザー間でのコミュニケーションを容易にするために共通の用語「パラデータ」を提案したのである。

これらのパラデータを保存することの本来の目的は、システムエラーを検出し、修復することにあったが、調査プロセスについて洞察する上での効用があることがすぐに判明した。最初に注目した点は、測定誤差の問題についてであり、聴き取りを行った際に調査員とコンピュータの間で生じる相互作用の間接的な記録を提供している点であった (Couper, Sadosky, and Hansen, 1994; Couper, 1995; Couper, Horm, and Schlegel, 1997参照)。Couper, Sadosky, and Hansen (1994, p. 849) は、次のように書いた。

キーストロークファイルは、調査のやりとりの最中に調査員とコンピュータの間の相互作用についての豊富なデータを私たちに提供する。私たちの研究では、そのような行動を理解すること、CAIにおける調査員のパフォーマンスの側面を評価するため系統的な方法でこ



これらのファイルの用い方を見つけること、そしてCAIの利用において調査員が引き起こす誤りや非効率を最小にするためCAIの道具や調査員の訓練において必要とされる改善点を特定することに注目している。

パラデータに対する関心が高まることに並行して、調査データの収集に対して技術の導入が進んだ。当初のパラデータ分析は、調査員がCAIのソフトウェアとどのようにやりとりしているかに注目していた。CATI(computer-assisted telephone interviewing)は1970年代以降存在していたのに対し、CAPI(computer-assisted personal interviewing)は1980年代後半に導入され、回答者との対話中のラップトップコンピュータの複雑なソフトウェアを管理する調査員の能力についての関心を高めることとなった。初期のパラデータ研究は(調査員が直面しうる困難を特定するため)調査員レベルの行動に目を向けるとともに、(困難を生じさせる質問文や機能を特定するため)項目レベルの行動にも目を向けていた。例えば、Couper, Hansen, and Sadosky (1997)は、調査員のコンピュータ経験は、非標準的行動を実行する様々なファンクションキーの使用に関連していたことを発見した。項目レベルでは、どの項目が調査員に最も助けを求めるということを生じさせたかを明らかにするためにパラデータ分析が使用された。これらの分析は、コンピュータ化された調査機器の有用性——調査員訓練を追加する必要性を指摘したり、あるいは問題のある質問や再設計の必要がある機能を特定したりしつつ——に焦点をあてていた。Couper, Hansen, and Sadosky (1997, p. 281)は「……キーストロークファイルは、システムを利用する際の調査員の能力の差異を明らかにするのと同じくらい(またはそれ以上)、コンピュータ化された調査機器やシステムの不備を明らかにするかもしれない」と述べている。

私は1998年の統計連合大会(JSM, Joint Sta-

tistical Meetings)のプレゼンテーションでパラデータの概念を初めて導入したが、(私にとって同僚でもあり、助言者でもあるロバート・グロブスのすすめがあつて)新しい用語を導入することは思いとどまった。そのため、公表された大会の抄録には「パラデータ」の用語は記載されなかった(Couper, 1998)。「パラデータ」の用語は、その2年後まで印刷物に記載されなかったのである。調査用具の有用性テストの論文(Couper, 2000, p.393-394)において、私は次のように書いた。

コンピュータ支援型のデータ収集法の利点の1つは、大量の自動的なデータが調査過程から生成され、その後でこれらのデータを調査過程を評価するために使用することができる点である。私は、(データを記述する)メタデータと区別するため、これらの情報源を「パラデータ」(調査過程を記述する補助データ)と名づける。パラデータの情報源は、回収率、ケースごとのコール数²⁾、平均的な聴取時間の長さなどのケース管理情報を含んでいる。有用性評価のために最も役立つパラデータは、キーストロークファイルやトレースファイル(監査証跡とも呼ばれる)である。

ほぼ同時期のことだが、ACASI(audio computer-assisted self-interviewing)という新しい技術の導入に際して、CAIパラデータの分析がその回答利用を検証するため利用された(Caspar and Couper, 1997参照)。調査員の観察もまたCASI(computer-assisted self-interviewing)の評価に使用されていた(Couper and Rowe, 1996参照)。ACASIとCASIの回答者に焦点をあてたことは、ウェブ調査の発展(後述する内容参照)に伴う、パラデータへの関心の高まりを予見させるものであった。

しかし、パラデータの概念は、1990年代後半と2000年代初頭に次の2つが並行して発展す

るに至って初めて本当に勢いを増し、広範な関心を得るようになった。1番目がウェブ調査の発展であった。2番目が大規模調査を管理および監視することのできるコールレコードデータ³⁾への新たな関心の高まりだった。パラデータ概念の定義や意味の問題に戻る前に、以下ではこれらの発展についてより詳しく述べる。

3 概念の拡大

前述のように、2つの発展がパラデータという用語の利用拡大をもたらしたが、同時にこれはパラデータを構成するものの定義の拡大ももたらした。本節ではこれらの発展のそれぞれについて述べる。

1番目の発展は、1990年代後半におけるウェブ調査の導入と急速な成長であった。コンピュータ支援型調査で獲得されうるキーストロークデータの類は、今ではウェブ調査を完了すると同時に、回答者から直接獲得される可能性もある。以前であれば、紙媒体の自記式調査で入手できなかった回答者行動について洞察することを可能にしている。

(私の知る限り)ウェブ調査にパラデータの利用語を用いたのはJeavons (2001)が最初であった。彼は「ウェブは、研究者に金鉱山を提供する。すべてのクリック、反応、ページ移動と回答の遅れが記録され、分析に使用されうる」(p. 177)と書いた。Heerwegh (2003;あるいはHeerwegh, 2008, 2011参照)は、サーバ側パラデータ (SSP, server-side paradata) とクライアント側パラデータ (CSP, client-side paradata) とを区別し、ウェブ調査でCSPを得るために使用できるJavaScriptコードを作ったことで知られる。自記式であるウェブ調査部分におけるCSPは、CAIにおけるキーストロークファイルに相当する。Heerweghの貢献は、ウェブ調査のパラデータの利用の急速な拡大につながった。他の者たちはHeerweghのJavaScriptコードを拡張し、汎用

スクリプトをウェブ調査で使用できるようにした (Kaczmirek, 2009参照; <http://kaczmirek.de/ucsp/ucsp.html>)。ウェブ調査のパラデータは今ではオンライン調査におけるデータ収集の不可欠な部分である (本特集の大隅昇の論文参照)。ウェブ調査で使用されるパラデータは、典型的には回答者レベルや質問項目レベルにおける、データの品質や測定誤差に関するものに集中している。

パラデータの普及につながった2番目の発展は、家族の成長に関する全米調査 (NSFG, National Survey of Family Growth) を実施する際にグローブスと仲間たちがパラデータという用語を使用したことだった。Groves et al. (2005)は、2002年のNSFGの第6次調査を管理するためにパラデータを広範囲に使用した。CATIにおいて架電記録データが長年使用されてきた一方 (例えば、Weeks et al., 1980; Greenberg and Stokes, 1990参照)、CAPIの台頭と (標本の住所情報が記載された、調査員が架電や訪問の詳細を記録する) 紙媒体の「表紙」からコンピュータによる標本管理システムへの移行がパラデータに対する新たな関心を生み出した。グローブスのこの用語に対する当初の抵抗を考えると、皮肉なことであるが、この概念を世間に広め、調査研究において広く受け入れられるようにしたのは、部分的とはいえグローブスの尽力による。

Groves and Heeringa (2006) のレスポンス・デザイン (responsive design) に関する論文は、コールレコード・パラデータ概念をさらに調査研究の主流に向かわせることに役立った。標本世帯に接触し、抽出対象者からの協力を得ることの難しさが増し、調査データ収集費用が上昇する中、研究者は既存のパラデータをより多く利用し、リアルタイムで調査を監視したり管理したりするための追加的な観測データを収集し始めた。このアプローチ——レスポンス・デザインやアダプティブ・デザイン (adaptive design) など様々に呼ばれる (Calinescu and Schouten,



2016; Couper and Wagner, 2011; Schouten, Calinescu, and Luiten, 2013; Wagner et al., 2012 参照) —は北米およびヨーロッパで今や広く使われている。他記式調査で用いられるパラデータは、通常は標本ケースレベルのカバレッジ、サンプリング、および調査不能誤差に関して集中している。Groves and Heeringa (2006) は、レスポンス・デザインやアダプティブ・デザインでは、データ収集時になされる決定を導くためには豊富なプロセスデータまたはパラデータが必要であると述べている。このようなパラデータは、費用や誤差の代理指標として機能する。

4 パラデータの定義と意味

パラデータの概念の進展を考えるにあたって、この用語の意味と定義の問題に戻ろう。ここでは手短かに「パラデータ (paradata)」の用語の起源を説明し、その意味が時間の流れとともにどのように発展してきたかを議論する。「パラデータ (paradata)」という単語を創出するために使用された接頭辞の「para」はラテン語に由来しており、「次に」あるいは「脇に」を意味している。それはまた (救急救命士・救急医療隊員 [paramedic] やパラリーガル [paralegal, 弁護士 の付随業務を行う法律事務職員] と同様) 「補助的な」や「二次的な」という意味もある。私は、「データ」(data, 調査で収集するまたは生成するとされているもの) や「メタデータ」(調査概要, コードブック, 値ラベルなどを含むデータに関するデータ) といった既存の用語に追加するために、「補助的な」や「追加的な」を意味する用語を選んだ。コンピュータ支援システムにおいて自動的に生成される副産物、すなわち調査プロセスに関する情報を提供する補助的なデータに焦点があたるような用語を作ったのである。

広範囲に使用されているにもかかわらず、「paradata」という用語はいまだに正式に定義されていない。聴き取りプロセスにおける自動

化されたデータとしてのより狭い概念は、他の様々な種類のデータや入手先を包含するように成長してきている。例えば、Olson (2013) は、「パラデータは拡大し、調査票自体から得られるデータ以外の調査員によって記録された情報も含むようになった」と述べている。それゆえ、自宅周辺に関する事柄、住居単位、世帯構成、あるいは回答者に関する事柄であろうと、調査員の観察したことは今やパラデータと見なされるのである。

Groves and Heeringa (2006, p.448) は、より一層包括的な定義を用いている。

(パラデータには) 標本個体に対する接触の試みの記録、標本ケースにたどり着くために要した移動距離と時間、聴き取り中に調査員が様々な作業に費やした時間といったデータ収集管理データを含めることができる。さらにパラデータは、標本個体における都市化水準、複数単位構造の存在、あるいは標本抽出枠に記録されている人口統計学的情報といった、データ収集資源を割りあてての有用な標本抽出枠に関するデータも含めることができる。パラデータには、標本ケースの調査への参加を得ることがどれほど難しいかの代理指標になるかもしれない新しい (de novo) 観察——調査のキー変数についての個体のもっともらしい特徴、あるいは回答を完了したケースの測定誤差が典型的なものであるかどうかといったこと——も含めることもある。

彼らの見方によれば、パラデータは、調査プロセスを管理し、評価することに用いる他の種類の補助的なデータ (例えば、センサスデータ、集計レベルの商用データ) も含んでいる。この意味で、パラデータは、出所よりも使用方法や目的によって定義されるものである。パラデータを大規模な一連のプロセスデータの部分集合とみなす者がいる一方で、両者を同等のものとみな

す者もいる (Couper and Lyberg, 2005; Kreuter, Couper, and Lyberg, 2010 参照)。

データ、メタデータ、パラデータ間の境界がぼやけている時もある。例えば、調査完了時間は、よくパラデータとみなされるが、もし結果的に調査完了時間の変数が分析データセットに含まれており、実質的な分析で使用される場合は、調査完了時間は調査データとみなされる可能性がある。同様に、調査のメタデータ (例えば、回収率、平均完了時間) は、生のパラデータから集計されるが、調査全体を記述するために使用される。Scheuren (2001) はミクロとマクロのパラデータの区別を行った。彼はマクロレベルのパラデータを調査データ収集に関する概要の記述統計量 (例えば、回収率、中断率、調査完了時間など) とみなす一方、個々のケース、コール、項目に関する詳細なデータをミクロレベルのパラデータとみなした。しかし、この区別が勢いを得ることはなく、マクロレベルのパラデータは、今では調査のメタデータとみなされるのが普通である。その意味では、パラデータについての私の記述で「回収率」を含めていたこと (Couper, 2000) は誤りだった。回収率はパラデータから得られるものであるが、通常はパラデータそのものとはみなされない。

したがって、パラデータはある部分はデータの出所によって定義される (すなわち調査過程の副産物として収集され、生成されるデータ) が、ある部分はその利用の観点からも定義される (すなわちデータ収集過程を私たちが理解するのに役に立つデータ)。

パラデータの範疇に何が含まれ、何が含まれないかについての論争は継続している。この言葉の本来の意味では、自動的に生成されるデータに焦点をあてていた。しかし、コールレコードデータなどのシステムデータ (例えば、架電/訪問の日時) と調査員によって入力された情報 (例えば、架電/訪問の結果、調査への協力を望まない調査対象者が述べた理由など) が混在し

ている。このことを考慮して、その他の調査員の観察内容も今ではパラデータとみなされるのが一般的である。調査を実施するために使用される道具は、常に進化しており、システムによって生成される可能性のあるデータの種類の拡大している。私自身の見解は、これらの区別は重要ではなく、パラデータの正確な定義は必要ではないというものである。

パラデータの最も広い定義はウィキペディアに記載されている。「調査のパラデータとは、調査データが収集されたプロセスについてのデータである」 (<https://en.wikipedia.org/wiki/Paradata> 参照) と。調査の方法と技術の変化が数年の間に狭めの定義を時代遅れにするので、私はこのように幅広いパラデータの見方に幾分か満足している。したがって、パラデータの定義が変化し進化してきたにもかかわらず、この概念には有用性が残っている。以下の各節で見ると、概念は拡大と進化が継続していくのである。

5 パラデータが円熟の域に達する

結果としてパラデータの問題は2つの大きくかけ離れた平行な立場をたどってきた。一方の立場は、他記式調査における調査員の観察と、しばしばコンピュータ制御の標本管理システムを介して生成されたコールレコードに注目する。これらのパラデータはしばしば代表性の誤差 (標本誤差、カバレッジ誤差、無回答誤差からなる) と大規模な調査プロジェクトの効率的な管理に焦点をあてる。もう一方の立場は、ウェブ調査におけるサーバ側パラデータとクライアント側パラデータ、そしてコンピュータ支援型調査におけるキーストロークレベルのパラデータに注目する。これらのパラデータは概して測定誤差に焦点をあてる。

2000年代の後半、2つの立場を合流させ、調査過程を評価・改善するためにパラデータを収集・分析することを正当化するのに役立つ調査パラ



データに関する複数のワークショップや短期コースが開催された。これらの行事は書物 (Kreuter, 2013) の出版につながったし, *Journal of the Royal Statistical Society*, シリーズ A での特集号「社会調査研究におけるパラデータの使用」につながった。これらはウィキペディアのページ (<https://en.wikipedia.org/wiki/Paradata> 参照) や『調査研究法百科事典』の項目 (Heerwegh, 2008 参照) とあわせて, 調査研究方法論の語彙の一部としてパラデータを確立させるのに役立った。パラデータに関するワークショップと短期コースは, 調査研究法に関するカンファレンス (例えば, 統計連合大会, 欧州調査研究会 [ESRA, European Survey Research Association] 会議と米国世論調査学会 [AAPOR, American Association for Public Opinion Research] 年次大会) によって提供されており, パラデータのウェビナー (ウェブ上のカンファレンス) は, これらの団体や他の専門家の団体によって提供されている。

パラデータが調査研究データ収集の主流の一部となったので, 獲得可能なデータの種類を増強するように, そして調査管理と評価・方法論研究の両方にとってデータがより有用になるように努めてきた。

パラデータという用語の利用もまた統計的調査の領域の枠を超えて拡大している。デジタル取引の増加と普及に伴い, 獲得され, 分析されるプロセスデータの量は, ますます増大している。監視のためにこのようなデータを使用する場合——例えば, 米国国家安全保障局 (NSA, National Security Agency) の電話のログファイル (誰が誰にかけたか, 通話時間など) の分析——あるいはビッグデータ分析に使われる場合, それらは通常はメタデータと呼ばれる。同様に, ウェブサーバーのログファイルやソーシャル・メディア・プラットフォームにおけるユーザー行動の記録は, サイトを改善すること, ユーザーの行動の理解することの両方を可能にするデータの発生源としてますます役に立つようになっている。し

かし, このようなデータを (しばしば調査と組み合わせながら) 研究目的のために使用する場合, それらのデータはパラデータと呼ばれることが多い。オンライン介入の分野 (例えば, Glasgow et al., 2007; Strecher et al., 2008; Resnicow et al., 2010 参照) から 1 つ例をあげよう。パラデータは, 参加者の関与とウェブサイトの利用の程度を評価するため, そしてどれほどの時間をそのサイトで過ごすか, どれほど多くのページを訪ねるか, どのオプション機能を利用するかを測るといった目的などに使用される。これは, 対照比較試験において薬の「投与量」や治療における暴露の程度を測定すること同様に役に立つものであり, 上記のオンライン介入の成功の理由を評価するために使用される。

2 番目の例は, 学習リソースの分析に関する分野から来ている。全米科学デジタル図書館 (NSDL, National Science Digital Library) は, 2010 年にデジタル学習オブジェクトとユーザーとの相互作用の呼び方にパラデータの用語を採用した (<https://en.wikipedia.org/wiki/Paradata> [learning_resource_analytics] 参照)。NSDL は, 統計的調査の領域における区別と, 同様にパラデータとメタデータを区別している (<https://wiki.ucar.edu/display/nsdl/docs/Paradata> 参照)。利用されている方法とツールは, ウェブ調査のパラデータにおける方法やツールと非常に似ている (例えば, Kilburn and Earley, 2015 参照)。しかしながら, ユーザー行動の追跡と研究ができるウェブサーバーのログファイルとクッキー (Cookie) の使用の増加にもかかわらず, パラデータという用語の利用の大部分は, まだ調査の領域に限定されている。

6 課題とチャンス

調査を評価および管理する上で, 様々な種類のパラデータの有用性が示されている一方で, まだ調査パラデータの広範な使用を制限する課

題が数多くある。

パラデータの利用拡大への障壁の1つは、パラデータの形式がいまだ使用されている特定のソフトウェアやシステムに特有の形式であるということである。つまり、パラデータの普遍的な基準がないということである。見方によれば、これは、監査証跡、トレースファイル、ログファイル等の類の初期の頃に似ている。当時は、すべてのシステムが独自のバージョンのパラデータを生成していたからである。ウェブ調査においては、依然として（調査のソフトウェアに管理されている）サーバー側パラデータにもあてはまるが、（標準のJavaScriptのクエリを使用している）クライアント側パラデータは、システム間および調査間で比較できるようなデータを次第に生み出すようになってきた。これについての2つの例として、Heerweghのクライアント側パラデータ（CSP）コード（<https://perswww.kuleuven.be/~u0034437/public/csp.htm>参照）とKaczmirekのユニバーサルクライアント側パラデータ（UCSP）コード（<http://kaczmirek.de/ucsp/ucsp.html>参照）がある。Diedenhofer and Musch (2014) のPageFocusコード等の特定のツールは、コード共有サービスであるGitHubにおいて入手可能となっている（<https://github.com/deboer/kePageFocus>参照）。このような動向は、ウェブ調査においてCSPがより広範に使用されるようになることを促している。他記式調査においては、パラデータの入手はまだ利用される調査ソフトウェアやケース管理システムに固有であるが、特定のソフトウェアのユーザー間での共有は広まっている（例えば、Blaiseの監査証跡とセッションシステム：<http://blaise.com/>参照）。

広範囲にパラデータが使用されることへのもう1つの障壁は、パラデータがそれほど構造的になっていないという性質にある。これはコールレコードのパラデータにもキーストロークのパラデータにも該当する。コールレコードデータの観測はしばしば階層的であり、他の事項（例：

接触状況、適格性）を条件として得られる結果（例：聴き取り実施、拒否）を伴っている。ケースごとの観測数も均一でなく、最初の訪問や架電で調査される対象者もいれば、最終的な調査結果が判明するまで多くの訪問や架電を試みられる調査対象者もいる。コールレコード分析の最近の進歩は、これらの課題にも取り組んでいる（例えば、Durrant, 2016; Durrant and D'Arrigo, 2014; Durrant, D'Arrigo, and Steele, 2013; Durrant, Maslovskaya, and Smith, 2015; Durrant and Steele, 2009; Steele and Durrant, 2011 参照）。

キーストロークレベルのパラデータあるいはウェブ調査におけるCSPでは、素データは多くの場合文字列形式であり、有用な情報を抽出するためには慎重な構文解析を必要とする。例えば、1つのウェブページをとっても、回答者間で様々な異なる動作——異なる数のマウスクリック、スクロール動作、同じページへの訪問の繰り返し、回答変更の有無——が生み出される可能性がある。望ましい動作測定の集約化（例えば、ページ訪問、ページ、調査全体、回答者）の水準を決定することは、前述のパラデータの活用においては重要な手順である。

これには分析に適した形式でデータを準備するための多くの事前処理が必要である。また、様々な分析ツールも必要である。これらの両面について進展がある。パラデータの獲得がより標準化されていくにつれ、そしてパラデータのより広範な利用がされるにつれ、関連のある情報を抽出することができるようツールが用意されてきている。次第に調査機関は、入手予定のパラデータを有効に生かし、調査をリアルタイム管理するためパラデータの処理を自動化するため、ケースマネジメントシステムのコンピュータ化を推進している（Hubbard et al., 2010; Thalji et al., 2013 参照）。ミシガン大学調査研究センターで開発されたミシガン調査管理システム（MSMS, Michigan Survey Management System）は、ミックスモードの調査データ収集



管理のためのパラデータの取得と処理を中心に構築されている。同様に、分析方法の開発はパラデータの分析の役に立っている（前述の通り）。これらの成果のうち数件は、Kreuter (2013) に記載されている。

関連する課題としては、パラデータの二次分析を奨励するために、パラデータを学外の研究者に対してより容易に利用できるようにする方法を見つけることがある。他人が解析するためのパラデータのファイルを準備することにはかなりの時間と手間がかかるが、このようなファイルを利用できるようにすることは、パラデータの利用を増やすと同時に、パラデータを抽出し、集約するための標準化を進展させることにつながり、そしてまたパラデータの分析を一層刺激することであろう。

パラデータの概要を二次分析のために提供している組織がある。調査員の観察が調査データの一部として提供されることもよくある。より詳細なコールレコードデータは、要求に応じて利用可能になる、あるいは（開示上の理由のため、除去されたか、カテゴリを併合された変数を用いた要約形式で）パブリックユースのデータセットとして公表される。一例として、国民健康聴き取り調査 (NHIS, National Health Interview Survey; http://www.cdc.gov/nchs/nhis/nhis_2014_data_release.htm参照) のパラデータファイルがある。この公開用データファイルは、3つの情報源からのキー変数が含まれていた。①接触履歴管理用具 (CHI: Contact History Instrument)、個々の調査員が接触を試みた結果について、調査員からデータを収集するもの（すなわちコールレコード・パラデータ）、②聴き取りを終えた後に、調査員がやり終えた質問についての感想を聞くこと（すなわち観測パラデータ）、③接触履歴管理用具のモジュールごとのデータと時間の情報（すなわちタイミングパラデータ）。これらはケースレベルで集計される。パラデータの利用可能性が向上することで、パブリックユースの

フォームであろうが、制限付データであろうが、研究者によるパラデータのより広範な使用を容易にするだろう。

パラデータの利用が増加するにつれて——とりわけ調査員の観察が関わってきたり、コールレコードに関する調査員による記録を必要としたりすることで——パラデータの質についての懸念が生じている。多くの観察は、調査員による主観的な判断を必要としており、調査員がこれらの判断をどれだけうまく行うかはばらつきがあるという証拠がある（例えば、Casas-Cordero et al., 2013; Eckman, Sinibaldi, and Montmann-Hertz, 2013; Sinibaldi, Durrant, and Kreuter, 2013; West, 2013）。最近の研究は、調査員に作成されたパラデータの質の評価と、調査員の訓練とフィードバックを向上させることにより、データの質を向上させる努力に焦点をあてている（例えば、West and Kreuter, 2013, 2015, 2016）。これらの評価についての取り組みは、コールレコードのパラデータか調査員の観察かにかかわらず調査員によって提供されるパラデータの質の向上につながっている。

パラデータの利用に関するもう1つの課題は、インフォームドコンセントについての問題である。コールレコードや調査員の観察には、標本対象者または回答者との間にはやりとりがない。（例えば）特にもしこのような観測が、パブリックユースのデータセットの一部分となる場合は、調査への不参加の理由であるとか、面接の様子や回答者宅の状況を記録することには、インフォームドコンセントについての疑問が生じるだろう。キーストロークレベルのデータとウェブ・パラデータの場合、このようなログファイルの取得はシステム障害から回復するための日常的な手順であるが、他の目的のためのデータ分析の利用が増えたことや、調査過程の追加的な詳細を取得するパラデータの役割が拡張することにより、研究目的のためにこれらのデータを利用することについての倫理的な懸念が高まっている

という議論がなされている。

Heerweghは、CSP開発の早い段階でこの問題に言及していた(<https://perswww.kuleuven.be/~u0034437/public/csp.htm>参照)。彼は「真に(方法論的に)必要とする研究のためにのみCSPを使用しなさい。回答者に対する『スパイ行為』のためだけにCSPを使用するべきではないし、ウェブ調査の回答者によって提供された確定的な回答を差し替えるためにCSPからの情報を決して使用してはならない」と書いていた。Singer and Couper (2011; Couper and Singer, 2013も参照)は、パラデータの取得について回答者に知らせることの影響を調査することを試みた。二人の研究は、パラデータとは何か、パラデータがどのように用いられるかを説明することが難しいことを明らかにした。

この問題についての最新の見解は、このようなパラデータ調査ですでに提供されることの他に回答者についての個人的な事柄を明らかにすることはなく、そしてパラデータが(回答者を評価するためではなく)方法論上の目的のために使われている限り、回答者に知らせることなくパラデータを使用することは容認可能であるということである。これは、回答者のウェブサイトの技能を改善するためにウェブサイトを介して回答者を追跡せずに、クッキーを用いることに似ている。しかし、タイミングパラデータはますます回答者を「スピード違反者(speeders)」や「満足者(satisficers)」(すなわち、調査の質問を読んだり答えたりする労力を割かない人々)に分類するために使用されているので、パラデータの収集を調査の回答者に伝えるかどうか(およびその方法)の問題は、再び浮上するかもしれない。

上記の調査パラデータを広範に導入する上での課題は、潜在的な好機と見ることができる。より多くの研究者がパラデータを使用し、彼らの研究を共有するにつれて、パラデータの需要が増加する。これは、パラデータの収集、処理お

よび解析のための基準とシステムを開発するための取り組みにつながる可能性がある。言い換えれば、これは研究者が取得済みパラデータを文書化し、他者とパラデータファイルを共有することを容易にし、さらにパラデータの使用を増加させることであろう。

最後の課題——と機会——は、パラデータ利用における2つの立場を一緒にすることである。これについての証拠は、パネル研究と縦断的研究において見受けられる。そこでは、測定誤差に焦点をあてていたパラデータ(例えば、項目欠測データ率、回答時間)が、後に続く質問での無回答や脱落を予測するために用いられていた。同様に、消極的な回答者を心変わりさせることによって回収率をあげることで、それらの回答者が提供するデータの品質との間にあるトレードオフを調査している研究者もいる。重ねて言うが、単一の誤差発生源に注目するのではなく、大局的な総調査誤差の展望をもって、より多くの研究をこれらの方針に沿って実施することは2つの立場を一緒にすることに役立つであろう。

7 パラデータの価値を高める

レスポンス・デザインやアダプティブ・デザインが使用されるにつれて、パラデータの需要は成長しつづけるであろう。同様に、コンピュータはますます調査データ収集のすべての側面において使用されるにつれて、生成されるパラデータは拡大を続け、パラデータの成長をさらに加速させる。パラデータは、リアルタイムの複雑なマルチモード調査を管理するために必要である。パラデータは、大規模な調査データを収集する過程においてますます不可欠な部分になるので、私たちは、誤差発生源を評価するだけでなく無回答調整に必要な評価を行うため、パラデータの質を評価し、収集されて利用されるパラデータの量と種類を広げることにはたゆまず奮励努力することだろう。測定誤差に対するパラデータ



の当初の注目は、カバレッジ誤差や無回答誤差、標本誤差と調査の費用と労力を含むものへと拡大している。

パラデータを構成するものの進化とともに、高まりつつあるパラデータの重要性のもう1つの例は、ウェブ調査の世界から来ている。ウェブの利用者はますますウェブ調査に回答するのにスマートフォンを使用するようになっていくので、モバイル機器上で調査に答えようとする際、このようなユーザーを識別し、モバイルユーザーが遭遇する問題を特定するために、パラデータの需要が成長している。これは、回答者によって使用されているデバイス、ブラウザ、およびオペレーティングシステムを識別するユーザーエージェント文字列の解析に端を発している (Callegaro, 2010参照)。しかし、ウェブ調査のパラデータは、スマートフォン上のスクロール、ピンチ・ズーム、向きの変更、その他のユーザーの行動を含めるに至るまで、さらに広がった (例えば、Couper and Peterson, 2016参照)。同様に、パラデータは、特にスマートフォン上での回答者のマルチタスク行動を検出することに使われている (Sendelbah et al., 2016; Diedenhofen and Musch, 2014)。モバイルデバイス向けにウェブ調査を最適化する必要が増加するにつれて、パラデータは、そのような機器を動かすためにリアルタイムで使用されている。回答者に対して利用可能なデバイスと技術が広がり続けるにつれて、ユーザーの経験値を向上させることと、様々なプラットフォームにおいて回答者が複雑な調査用具をどれほど効果的に使用できるかを理解することの両方の目的で、パラデータの利用が増加することであろう。

最近のパラデータ利用の増加にもかかわらず、パラデータは補完的な役割を提供するものであることを記憶しておくことは重要である。パラデータは、標本抽出枠に関わるデータや管理データまたはその他の補助データなどを含め、他の情報源に代わるものではない。しかし、パラ

データは調査の設計と実査を評価し、改善するための情報の追加的情報源を提供する。同様に、パラデータは、評価や品質管理の他の方法に代わるものではない。例えば、パラデータは、特別な課題に直面している調査員を特定することや特段の注意が必要な地域を識別することに役立つ。パラデータは、より費用のかかる方法を用いて追加の監視の必要があるかもしれない調査員を識別できる。同様に、調査票の設計においても、パラデータは潜在的な問題がある箇所を指摘する上で有用である。これらの問題には、認知的インタビュー、アイトラッキング研究、実験などのより集中的な方法を使用しながら対処することが可能である。したがって、パラデータは調査研究者に対して一連の追加的なツールを提供している。パラデータは、最も重要な課題に対して資源と注意を向けることができるように、大規模な調査が直面している問題を定量化するのに有用である。

8 要約と結論

約20年の間に、パラデータ概念は、CAIの調査過程の特定の部分を記述するという狭義の用語から——本特集の論文によって示されるように、CAPI、CATI、ウェブ、紙の調査票を含む——データ収集の多くのモードに適用できる幅広い用語に成長してきた。コンピュータ管理の調査用具の測定誤差と利用可能性に焦点をあてていたものから調査過程の他の多くの側面を網羅するものにまで成長してきた。パラデータの概念は、いまや成熟したものとなり、調査研究用語の一部として受け入れられている。パラデータの収集、処理および利用は成長を続けている。調査がますます複雑になり、プロセスのデジタルデータが利用可能になるにつれて、パラデータは調査の設計、管理および評価においてますます中心的な役割を果たすことであろう。こうした目的のためにパラデータの有用性を最大限に

発揮するためにはなお多くの課題が残っているが、このパラデータの領域における研究・開発は調査研究の分野を前進させているのである。

【訳者謝辞】この翻訳の草稿に対して大隅昇先生から有益なコメントを頂きました。謝意を表します。

訳者注

- 1) CAI, CATI, CASIといった用語については定訳がないため、あえて訳出せず、略語を用いた。
- 2) コール(call) には、電話調査における架電と面接調査などにおける訪問の両方の意味がある。

- 3) コールレコード(訪問記録や架電記録)。面接調査と電話調査の両方を想定して使われる場合はコールレコードとし、文脈から訪問記録か架電記録か明確にできる場合はいずれかに訳出した。

文献

- Calinescu, M., and Schouten, B. (2016), "Adaptive Survey Designs for Nonresponse and Measurement Error in Multi-Purpose Surveys." *Survey Research Methods*, 10 (1): 35-47.
- Callegaro, M. (2010), "Do You Know Which Device Your Respondent Has Used to Take Your Online Survey? Using Paradata to Collect Information on Device Type." *Survey Practice*, 3 (6).
- Casas-Cordero, C., Kreuter, F., Wang, Y., and Babey, S. (2013), "Assessing the Measurement Error Properties of Interviewer Observations of Neighbourhood Characteristics." *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176 (1): 227-249.
- Caspar, R. A., and Couper, M. P. (1997), "Using Keystroke Files to Assess Respondent Difficulties with an ACASI Instrument." *Proceedings of the Section on Survey Research Methods*, American Statistical Association. Alexandria: ASA, pp. 239-244.
- Couper, M. P. (1995), "New Tools for Evaluating Interviewer Performance in CAI." Paper presented at the International Field Directors and Field Technologies Conference, Fort Lauderdale, May.
- Couper, M. P. (1998), "Measuring Survey Quality in a CASIC Environment." Invited paper presented at the Joint Statistical Meetings of the American Statistical Association, Dallas, August.
- Couper, M. P. (2000), "Usability Evaluation of Computer-Assisted Survey Instruments." *Social Science Computer Review*, 18 (4): 384-396.
- Couper, M. P., and Hansen, S. E. (2001), "Computer-Assisted Interviewing." In J. F. Gubrium and J. Holstein (eds.), *Handbook of Interviewing*. Thousand Oaks, CA: Sage, pp. 557-575.
- Couper, M. P., Hansen, S. E., and Sadosky, S. A. (1997), "Evaluating Interviewer Performance in a CAPI Survey." In L. Lyberg et al. (eds.), *Survey Measurement and Process Quality*. New York: John Wiley, 267-285.
- Couper, M. P., Horm, J., and Schlegel, J. (1997), "Using Trace Files to Evaluate the National Health Interview Survey CAPI Instrument." *Proceedings of the Section on Survey Research Methods*, American Statistical Association. Alexandria: ASA, pp. 825-829.
- Couper, M. P., and Lyberg, L. E. (2005), "The Use of Paradata in Survey Research." In *Proceedings of the 55th Session of the International Statistical Institute*, Sydney, Australia, April [CD].
- Couper, M. P., and Peterson, G. (2016), "Why Do Web Surveys Take Longer on Smartphones?" *Social Science Computer Review*, available online, DOI: 10.1177/0894439316629932.
- Couper, M. P., and Rowe, B. (1996), "Evaluation of a Computer-Assisted Self-Interview Component in a Computer-Assisted Personal Interview Survey." *Public Opinion Quarterly*, 60: 89-105.
- Couper, M. P., Sadosky, S., and Hansen, S. E. (1994), "Measuring Interviewer Behavior Using CAPI." *Proceedings of the Section on Survey Research Methods*, American Statistical Association.



- ation. Alexandria: ASA, pp. 845-850.
- Couper, M. P., and Singer, E. (2013), "Informed Consent for Web Paradata Use." *Survey Research Methods*, 7 (1): 57-67.
- Couper, M. P., and Wagner, J. (2011), "Using Paradata and Responsive Design to Manage Survey Nonresponse." *Proceedings of the 58th Congress of the International Statistical Institute*, Dublin, Ireland, August.
- Diedenhofen, B., and Musch, J. (2014), "PageFocus: A New JavaScript to Detect Cheating in Online Tests." Paper presented at the General Online Research Conference, Cologne, March.
- Durrant, G. B. (2016), "Using Quantitative Paradata to Improve the Data Quality in Social Surveys." In R. Edwards, A. Phoenix, H. O'Connor, and J. Goodwin (eds.), *Working With Paradata, Marginalia and Fieldnotes: The Centrality of By-Products of Social Research*. London: Edward Elgar.
- Durrant, G. B., and D'Arrigo, J. (2014), "Doorstep Interactions and Interviewer Effects on the Process Leading to Cooperation or Refusal." *Sociological Methods and Research*, 43(3): 490-518.
- Durrant, G. B., D'Arrigo, J., and Steele, F. (2013), "Analysing Interviewer Call Record Data by Using a Multilevel Discrete-Time Event History Modelling Approach." *Journal of the Royal Statistical Society, Series A*, 176 (1): 251-269.
- Durrant, G. B., Maslovskaya, O., and Smith, P. W. (2015), "Modelling Final Outcome and Length of Call Sequence to Improve Efficiency in Interviewer Call Scheduling." *Journal of Survey Statistics and Methodology*, 3 (3): 397-424.
- Durrant, G. B., and Steele, F., (2009), "Multilevel Modelling of Refusal and Non-Contact in Household Surveys: Evidence from Six UK Government Surveys." *Journal of the Royal Statistical Society, Series A*, 172 (2): 361-381.
- Eckman, S., Sinibaldi, J., and Montmann-Hertz, A. (2013), "Can Interviewers Effectively Rate the Likelihood of Cases to Cooperate?" *Public Opinion Quarterly*, 77 (2): 561-573.
- Glasgow, R. E., Nelson, C. C., Kearney, K. A., Reid, R., Ritzwoller, D. P., Strecher, V. J., Couper, M. P., Green, B., and Wildenhaus, K. (2007), "Reach, Engagement, and Retention in an Internet-Based Weight Loss Program in Three Healthcare Systems." *Journal of Medical Internet Research*, 9 (2): e11.
- Greenberg, B. S., and Stokes, S. L. (1990), "Developing an Optimal Call Scheduling Strategy for a Telephone Survey." *Journal of Official Statistics*, 6 (4): 421-435.
- Groves, R. M., Benson, G., Mosher, W. D., et al. (2005), "Plan and Operation of Cycle 6 of the National Survey of Family Growth." National Center for Health Statistics. *Vital and Health Statistics* 1 (42), available at: http://www.cdc.gov/nchs/data/series/sr_01/sr01_042.pdf.
- Groves, R., and Heeringa, S. (2006), "Responsive Design for Household Surveys: Tools for Actively Controlling Survey Errors and Costs." *Journal of the Royal Statistical Society, Series A*, 169 (3), 439-457.
- Heerwegh, D. (2003), "Explaining Response Latencies and Changing Answers Using Client-Side Paradata from a Web Survey." *Social Science Computer Review*, 21 (3): 360-373.
- Heerwegh, D. (2008), "Paradata." In P. J. Lavrakas (ed.), *Encyclopedia of Survey Research Methods, Volume 2*. Los Angeles: Sage, pp. 574-575.
- Heerwegh, D. (2011), "Internet Survey Paradata." In M. Das, P. Ester, and L. Kaczmirek (eds.), *Social Research and the Internet*. New York: Taylor and Francis, pp. 325-348.
- Hubbard, F., Couper, M. P., Wagner, J., Gu, H., Benson, G., and Schulz, P. (2010), "Creating Effective Paradata Dashboards to Help Reduce Survey Costs and Errors." Paper presented at FedCASIC, Washington, D. C., March.
- Jeavons, A. (2001), "Paradata: Concepts and Applications." *Proceedings of the ESOMAR Net Effects 4 Internet Conference*, Barcelona, Spain, February. Amsterdam: ESOMAR, pp. 177-187.
- Kaczmirek, L. (2009), *Human-Survey Interaction: Usability and Nonresponse in Online Surveys*. Cologne: Herbert von Halem Verlag.
- Kilburn, D., and Earley, J. (2015), "Disqus Website-based Commenting as an E-research Method: Engaging Doctoral and Early-career Academic Learners in Educational Research." *International Journal of Research and Method in Education*, 38 (3): 288-303.
- Krueger, B. S., and West, B. T. (2014), "Assessing

- the Potential of Paradata and Other Auxiliary Information for Nonresponse Adjustments." *Public Opinion Quarterly*, 78 (4): 795-831.
- Kreuter, F. (ed.) (2013), *Improving Surveys with Paradata: Analytic Uses of Process Information*. New York: Wiley.
- Kreuter, F., Couper, M. P., and Lyberg, L. E. (2010), "The Use of Paradata to Monitor and Manage Survey Data Collection." Invited paper presented at the Joint Statistical Meetings, Vancouver, August.
- Olson, K. (2013), "Paradata for Nonresponse Adjustment." *Annals of the American Academy of Political and Social Science*, 645 (1): 142-170.
- Resnicow, K., Strecher, V. J., Couper, M. P., Chua, H. F., Little, R. J. A., Nair, V., Polk, T., and Atienza, A. A. (2010), "Methodological and Design Issues in Patient-Centered eHealth Research." *American Journal of Preventive Medicine*, 38 (1): 98-102.
- Scheuren, F. (2001), "Macro and Micro Paradata for Survey Assessment." Washington, D.C.: Urban Institute, unpublished paper.
- Schouten, B., Calinescu, M., and Luiten, A. (2013), "Optimizing Quality of Response Through Adaptive Survey Designs." *Survey Methodology*, 39 (1): 29-58.
- Sendelbah, A., Vehovar, V., Slavec, A., and Petrovčič, A. (2016), "Investigating Respondent Multitasking in Web Surveys Using Paradata." *Computers in Human Behavior*, 55 (Part B): 777-787.
- Singer, E., and Couper, M. P. (2011), "Ethical Considerations in Web Surveys." In M. Das, P. Ester, and L. Kaczmirek (eds.), *Social Research and the Internet*. New York: Taylor and Francis, pp. 133-162.
- Sinibaldi, J., Durrant, G. B., and Kreuter, F. (2013), "Evaluating the Measurement Error of Interviewer Observed Paradata." *Public Opinion Quarterly*, 77 (1): 173-193.
- Steele, F., and Durrant, G. B. (2011). "Alternative Approaches to Multilevel Modelling of Survey Non-Contact and Refusal." *International Statistical Review*, 79 (1): 70-91.
- Strecher, V. J., McClure, J., Alexander, G., Chakraborty, B., Nair, V., Konkel, J., Greene, S.M., Couper, M. P., Carlier, C., Wiese, C., Little, R. J. A., Pomerleau, C., and Pomerleau, O. (2008), "The Role of Engagement in a Tailored Web-Based Smoking Cessation Program: Randomized Controlled Trial." *Journal of Medical Internet Research*, 10 (5): e36.
- Thalji, L., Hill, C. A., Mitchell, S., Suresh, R., Speizer, H., and Pratt, D. (2013). "The General Survey System Initiative at RTI International: An Integrated System for the Collection and Management of Survey Data." *Journal of Official Statistics*, 29 (1): 29-48.
- Wagner, J., West, B. T., Kirgis, N., Lepkowski, J. M., Axinn, W. G., & Ndiaye, S. K. (2012), "Use of Paradata in a Responsive Design Framework to Manage a Field Data Collection." *Journal of Official Statistics*, 28 (4): 477-499.
- Weeks, M. F., Jones, B. L., Folsom, R. E. and Benrud, C. H. (1980), "Optimal Times to Contact Sample Households." *Public Opinion Quarterly*, 44 (1): 101-114.
- West, B. T. (2013), "An Examination of the Quality and Utility of Interviewer Observations in the National Survey of Family Growth." *Journal of the Royal Statistical Society, Series A*, 176(1): 211-225.
- West, B. T., and Kreuter, F. (2013), "Factors Affecting the Accuracy of Interviewer Observations: Evidence from the National Survey of Family Growth." *Public Opinion Quarterly*, 77 (2): 522-548.
- West, B. T. and Kreuter, F. (2015), "A Practical Technique for Improving the Accuracy of Interviewer Observations of Respondent Characteristics." *Field Methods*, 27(2): 144-162.
- West, B. T. and Kreuter, F. (2016), "Strategies for Increasing the Accuracy of Interviewer Observations of Respondent Features: Evidence from the U.S. National Survey of Family Growth." Draft paper under review.



特集論文

3

訪問調査における 調査員訪問記録の活用について 事例紹介として

前田忠彦

統計数理研究所データ科学研究系 准教授

1 はじめに

面接法などの調査員が介在する調査では、調査員が対象者から取得する回答だけではなく、完了者と調査不能者の双方について記録することが可能な変数、例えば対象世帯の住居に関する情報や、調査員自身の活動記録なども、分析対象とすることが可能である。面接法による社会調査の回収率が長期的に低落傾向にある状況下で、回収率向上や調査不能バイアスの検討に資するべく、次の二つの側面を考えることの意義は大きいだろう。

1. 調査員の行動分析: 調査員の行動は、回収率に影響する要因のうち調査主体が積極的に介入することが可能な要素であり、その部分のさらなる研究が求められる。
2. 協力者の回答との関連の検討: 回収に要した訪問回数は、項目内容によっては対象者の回答内容とも関連する。

以下この二つの側面に関する簡単な検討事例を紹介する。本稿では、面接法や留置法など調査員による対象者の個別訪問を伴う調査において取得される、「パラデータ」と呼ばれるデータのうち、特に訪問記録から得られる情報の活用について、統計数理研究所が行った調査での経験を元に、多少の情報提供を行うこととした。訪問調査というのはあまり使われない語かも知れないが、上記の二方法を念頭においてい

る。パラデータの問題については本特集の他論文に詳しいので、本稿では「調査の実施に伴い付随的に得られる各種情報」程度に緩やかに考えておくことにしたい。こうした訪問調査においても、コンピュータ支援型調査が次第に行われるようになり、例えば私が関わったものでも、「2015年階層と社会意識全国調査(第1回SSP調査)」では、CAPI法(Computer Assisted Personal Interviewing: コンピュータ支援による個人面接法)が採用されている(SSPプロジェクト事務局, 2016)。その概要を記した同調査報告書中の前田(2016)でも明確に言及していないが、同調査ではCAPIであることを利用して、回答者が個々の設問に回答した日時のデータも取得している。こうした点ではWeb調査などと同様に、調査実施中にリアルタイムで得られるプロセスデータ(としてのパラデータ)の解析が可能だが、その紹介は別の機会に譲ることとしたい。

2 調査不能に関する研究を背景として

こうした調査付随情報の解析が要請される背景について、私見ではあるが調査主体としての経験に基づいて次のようなことを考えている。

統計数理研究所による日本人の国民性第13次全国調査(中村・土屋・前田, 2015; 以下KS13と略記)では回収率が50%となっており、1953年以来的調査で過去最低を記録した。統計数理研究所が実施した他の関連調査でも回収率の低下傾

向に歯止めがかかっていない。調査不能（の増加）の主な理由は一時不在と拒否（の増加）、特に後者であり、KS13では不能総数に対して一時不在 20%、拒否 59%である。この水準の回収率では、調査不能バイアスが調査精度に対する深刻な脅威となりうる。事後的な調査不能バイアスの調整 (bias adjustment) のための方法を、調整自体の当否を含めて慎重に検討する必要がある。例えば Fushiki and Maeda (2014)、伏木・前田 (2015) はそうした試みの一端であり、中にはバイアス調整の効果が大きい項目も例示されている。しかし、バイアス調整はあくまで補助的な手段であり、やはり調査不能自体を減らし回収率の点で質の良い調査を実施することが第一義的な重要性をもつものと考えている。つまり、調査不能に寄与している様々な要因の精査が求められる。調査不能の寄与要因は調査デザイン全体に関わると言えるほど多様であるが、ここでは a) 対象者の個人属性や世帯的属性、b) 調査地点の要因、c) 調査員側の要因、d) 調査実施側の実施管理面の要因、などを想定しておこう。

本稿での主題は d) 要因の検討である。つまり、調査の有効／不能の説明要因たりうる側面の一つとして、調査員の活動状況について若干の分析結果を紹介するのだが、その前にそれ以外の要因についても多少のコメントを残しておこう。

a) については従来からも多数の検討が行われてきたはずである。曰く、女性よりも男性のほうが、また都市度の高い地域のほうが、そして若齢層のほうが、住居形態としては戸建て住宅よりも集合住宅居住者のほうが、調査不能率が高い。一例として前田 (2005) を参照していただきたい。この論文を含めて、私の分析経験によれば、都市度とは別に地域性として関東圏や近畿圏は不能率が高い。また集合住宅の中でもオートロック付きのマンションなど、何らかの入場制限を伴う住居は更に不能の可能性が高まる。こうした住居情報は、調査対象者宅を訪問してはじめて得られる情報であるから、調査員観察

に基づくパラデータの例である。本稿の主題とはしないが、近年の統計数理研究所の調査では対象者宅の様子の観察として、表札の有無や、住居の様子（新しさ、大きさなど）を記録するようにしている。これらの住居の様子も経験的には（戸建てか集合住宅かという住居形態の情報に加えて）調査不能の説明要因になりうるようである。実際 Fushiki and Maeda (2014) ではバイアス調整の文脈の中ではあるが、その一部を調査不能の説明モデルの中で利用している。

b) や c) の要因について近年統計数理研究所が実施した調査を題材とした検討事例として、b) については Matsuoka and Maeda (2015)、c) については松岡・前田 (2015) がある。これらの研究は調査地点の要因として、地域を都市規模や地方ブロックなどの大まかな要因で区分した変数だけではなく、調査地点毎に得られる地点属性を用いて多水準分析を行っている。結果としてはそうした地点変数や調査員変数の効果が一部認められ、そのことから調査管理についての提言などが行われている。二つの論文の中では地点についての公的統計データを利用しているが、例えば地点の様子についての調査員観察によりデータを取得するならば、そうしたもののパラデータに含めて考えても良いのだろう。

以下 d) 調査実施側の実施管理面の要因としての、調査員の訪問状態についての検討事例を紹介する。繰り返しになるが、実施プロセスにおいて少しでも有効回収を増やすための調査管理法の研究が求められる。

3 調査員の訪問記録に関する分析

本節では、主に KS13 の調査データに基づく訪問記録の検討結果を紹介するが、参考までに、方法は多少異なるが調査年の近い別の調査での検討結果も紹介しよう。



分析方法など

(1) 調査データ

KS13では、計画標本400地点6,400名に対し、計19,740回の訪問が記録された。回収数は3,170、回収率は49.5%である。調査期間は2013年10月末～12月上旬で足かけ8週、7回の週末を含む比較的長い期間をかけての調査となっている。

2013年1月～2月に実施された統計数理研究所による「国民性に関する意識動向(2012年度)調査」についても一部の結果を参照する。この調査は面接法(計画標本サイズ3,000、回収数1,528、回収率50.9%)、留置法(計画標本サイズ3,500、回収数2,114、回収率60.4%)という2種類の方法で実施された。以下DK2012(面接)、DK2012(留置)と略記する。

上記いずれの調査も実査のプロセスは、専門調査会社に委託実施したものであり、委託先はKS13とDK2012で異なる。また実査はコンピュータ支援型ではなく、従来法として紙の調査票を利用したものになっている。

(2) 訪問記録

各調査での訪問記録の内容は概ね共通である。例えばKS13では、調査票の表紙ページが訪問記録のための記録票になっており、8回分の訪問状況を記録できるようになっている。調査員は各回の訪問日時、接触した対象(本人、家族、それ以外、接触なし)を記録する。調査票の最後に調査完了日時を記録することになっており、完了票についてはこの訪問記録の最後の訪問と調査票末尾の完了記録が一致するのが本来の形であるが、記録に多少の齟齬が生じていることも多い。この意味では、調査員が手動で紙の調査票に記録する形での訪問記録は精度が充分とは言えないかも知れない。こうした点はコンピュータ支援型調査での改善が期待されるだろう。なお、8回目の記録は、絶対数も少ないが、少し性格が異なる可能性がある。というのも、8回より多く訪問した場合、最後の活動の記録を

付けるよう調査員に指示したためである。

(3) 分析方法

ここでは保田・宍戸・岩井(2008)と保田(2009)に示された方法による調査員活動記録の分析を、主にKS13のデータを用いて例示する。その方法とは調査員の対象者への個々の訪問回の結果を、A:接触成功→協力(完了)、B:接触成功→拒否(非回収)、C:接触失敗→接触断念(非回収)、D:接触失敗→訪問継続、と大分類し、(1)接触成功率 $= (A+B) / (A+B+C+D)$ 、(2)協力獲得率 $= A / (A+B)$ 、(3)断念率 $= C / (C+D)$ 、(4)訪問継続率 $= D / (C+D)$ 、などの指標を検討しようとするものである。保田ほか(2008)の提案を少し変更し、(3)の断念率を(3)a不在に由来するものと、(3)bその他の理由に分けてより細かく整理するものとする。

結果

(1) 回収結果別の訪問回数

訪問記録の分析例を示す前に、そもそものような回収経過を示したのかについて見てみることにしよう。図1にDK2012(面接)の回収経過を示したが、回収と拒否については週末に伸びを示し、ともに上記説明のように対象者の接触を伴う結果であることを反映して、累積曲線が似た形状を示している。KS13(面接)の回収経過も同じようなものであるが、DK2012(留置)の場合は様子が異なる。

訪問回数の記録について三つの調査の基礎情報を示しておく。表1は一対象者当たり、調査結果別(完了、拒否、短期不在、その他の不能に大別)の訪問回数の記述統計である。例えばKS13では完了までに平均2.54回の訪問が必要としたのに対して、拒否はそれより0.5回ほど多い。資料は割愛するが、面接調査の回収と拒否(ともに対象者への接触が必要)での累積訪問数は、KS13でもDK2012(面接)でも週末の伸びが明確に観察されるのに対し、DK2012(留置)では明確には観察されず、調査員の活動方

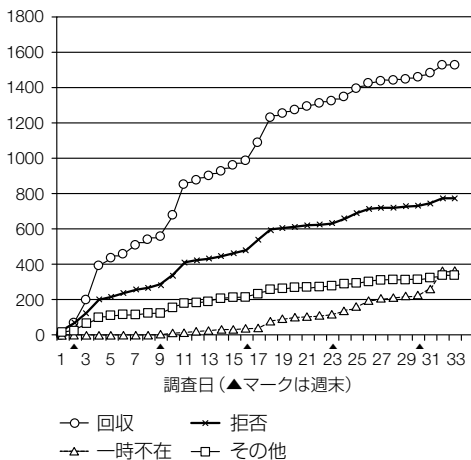


図1 動向2012調査における経過日数(調査日)と回収などの経緯

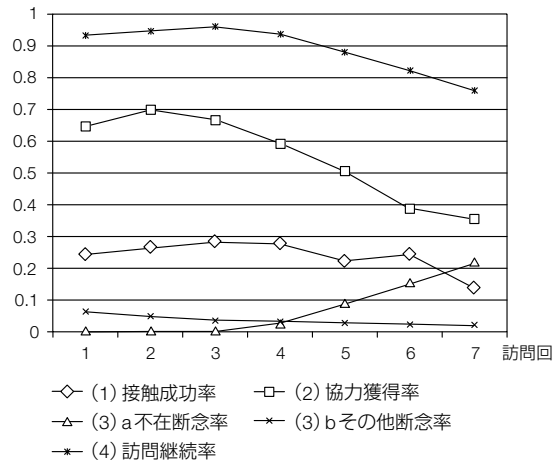


図2 訪問回別の4指標の整理
(日本人の国民性第13次全国調査)

法が調査法によって異なることを示している。

DK2012で見れば留置法では調査員の訪問回数は少し異なり、回収と一時不在で訪問回数の記録が面接法より多い。また表には示されないが面接法では休日に回収されたケースが全体の67%であるのに対し、留置法の場合では47%である。留置法では週末に対象者に接触し調査票を配付、平日に回収するというパターンも多く含まれることが推察される。

訪問記録の分析結果

以下の訪問記録の検討はKS13に基づく。

第1項の「分析方法」で定義した(1)～(4)の指標を、訪問回別に整理したものが図2である。接触成功率は、4回目までは少しずつ伸びるがそれ以降は逓減気味であり、訪問時期としては調査開始後の第4週目平日まではほぼ同じ水準を保つが、その後漸減する。協力獲得率もその辺りから漸減しており、これらを見ると、4週目までの活動にくらべてその後の活動で効率が落ちる。より協力の獲得が困難な対象者が残った結果であると推察される。

住居形態別に接触成功率と(接触に成功した後の)協力獲得率を見たものが図3と図4である。

表1 三つの調査での調査結果別の訪問回数に関する基礎統計

結果分類	KS13		DK2012(面接)		DK2012(留置)	
	平均	SD	平均	SD	平均	SD
1. 回収	2.54	1.61	2.27	1.46	2.92	1.25
2. 拒否	3.07	1.99	2.35	1.56	2.23	1.56
3. 一時不在	6.70	1.33	5.59	1.53	5.84	1.48
4. その他	2.29	1.68	2.29	1.64	2.22	1.64
全体	3.08	2.11	2.70	1.86	2.95	1.67

接触成功率は、一戸建>>集合住宅>オートロック有の集合住宅(>>は大差あり,>は僅差)となるが、接触に成功した後の協力獲得率は住居形態別には大差がなさそうである。一般に接触が困難とされるオートロック住宅の場合の接触成功率は、訪問回数が浅いうちは通常の集合住宅とも差があるが、回数を重ねると差が縮まる傾向にある。協力獲得率も初回訪問時では他にくらべて劣るが、2回目以降は遜色なくなっている。

4 協力者の回答と訪問回数の連関の検討

検討の目的と留意点

ところで面接法や留置法などの調査員が対象

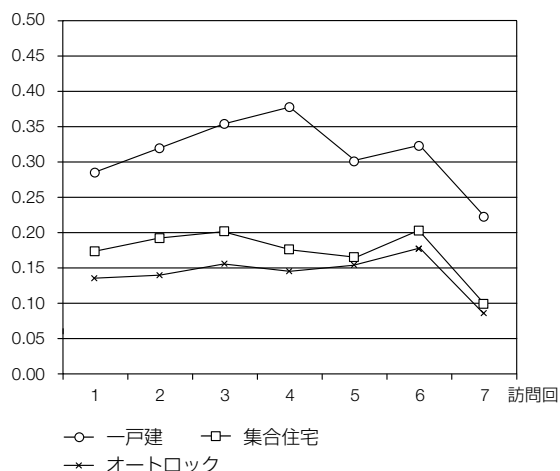


図3 住居形態別の「(1)接触成功率」の訪問回別推移

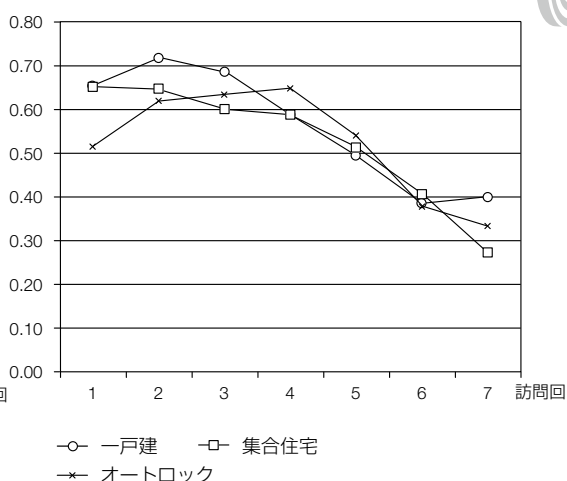


図4 住居形態別の「(2)協力獲得率」の訪問回別推移

表2 三つの調査における訪問回数の平均

結果分類	KS13		DK2012(面接)		DK2012(留置)	
	平均	SD	平均	SD	平均	SD
1. 完了	2.54	1.61	2.27	1.46	2.92	1.25
2. 拒否	3.07	1.99	2.35	1.56	2.23	1.56
3. 一時不在	6.70	1.33	5.59	1.53	5.84	1.48
4. その他	2.29	1.68	2.29	1.64	2.22	1.64
全体	3.08	2.11	2.70	1.86	2.95	1.67

自宅を訪問する「訪問調査」では、有効回収に至るまでに要する訪問回数は、対象者の接触の容易さの指標となるだろう。こうした訪問回数は、対象者の属性によって異なると思われるが、そうした属性間の違いを考慮した上で、回答内容に対しても何らかの関係があるのだろうか。つまり回答を得るために接触することが容易な対象者と、そうではなかった対象者の間で、何らかの回答傾向の違いが、属性の違いを超えて観察されるのだろうか。この問題についても簡単に見ておくことにしよう。

以下完了票の回答内容と訪問回数の連関を検討結果の一部を紹介するのだが、あらかじめ述べておくと、訪問回数の多さの意味するところは、単純ではない。少ない訪問で協力が得ら

れた対象者は、在宅率も高いなど接触が容易で、かつ調査にも協力的な人たちだろう。他方、協力を得るために多数の訪問が必要であった人たちには、恐らく (a) 単に不在率が高くなかなか会えなかったけれども、調査そのものには協力的な方であったために会えれば容易に協力が得られた対象者と、(b) 接触は当初からできていたが調査には消極的で、訪問の繰り返しを通じてようやく説得に成功した（潜在的にはそれほど協力度が高くない）対象者など、複数のタイプを含んでいるからである。こうした点に留意した上で、結論を下すにあたっては慎重な態度が必要だろう。

なお調査不能バイアスの検討の文脈では前段のような「調査への協力傾向」といった潜在的な要因も考慮する価値がある。訪問回数そのものではなく別の項目に基づくものであるが、こうした点を検討した研究例もある（例えば、土屋(2010)などを参照）。

訪問回数の分布と回答との連関の例

さて、前節と同様に、三つの調査について、今度は完了票のみについての訪問回数の分布（5回以上をまとめる）をまとめたものが表3であ

表3 三つの調査での完了票についての訪問回数の分布

回数* ¹	日本人の国民性第13次全国調査				国民性に関する意識動向2012年度調査			
	K型		M型		面接		留置* ²	
	度数	%	度数	%	度数	%	度数	%
1回	508	31.9	509	32.2	526	34.4	896	42.4
2回	405	25.5	436	27.6	398	26.0	548	25.9
3回	317	19.9	280	17.7	274	17.9	360	17.0
4回	183	11.5	177	11.2	161	10.5	159	7.5
5回以上	178	11.2	177	11.2	169	11.1	151	7.1
合計	1,591	100.0	1,579	100.0	152.8	100.0	2,114	100.0

* 1 5回以上の訪問は一つのカテゴリにまとめた。

* 2 意識動向2012年度調査の留置票の訪問回数は初回訪問を除いた回数を数えた。

表4 訪問回数との連関が強かったカテゴリの例(M型)

	訪問回数					全体
	1回	2回	3回	4回	5回以上	
# 2.30d 交通事故の不安を“感じる”	59.7%	56.7%	54.6%	48.6%	49.2%	55.5%
# 2.13 若いときは楽しむより“将来に備えよ”	70.1%	65.6%	64.3%	62.7%	53.1%	65.1%

* # で始まる番号は、日本人の国民性調査での項目の整理番号

表5 訪問回数との連関が有意な項目数

検定結果	日本人の国民性第13次全国調査				国民性に関する意識動向2012年度調査			
	K型調査票		M型調査票		面接法		留置法	
	質問	属性	質問	属性	質問	属性	質問	属性
1: N. S.	67	5	54	6	52	4	61	4
2: $p < .05$	2	3	7	5	7	5	5	411
3: $p < .01$	1	10	4	7	16	5	10	6
計	70	18	65	18	75	14	76	14

る。既に述べたように留置法では調査票を渡す訪問が含まれるという意味で、面接法と訪問プロセスの内容が異なるので、便宜的に、2回目までの記録を1回目の訪問にまとめている。また全ての調査で5回以上の訪問回は少ないので、それらを一つのカテゴリにまとめている。

以下の連関の検討については少しイメージがわかりにくいかも知れないので、表4にはKS13のM型調査票から、連関が強くなった項目(カテゴリ)の集計例を二つ例示してみた。項目の内容は“交通事故に対する不安がある”，考え方として“若いときは(楽しむことよりも)将来に備えるべき”というようなものである。1回目か

ら5回目(以上)にかけて、ほぼ単調に変化しており、より詳細に検討すると訪問回と年齢の間の連関で説明される部分も多いようである。この結果が年齢との連関だけで説明されるとすれば、年齢以上に重要な情報を含まないことになる。しかし、それ以上の情報を含む(調査への協力傾向など、回答者の態度に関する何らかの情報を担う)とすれば、例えば調査不能バイアス調整のために活用しうる可能性がある。こうした点を統計的な検討課題とすべきということになる。

連関の検討結果

さて、訪問回数5回以上をまとめて、KS13に



表6 意識動向2012年度調査で、訪問回数との間の連関が強かった上位の項目の例

順位	面接	留置
1	問29.1. この3ヶ月の行動：地域活動に参加	問30.1. この3ヶ月間の行動：地域活動に参加
2	問29.11. この3ヶ月の行動：わからない、答えられない	問21. 居住年数
3	問22. 定住意向	問23. 近所とのつきあいの程度
4	問29.8. この3ヶ月の行動：地元の自治体のあり方について話題	問25D. 活動参加：地域の親睦
5	問29.3. この3ヶ月の行動：宗教施設に参拝	問24. 町内会・自治会の所属
6	問29.10. この3ヶ月の行動：特に該当するものはない	問17. 選挙への関心
7	問29.7. この3ヶ月の行動：地域の人と防災の話し合い	問25A. 活動参加：地域防犯
8	問21. 居住年数	問25C. 活動参加：地域の環境保全
9	問30. 調査協力理由	問8. 日本と個人の幸福
10	問25C. 活動参加：地域の環境保全	問13E. 不安感：原子力施設の事故
11	問25D. 活動参加：地域の親睦	問9. 幸福がためになることか
12	問16. 日本人全体の生活水準の変化	問25B. 活動参加：地域防災
13	問25B. 活動参加：地域防災	問19C. 日本の「経済力」

については同調査で利用している2種類の調査票(K型, M型)毎に、協力に要した訪問回数と回答者の回答の連関の検討結果をまとめたものが表5である。ここでの検討は訪問回数×回答内容のクロス表の χ^2 乗検定で有意であるか否か、という形で簡便に行っている。結果のまとめは、いわゆる属性に関する項目と一般の意識などを質問した項目(以下単に「質問項目」とに分けて示してある。また複数回答可の項目は便宜的に各回答選択肢の選択の有無を一つの項目とした単一回答項目であるかのように扱っている。DK2012で有意な連関を示す項目が多いことの原因の一部は、このような複数回答可の項目が含まれていたことである。なお、ここでの属性項目の中には、性別・年齢は含めていないが、年齢については訪問回数と明らかな連関があることが既に分かっている。

さて、全ての調査を通じて属性項目は概して訪問回数と連関しやすく、例えば「就業状態」のような項目はその典型である(当然の結果ではあるが、有職者のほうが訪問回数が増えるという方向の連関である)。なおKS13の属性項目はKM両調査票で共通、またDK2012調査についても面接法と留置法で属性項目は共通であるが、両調査票で概ね同じ項目が有意になった。訪問回数自体を被説明変数として複数の属性要

因を説明変数とした回帰分析を行うと、訪問回数が増えることに寄与する属性変数は面接法と留置法とでは多少異なるようであるが、就業状態が非常に強い要因であることは共通する。

さて、表6にはDK2012調査で「質問項目」のうち、訪問回数との連関が強かった項目の例を挙げた。紙幅の関係で質問文そのものを表示することはできないが、上位のほうから、項目内容の要点が分かるような短文の形で示してある。この調査では、「日本人の国民性調査」とは異なる独自項目を多数導入しており、そうした項目のうち「地域活動の多さ」と関わるものについての連関が多く観察される。

では訪問回数は、年齢(やそれ以外の属性)とは異なる独自の情報を持っているのか。詳細は割愛するが、前述の連関が強い項目について、訪問回数を説明変数に投入したロジスティック回帰分析により分析したところ、面接・留置ともに、属性要因(性別・年齢・都市規模・住居形態・学歴・就業状態・世帯内地位など)をコントロールした上でも、一部の項目では、訪問回数が少ない人のほうが地域活動に熱心というような方向の有意な連関が残った。つまり、少なくとも一部の項目に対しては、属性要因以上の説明要因となるように思えるという結果である。ただし、有意な連関といっても、効果量としては概して小

さく、「細かい糸でつながっている」というような状態の結果だった。したがってこの結論も暫定的なものと理解していただきたい。

5 おわりに

調査員の活動は、調査実施者が調査不能に関連する要因として直接制御しうる数少ない要因である（保田ほか, 2008）。例えば面接調査ではオートロック付きの集合住宅居住者の接触困難が強調されることが多いと思われるが、図3や図4に基づくならば、それ以外の集合住宅と大きな違いがあるわけではないとも受け取れる。接触が可能になった後での協力率も、他の住居形態と大きく違うわけでもない。例えば調査員教育の場で、説明のニュアンスを多少変えても

良いかも知れない。

調査協力者から協力を得るために必要となった訪問回数についての結論は、やや控えめに考えておきたい。既に述べたことであるが、協力を得るまでに必要だった訪問回数の意味は訪問経過に応じて複数ありうるので、例えば「調査への（非）協力傾向」という単一の変数——訪問回数が多い回答者ほど非協力的と見なされるような——と考えると検討事例数としても本稿で扱った調査だけでは全く不十分で、より多くの調査での検討の積み重ねが必要と考えている。パラデータの一部として訪問回数をバイアス調整に利用することを試みた先行研究もあるが（Olson, 2013）、ここではそうした方向の検討が日本の調査でも必要になるということのみを述べて、本稿を閉じることにしたい。

文献

- Fushiki, T. and Maeda, T., 2014, "Nonresponse Adjustments for Estimates of Proportions in the 2010 Survey on Stratification and Social Psychology", *Behaviormetrika*, 41(1): 99-114.
- 伏木忠義・前田忠彦, 2015, 「調査不能を伴う社会調査における推定——国民性に関する意識動向調査を題材に」『新潟大学教育学部研究紀要 自然科学編』7巻2号, 63-71。
- Matsuoka, R. and Maeda, T., 2015, "Neighborhood and Individual Factors Associated with Survey Response Behavior: A Multilevel Multinomial Regression Analysis of a Nationwide Survey in Japan", *Social Science Japan Journal*, 18(2): 217-232.
- 前田忠彦, 2005, 「郵送法の特徴に関する一研究」『統計数理』53 (1) : 57-81。
- 前田忠彦, 2016, 「第1回SSP調査の設計と実施概要」SSPプロジェクト事務局（吉川 徹・伊藤理史）編『2015年階層と社会意識全国調査（第1回SSP調査）報告書』3-9。
- 松岡亮二・前田忠彦, 2015, 「『日本人の国民性 第13次全国調査』の欠票分析——個人・地点・調査員の特性と調査回収状況の関連」『統計数理』63 (2) : 229-242。
- 中村 隆・土屋隆裕・前田忠彦, 2015, 「国民性の研究 第13次全国調査——2013年全国調査」『統計数理研究所調査研究リポート No.116』。
- Olson, K., 2013, "Paradata for nonresponse adjustment", *The Annals of the American Academy of Political and Social Science*, 645(1): 142-170.
- SSPプロジェクト事務局（吉川 徹・伊藤理史）編, 2016, 『2015年階層と社会意識全国調査（第1回SSP調査）報告書』SSPプロジェクト。
- 土屋隆裕, 2010, 「調査への指向性変数を用いた調査不能バイアスの二段補正——『日本人の国民性 第12次全国調査』への適用」『統計数理』58 (1) : 25-38。
- 保田時男・宍戸邦章・岩井紀子, 2008, 「大規模調査の回収率改善のための調査員の行動把握——JGSSにおける訪問記録の分析から」『理論と方法』23 (2) : 129-136。
- 保田時男, 2009, 「JGSSにおける調査員の訪問記録の分析」『JGSSで見た日本人の意識と行動——日本版General Social Surveys研究論文集 [8]』79-90。



特集論文

4

なぜ調査員の訪問記録を
分析するのか

保田時男

関西大学社会学部 教授

私はある時期、調査員の訪問記録の分析に大きな関心を傾けていた。一人ひとりの調査員がいつどのようなパターンで対象者の家を訪問し、いつ接触到成功・失敗して、どのように調査を終えたのか、というパラデータについて、その整備方法や分析方法を考え、実際にかなり綿密な分析を行ってきた。

その意味で今回の特集で執筆依頼があったことに驚きはしないのであるが、正直な気持ちを告白すると、執筆にはあまり気が進まなかった。編集委員からの提案は、たとえば私も深く関わっているNFRJ (National Family Research of Japan; 日本家族社会学会が5~10年おきに実施している全国調査) など、留置調査の訪問記録について新たな分析を行うことであった。この種のデータは十分に分析されずに放置されているくらいがあるので、その見直しには一定の意味があるのかもしれないと感じて引き受けたものの、いざとなるとやはりなかなか分析意欲が高まらない。

そこで落ち着いてその理由を考えていたところ、むしろここで考えたことをまとめた方が有意義ではないか、という思いに至った。なぜ訪問記録の分析をいまは行わないのか。あるいはなぜ「行すべき」ときがあるのか。前置き(言い訳)が長くなったが、本稿は、訪問記録の分析の意義を見つめ直した自省的なエッセイである。

大枠の考え方は過去の論考でもすでに示しているが、改めてそのエッセンスをまとめ直すとともに実際的な手続き(4節)についても概観す

る。規模の大小を問わず、社会調査の企画・設計を行う立場の方々に読んでいただけるとありがたい。

1 訪問記録の分析は事例研究である

最初に、なぜいま私が訪問記録の新たな分析を行わないのかという理由を一言で述べておくと、それは、訪問記録の分析は本質的に「事例研究」であると考えているからである。訪問記録の分析は、調査員の適切な制御を考えることにつながり、調査の改善(直接的には回収率の向上)に寄与する。しかし、私はその分析成果は相当な程度、事例的なものであって、安易な一般化には慎重であるべきと考えている。つまり、ある調査の訪問記録からわかった傾向は、その調査の継続調査であるとか、サンプリングの様式や調査会社が同様の調査など、類似性の高い調査の改善のためには役立つが、種類の異なる調査のためにはそう役立つものではない、ということである。

その意味で、たとえばNFRJの訪問記録の「事例研究」を、まさにいま行うべき理由は私には見つからなかった。直近に行われたNFRJ08はすでに実施から8年が経過しており、実査時の記録や記憶は十分に活用できない。また、それでも改めて分析を行うならば、逆に次期調査(NFRJ18)の調査票設計を具体化するもう少し後の段階の方が意義がある。いま訪問記録の分

析を行うことは、いかにも中途半端に感じられた。

そうは言っても、基本的に大きな傾向はそう変わらないのではないか、と思われるかもしれない。それはそのとおりである。しかし、我々(調査企画者)が訪問記録から知らなければならないことは、一般的な傾向ではなく、調査の改善につながるような「この調査ならではの」特殊な傾向のはずである。

かつて訪問記録の分析に手を付け始めた際に私がもっとも驚いたのはJGSS-2003の分析結果であった(保田, 2005)。この調査では難度が異なるA票とB票が用意され、各調査員は全員A票とB票をほぼ半数ずつ担当していた。当然ながら難度が高いB票の方が回収率が悪かったのであるが、訪問記録を分析すると、B票の対象者は拒否による非回収だけでなく、なんと不在による非回収の割合も高いことがわかった。対象者の不在率は、調査票の内容とはまったく何の関係もないはずなのに、A票に比べてB票の不在率は2割以上高かったのである。

これはつまり、難度の高いB票の場合は、調査員が少ない回数で早めに訪問を切り上げて「不在」と扱っていたということである。このことから過度に難度の高い調査票は、調査員の訪問の意欲にも影響する(拒否される可能性が高いので粘り強く訪問する気持ちが弱くなる)ことがわかり、改善のヒントにもなった。

訪問記録の分析の意義は、このように個別具体的な調査の改善につながる「事例研究」が第一である。その意味で、その調査事例についてもっとも知識を有している個々の「調査企画の当事者(あるいはそれに近い者)」が、適切な時期に分析を行い、後の改善に活かすことが望ましい。

2 調査員の制御の重要性

では、そもそも調査の改善(直接的には回収率の向上)のために、訪問記録の分析をなぜかく

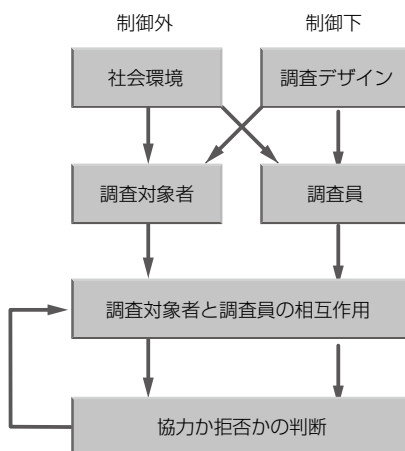


図1 調査協力の概念枠組み
出典: Groves and Couper (1998:30)

も重視するのか。私はGroves and Couper (1998: 30) の図式からその重要性が読み取れることを繰り返し主張している(図1)。この図式は訪問調査での調査協力が決まるまでの過程を整理しているが、ポイントは調査企画者が制御できる部分(右側)と制御できない部分(左側)を明確に区別していることである。つまり、調査企画者は回収率を向上させるために調査デザインや調査員の動かし方を改善することはできるが、社会環境や調査対象者の在り方に変化を求めることはできない。その意味で言えば、多くの回収率の分析が調査対象者の属性(性別や年齢、居住地など)の影響を調べることに注力していることは奇妙なことである。その種の分析はサンプルの偏りを調べるためには役立つが、回収率の改善を図るためには直接的には役立たない。

調査員こそが、調査デザイン(調査票やサンプルの設計)以外に制御可能な数少ない要素であり、調査を改善するためには調査員の行動をよく把握しなければならない。しかし、実査を調査会社に委託している場合には調査企画者が直接的に調査員に接触することはほとんどなく、その実際の行動を知ることは容易ではない。



その中で訪問記録は比較的簡便に調査員の客観的な行動パターンを知ることができ、また調査員に新たな負担を課す面も小さい(交通費の請求に際して訪問の記録はもともと必要なため)。非常に効率的でかつ効果的に調査員の行動様式を把握できる手段となるのである。

3 訪問記録の分析成果

実際的には、訪問記録の分析によってどのような知見を得ることができ、どのように調査の改善に活かすことができるのか。私が関わった過去の分析成果を紹介しておこう。

2005年SSM調査の予備調査という位置付けで実施された「2003年SSM日本予備調査」は、非回収(欠票)の状況について比較的詳しい情報が記録されており、非回収の決定日時、訪問回数、最終的に誰と会えたか、といった訪問記録もある程度収集されていた。2005年の本調査の改善に役立てるという明確な目標からこのデータを分析した経験は、私が訪問記録分析の有効性を強く感じる端緒となった(保田, 2005)。

いくつかの分析結果の中で意外であったのは、調査員1人の担当票数と訪問・回収状況との間の関係性であった。担当票数が多いと回収率が悪いことは予想どおりであったが、反対に担当票数が少なすぎる(11票以下)場合も回収率が悪く、12~14票の場合でもっとも回収率がよかったのである。訪問回数を分析すると、担当票数が少なすぎる場合には、早い時点で訪問を打ち切っていた。このことから、少ない票数のために調査地に赴くことは、調査員にとって負担感が大きいのではないかと解釈できた。

この分析結果が示すことは、より厳しいルールで調査員の訪問を制御すべき、ということではない。多くの調査員は、調査要領で定められる以上に訪問を継続しており意欲的であった。しかし、最終的にその意欲がどこまで保たれるかという点において、調査員の負担感を考慮す

る必要があるということである。

ただし、このような訪問記録の分析はその後のSSM調査の改善に十分には役立てられなかった。1つの理由はこの記録が非回収(欠票)のケースについてのみ収集されており、回収票について同様の記録がなくデータが中途半端であったことにある。このことからわかるようにSSM調査の訪問記録は、調査員の行動を把握するためというよりは、非回収(欠票)によるデータの偏りを検討することを第一の目的として収集されていたようである。実際、調査企画の幹事による分析はそのような視点を中心としていたし(三隅, 2005)、その後の本調査(2005年SSM日本調査)で収集された同様の記録についても、回収打ち切りの判断に至る過程を調査の基礎情報として残すような意味合いで分析されている(三隅・三輪, 2008)。私がとった分析視点は大きく異なったのであるが、調査企画者に近い立場にはなかったこともあり、残念ながらその成果はあまり活かす機会がなかったという思いがある。

一方、同じ時期に行ったJGSS (Japanese General Social Surveys; 日本版総合的社会調査)の分析成果は、あとに続く調査の改善にかなり積極的に反映させることができた。先に述べたとおりJGSS-2003の訪問記録の分析から、難度の高い調査票で訪問が早く打ち切られ、非回収が増えているとわかったことは驚きであり、調査票や調査手続きの難度の見直しにつながった。

また、JGSS-2003の訪問記録の分析からは、なぜか「不必要な」予備標本への訪問が多く見られることが明らかになった。転居など一定の要件に該当した欠票は、予備標本に置き換えられることになっていた。しかし、訪問記録を確認すると、なぜかそのような理由がまったく発生していない場合でも予備標本の訪問・回収が行われているケースが比較的多く見られたのである。このような予備標本は当然無効なので、調査会社からの納品データには含まれない。そ

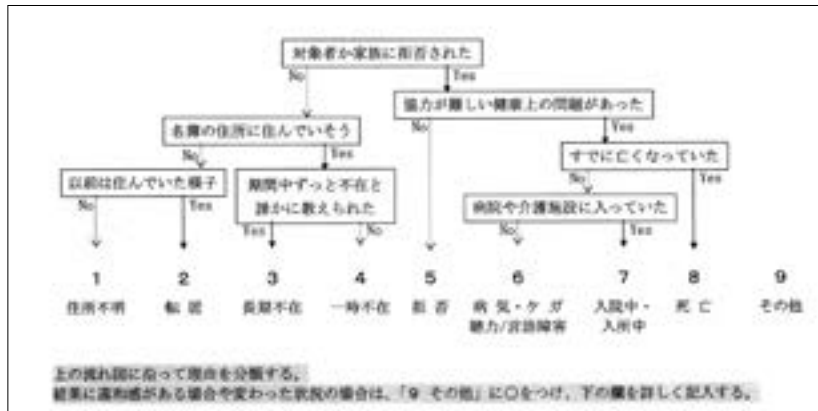


図2 JGSS-2006における欠票理由の分類フローチャート

	訪問日	訪問時間	話をした人	回収状況
1	月 日()	午前・午後 時 分頃	本人 家族のみ 不在	面接票 1 面接完了 (月 日) 2 面接不能 留置票 1 回収完了 (月 日) 2 回収不能 一方で調査不能の場合、次のページを忘れず記入する。
2	月 日()	午前・午後 時 分頃	本人 家族のみ 不在	
3	月 日()	午前・午後 時 分頃	本人 家族のみ 不在	
4	月 日()	午前・午後 時 分頃	本人 家族のみ 不在	
5	月 日()	午前・午後 時 分頃	本人 家族のみ 不在	
6	月 日()	午前・午後 時 分頃	本人 家族のみ 不在	
7	月 日()	午前・午後 時 分頃	本人 家族のみ 不在	
8	月 日()	午前・午後 時 分頃	本人 家族のみ 不在	
9	月 日()	午前・午後 時 分頃	本人 家族のみ 不在	
10	月 日()	午前・午後 時 分頃	本人 家族のみ 不在	
対象者本人に会えるまでの訪問を全て記録する。 10回を越えた訪問の記録は不要。 情報ができるまで、必ず5回以上訪問する。 (同日に2度訪問してもよいと数えるが、同日の訪問を禁止するわけではない)				留置票の受取予定日時

図3 JGSS-2006における訪問記録のフォーマット

のため、訪問記録を確認するまでこのような状態にはまったく気が付いていなかった。

なぜ調査員が無駄な予備標本を訪問していたのかというと、1つに非回収の理由が最終的にどのように分類されるか明確ではないので(一時不在に分類されるのか長期不在に分類されるのかなど)、念のために予備を確保しておくという側面があることがわかった。また、もう1つの理由として、調査員にとって予備標本への訪問の負担感が大きいことがわかった。つまり、何度か担当の調査地を訪問してその地点での回収

をほぼ終えたころになってから予備標本の必要が発生してしまうと、それから1票の予備標本のためだけに接触・回収という手続きを一からやり直さなければならなくなる。それならば、無駄になるかもしれないが、はじめからある程度の見込みで予備標本を回収しておいた方が負担が小さいというわけである。

このような知見をもとに、続くJGSS-2005からは予備標本の使用を全面的に取りやめた。その方が調査員の負担が軽くなり、総合的に見て回収率やデータの品質にプラスになると判断さ



れたためである。また、欠票理由の分類の仕方を明確にすることが、調査員の負担を減らす意味もあることがわかったので、欠票理由の分類をフローチャート式で明示することにした(図2)。さらに、これまでの訪問記録は訪問回数や最終的な訪問日時を記入するだけのものではあったが、より詳しく訪問のパターンを把握することの重要性が明らかになったので、毎回の訪問(日時、誰と会えたか)を記録するように改めた(図3)。また、訪問記録を記入するために調査員の負担が増えることがないように、これまで別紙(欠票調査票)に記載していた訪問記録の内容を調査票の表紙に配置し、調査資料の管理を容易にした。

このような工夫を凝らしたJGSS-2005であるが、結果的に回収率はうまく向上しなかった。しかし、さらに謝礼の渡し方や依頼状・封筒の見直しなどを行うことで改善の努力を続けたところ、JGSS-2006では回収率の格段の向上を見ることができた。JGSS-2003、2005、2006の公式回収率は、55.0%、50.5%、59.8%と推移した。

ここで重要なことは、訪問記録を精緻に収集していたおかげで、この回収率改善の内実を後から検証することができた、ということである。2005年は個人情報保護法の施行により社会調査への警戒心が過敏に高まった年である。この1年だけは、全国的に社会調査の回収率が一時的に低くなっている。その意味で、2005年から2006年への回復は、社会全体の流れであって、JGSSの改善成果ではないという疑いも持たれた。

保田・宍戸・岩井(2008)の分析は、訪問記録の分析から、このような社会全体の流れとは別に、調査員の訪問の粘り強さが改善されていることを明らかにした。つまり、JGSS-2006では前回調査(JGSS-2005)と比べて、訪問回数がかさんでも訪問継続率が維持されており、調査対象者に接触することが多くなっていることがわかった。接触後の協力獲得率の向上とはまったく別の改善成果である。

その他、JGSS-2006の訪問記録については、かなり多方面にわたる分析を行い、1人の調査員による訪問は1~2週間の短期間に集中しており、休日を過度に重視した訪問を行っていることや、同じ調査員は同じ時間帯ばかりに訪問していることなどがわかった(保田, 2009)。一方で、実際の調査対象者の在宅率は、平日と休日でそれほど大きな違いがあるわけではなく、休日の方が、接触できた際に協力を拒否される確率はむしろ高い。また、在宅時間は個人差が大きく、同じ時間帯ばかりの訪問では接触しにくい対象が出てくる。

これらの意味で、調査員の訪問のパターンには改善の余地があることがわかった。このような分析は、調査要領に調査対象者の在宅率の傾向を示すなどの工夫につながっている。

4 訪問記録の収集法・分析法

繰り返すが、訪問記録の分析は「事例研究」なので、上記のような傾向がどの調査でも同じように見られるとは限らない。その意味で、それぞれの調査において、調査企画者が自身の調査について訪問記録を収集し、分析するという行為自体がもっとも重要となる。

訪問記録の分析結果は安易に一般化できないが、その収集・分析の方法にはある程度一般的なプロセスが成り立つであろう。ここでは私の経験上のノウハウを簡潔に示す。

まず、前提として各訪問の結果は図4のように「A協力」「B拒否」「C接触断念」「D訪問継続」の4つに分類できる。ある訪問で対象者に接触できれば、その結果は、協力が得られるか(A)、拒否されるか(B)のどちらかである。一方、接触できなかった場合には、そこで訪問を打ち切るか(C)、再度訪問を試みるか(D)のいずれかである。

このように分類すると、訪問が成功する(調査票が回収できる)までのプロセスを以下の3つ

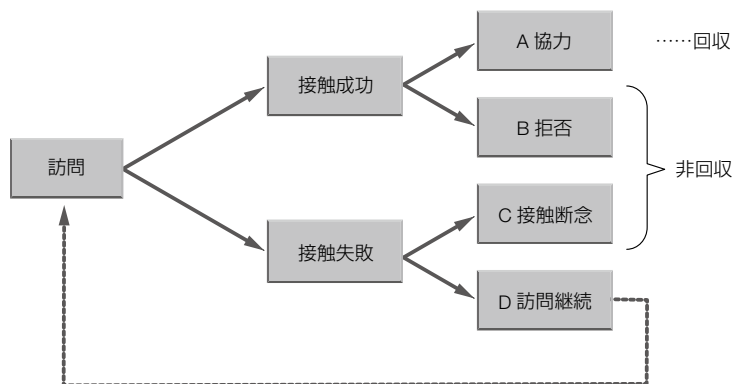


図4 各訪問の結果の分類

に分解して考えることができるようになる。

$$\text{接触成功率} = (A+B)/(A+B+C+D)$$

$$\text{協力獲得率} = A/(A+B)$$

$$\text{訪問継続率} = D/(C+D)$$

3つのいずれも高いほど回収率の向上に役立つが、それぞれ意味合いがまったく異なるので、これを分けて考えることができる点に訪問記録を詳しく収集する意義がある。

実際的には、すでに図3で示した形式で記録をとれば、1回ずつの訪問についてこのような分類が可能になる。次回の訪問が続いている場合は、当然その回の訪問結果は「D 訪問継続」ということになる。一方、その回で訪問が終了している場合には、調査票の回収状況と照らし合わせることで最終の訪問結果を分類できる。調査票が回収できていれば「A 協力」であるし、非回収でその欠票理由が拒否の場合には「B 拒否」、それ以外（一時不在など）の場合には「C 接触断念」となる。

このような判断を下すためには、訪問の記録が「対象者本人とはじめて会えるまで」で打ち切られていなければならない。つまり、たとえば留置調査で再訪問して調査票を回収したときのような訪問は訪問記録からは除かなければなら

ない。そのために、各回の訪問で「話をした人」を記録することが重要になる。たとえば、本人と話をした訪問の記録が2回続く場合には2回目の訪問は調査票回収のための再訪問で無効なので記録からは削除しなければならない。

訪問調査のデータにもこのようにいくつかの点からクリーニングが必要になる。たとえば、最終の訪問で「本人と話をした」に○が付いていないのに、調査票が回収できている場合には、最後の訪問の記録が抜けているのか、あるいは訪問は記録されているが、○を付け忘れているのか、他の記録（回収日時など）から判断しなければならない。あるいは、日付の記入間違いも意外と多い。ミスの内容を推測しやすくするために月日だけでなく曜日を記入してもらうことを勧める。また、「話をした人」という文言を用いているのは、「接触した人」や「会った人」のような別の文言だと、インターホン越しに話した場合などに、記入に戸惑う調査員がいるためである。話をした人の選択肢に「家族」があるのは、近年、本人ではなく家族に調査協力を拒否されるケースが多くなっていることによる。本人による拒否と家族による拒否はまったくメカニズムが異なるので注意が必要である。たとえば、JGSS-2006における各種の方策は対象者本人による拒否を減らすことには成功したが、家族に

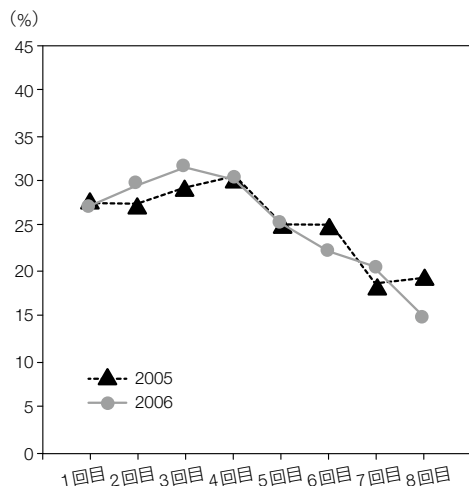


図5-a JGSS-2005と2006における各回の訪問結果から得た接触成功率

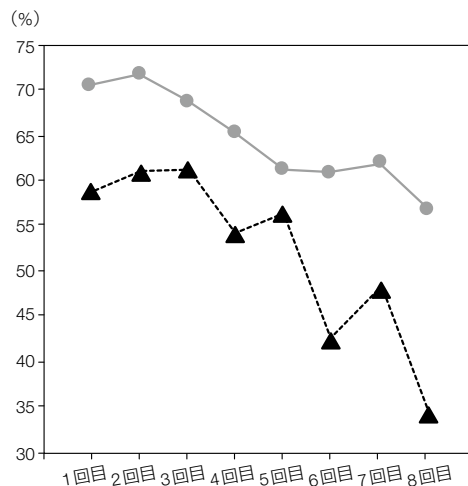


図5-b 協力獲得率

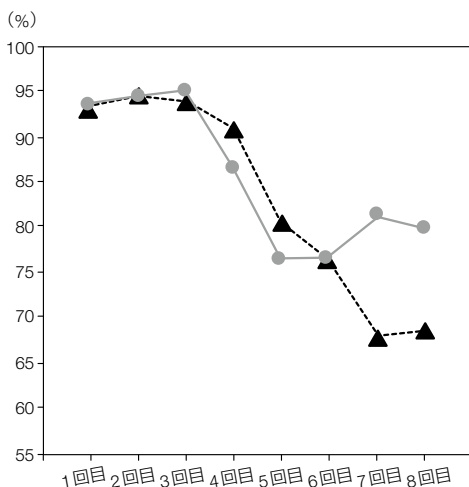


図5-c 訪問継続率

よる拒否を減らすことはまったくできなかった(保田・宍戸・岩井, 2008)。

適切にクリーニングができれば、そのデータはパネルデータの技法で入力・分析できる。つまり、各回の訪問を1レコードとするロングデータに整理して分析を行えばよい。先に示したように、各回の訪問結果を「A協力」「B拒否」「C接触断念」「D訪問継続」の4つに分類し、そこから算出される接触成功率、協力獲得率、訪問継続

率がどのようなパターンで変動するかを確かめることが分析の基本となる。たとえば、図5は各回の訪問における3つの確率の変動の仕方を、JGSS-2005と2006で比較している(保田・宍戸・岩井, 2008)。接触成功率にはほとんど違いが見られないこと、協力獲得率は一貫してJGSS-2006の方が高いこと(しかも後の方の訪問回で差が大きくなること)、JGSS-2006では7回目以降の訪問継続率が落ちていないこと(調査員が粘り強くなっていること)などがわかる。

訪問記録のデータに加えて調査員の情報(性別、調査員としての経験年数、担当票数など)や調査地点の情報(都道府県、市郡規模など)があれば結合して分析に活かすことができる。ただし、その場合はマルチレベルの層を増やした分析が必要になるし、調査員と調査地点の対応関係を考慮することに注意しなければならない。統計的にはやや複雑なデータとなる。

5 訪問記録の分析は将来への投資である

ここまで調査員の訪問記録を収集・分析する意義と具体的な成果・方法について概括した。

適切にデータを収集・分析することは意外と大変であり、もしかすると、訪問記録の分析のためにここまでの労力を割く余裕はないと感じるかもしれない。だが、経験則から言えば、訪問記録の分析にはそれだけの労力を払う価値がある。

さらに言えば、このような周辺のパラデータの収集・分析に労力を避ける余裕をもって調査企画を立てなければならないのだと思う。やや大げさな話になるが、日本の学術的社会調査は、全般的にぎりぎりの人員と予算で必要なデータをなんとか収集・分析しているくらいがある。調査の規模を縮小したり、テーマの範囲を絞り

込んだりしてでも、訪問記録を分析できるくらいの余裕を確保した方がよいのではないだろうか。その投資が、長期的には将来の調査を地道に改善することにつながっている。

【付記】日本版General Social Surveys (JGSS) は、大阪商業大学JGSS研究センター（文部科学大臣認定日本版総合的社会調査共同研究拠点）が、東京大学社会科学研究所の協力を受けて実施している研究プロジェクトである。

文献

Groves, R. M. and Couper, M. P. 1998, *Non-response in Household Interview Surveys*. Hoboken: John Wiley & Sons, Inc.

三隅一人, 2005, 「2003年SSM日本予備調査における回収状況の基礎分析」尾嶋史章編『現代日本におけるジェンダーと社会階層に関する総合的研究』(科研報告書): 277-294。

三隅一人・三輪 哲, 2008, 「2005年SSM日本調査の欠票・回収状況の分析」三輪 哲・小林大祐編『2005年SSM日本調査の基礎分析——構造・趨勢・方法』(2005年SSM調査シリーズ1): 17-29。

保田時男, 2005, 「回収率の改善のための調査員の制御についての考察」尾嶋史章編『現代日本におけるジェンダーと社会階層に関する総合的研究』(科研報告書): 319-327。

——, 2009, 「JGSSにおける調査員の訪問記録の分析」『日本版 General Social Surveys研究論文集[8]』: 79-90。

保田時男・宍戸邦章・岩井紀子, 2008, 「大規模調査の回収率改善のための調査員の行動把握——JGSSにおける訪問記録の分析から」『理論と方法』23 (2): 129-136。



特集論文

5

電話調査における
パラデータの活用

城川美佳

富山大学大学院医学薬学研究部(医学)
公衆衛生学講座 助教

1 はじめに

電話調査の実施においては、電話番号と利用者のリストを入手できない場合が多い。社会調査で対象者抽出に利用されることの多い住民基本台帳や選挙人名簿に、電話番号が記載されていないためである。代替リストとして電話帳が考えられるが、電話帳への掲載には職業による偏りがあることが知られており、また世帯人員数による掲載の有無の違いも指摘されている(城川・西川, 1998)。近年では、電話帳への掲載拒否の割合が増えたことに加え、NTT(旧電電公社)の回線を用いない電話回線も出てきたことから、電話帳は対象者抽出には適さないと考えられている。

1978年にジョセフ・ワークスパーグが発表したRandom Digit Dialing法(以下、RDD法)の変法(Waksberg, 1978)は、電話番号が、①数字の羅列によって作成されている、②地域を表す番号と回線を所有する世帯を表す番号で構成されている(図1)。この性質を持っていることを利用して、リストに頼らずに電話番号を用いて世帯抽出する手法である。この手法は社会医学的分野においても海外で利用されており、米国のBRFSS(The Behavioral Risk Factor Surveillance System: 健康リスク関連行動に関する調査)は本法で実施している代表的な例である。

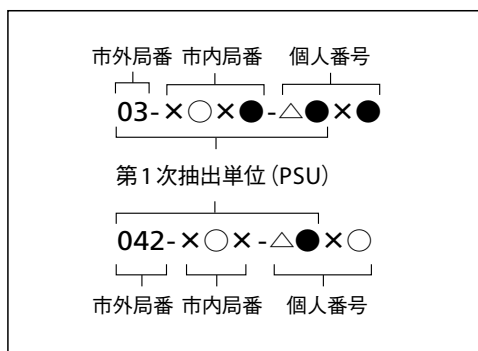


図1 電話番号の構成

本法はリストがなくても世帯抽出を可能とするが、一方で、電話番号が事業所あるいは居宅の区別なく割り当てられているため、実際に調査を行う際に業務用の回線か居宅が所有している回線かを確認する過程が加えられている。また、日本では行政区と市外局番・市内局番が示す地区に多少のずれがあるため、事前に市外局番と市内局番をリストアップしたうえで、調査対象地域から外れている可能性についても、確認する必要がある。

概して、事業所は連続して電話番号を取得することが多い。また、調査対象外の電話番号や未使用の電話番号も連続していることが考えられる。ワークスパーグ変法では、このような電話番号の特性を利用して、不必要に電話をかける回数を減らす工夫を取り入れている。すなわち、地域を示す番号(市外局番、市内局番)に個

表1 PSUの回答状況と回答を得るまでの電話回数

	未使用	地区外	業務用	FAX	拒否	不在	回収
1回	569		183	75	218	39	37
2回	6	1	23	4	63	4	20
3回			13	1	22		7
4回	1		5	1	11		4
5回			6		14		2
6回			5	1	4		6
7回			2		5	2	2
8回				1	2	106	4
9回					2	9	
計	576	1	237	82	341	150	82

※単位：回線

表2 電話回線の状況の比較

	未使用	地区外	業務用	FAX	拒否	不在	回収
PSU	38.8	0.1	16.0	5.5	23.0	10.8	5.5
本調査	26.5	0.0	8.6	5.9	19.9	28.5	10.7

※単位：%

表3 PSU抽出におけるアクセス回数別の本調査での実施状況

	1回	2回	3回以上	計
未使用	211 (22.7)	142 (22.0)	390 (31.8)	743 (26.5)
業務用	97 (10.4)	54 (8.3)	89 (7.3)	240 (8.6)
FAX	58 (6.2)	42 (6.5)	65 (5.3)	165 (5.9)
拒否	193 (20.8)	134 (20.7)	229 (18.7)	556 (19.8)
不在	242 (26.1)	200 (30.9)	355 (29.0)	797 (28.5)
回答	128 (13.8)	75 (11.6)	97 (7.9)	300 (10.7)
計	929	647	1,225	2,801

※単位：回線 (%)

人番号の上2桁を加えた8桁で示される番号を第1次抽出単位 (Primary Sampling Unit; PSU) とし、PSUに下2桁の数字を加えた代表電話番号に実際に電話をかけ、事業所や未使用などの対象外の電話番号、調査参加に同意を得られなかった電話番号のPSUを除外する。調査で使用するPSUを抽出したうえで、無作為に抽出した下2桁を加えた電話番号に電話調査を行うのである (PSUは地区を示す番号が付帯しているので、各PSUからの抽出数を割り当てることも可能である)。このような手順を踏んで実施する電話調査では、PSUの抽出、対象世帯への依頼、対象者への依頼、の各段階で情報を取得できる。さらには、追跡調査の実施では、電話番号をその番号が付置された電話回線を所有している世

帯の番号として利用することも可能である。

本稿では、電話調査におけるパラデータについて、私がRDD法を用いて2002-2006年に実施した社会医学領域での調査 (20歳以上の居住者を対象) で取得したものをを用いて概説する。

2 第1次抽出単位 (PSU) の状況と本調査での実施状況

先述したように、ワークスパーグ変法ではPSUに2つの数字を加えたものをPSUの代表番号として調査を実施し、回答が得られたPSUを用いて調査を進めることとしている。PSUの抽出において、代表番号の調査実施状況で各PSUでの調査実施可能性を予測させるものであれば、パラデータとして有用な情報となると考える。



表4 電話回線の状況

	2004年実施 東京都特別区 12月土・日の2日間	2005年実施 北海道S市 12月金・土の2日間
番号作成標本数	3,429	4,064
有効回収	305 (8.9%)	319 (7.8%)
対象の該当者がいる世帯数	449 (13.1%)	432 (10.6%)
有効回答率	67.9%	73.8%
未使用 (未使用, 移転, 故障含む)	906 (26.4%)	1,395 (34.3%)
業務用	209 (6.1%)	249 (6.1%)
FAX	153 (4.5%)	141 (3.5%)
拒否	786 (22.9%)	920 (22.6%)
対象者不在	65 (1.9%)	63 (1.6%)
呼出音のみ	585 (17.1%)	467 (11.5%)
留守番電話のみ	343 (10.0%)	392 (9.6%)
話し中	30 (0.9%)	35 (0.9%)
対象外 (該当者なし, 不明)	24 (0.7%)	33 (0.8%)
その他	23 (0.7%)	50 (1.2%)

※単位:回線 (%)

2002年に実施した調査 (東京都特別区を対象として実施) では、1,483群のPSUに調査を実施し、82群を抽出した。

除外されたPSUのうち、代表番号が未使用だったものは、ほとんどが1回アクセスしただけで確認できた (表1)。FAXの確認は、同一電話回線を通話とFAXに切り替えて使用している可能性も考慮して、複数回アクセスするという作業を加えたため、確認終了までのアクセス回数が未使用よりも若干多くなっている。業務用回線については、1回で判明したのは全体の77%であったが、最大で7回アクセスしている。拒否と回答が得られたPSUの結果が得られるまでのアクセス回数は、同様の推移が見られた一方で、対象者が不在と判明するまでに要したアクセス回数は、2回以下と7回以上に二極化された。対象者が不在の場合、電話対応者から対象者が在宅している時間や曜日に関する情報を収集して、次の電話に利用しているが、それでも直接話ができない場合が多いと考えられる。

PSU抽出時と本調査での実施状況とを比較すると、未使用、地区外、業務用回線の割合は減少しており、この作業の有用性が確認できた (表2)。また、PSU抽出時に回答を回収するまでのアクセス回数で1回、2回、3回以上の3群に分け

たとき、各群での調査時の実施状況による違いが見られた (表3)。すなわち、代表番号に1回アクセスして抽出された (調査が実施できた) PSUは、2回、あるいは3回以上アクセスが必要だったPSUに比べて回答を回収できた回線数が多かったが拒否の割合も多かった。不在のまま終了した回線の割合が少なく、比較的アクセスが容易だったことが推察される。一方で、PSU抽出時に複数回のアクセスを要した群に若年層の回答が多かった。本調査の際には、PSU抽出時のアクセス回数にこだわらず、むしろ複数回アクセスを要したPSUでは十分にアクセスすることが望ましいと考えられた。

3 電話調査における調査環境

表4は、2004年と2005年に実施した調査での電話回線の状況を示したものである。同一の調査目的、依頼文、調査票を用いて、東京都特別区と北海道都市部のある市 (S市) を対象地域としている。本調査で利用した電話番号の回線数に対して、S市では未使用が多く、東京都特別区で呼出音・留守番電話が多い傾向が見られるものの、概して同様の傾向を示している。パラデータから見る限り、都市部に居住する成人を対象

表5 調査実施状況とアクセス状況

		回答 (n=300)	拒否 (n=556)	計 (n=856)
アクセス回数	1回	210 (70.0)	462 (83.1)	672 (78.5)
	2回	51 (17.0)	57 (10.2)	108 (12.6)
	3回	17 (5.7)	22 (4.0)	39 (4.6)
	4回以上	22 (7.3)	15 (2.7)	37 (4.3)
曜日	月曜日	34 (11.4)	55 (9.9)	89 (10.4)
	火曜日	37 (12.4)	43 (7.7)	80 (9.4)
	水曜日	35 (11.7)	31 (5.6)	66 (7.7)
	木曜日	43 (14.4)	66 (11.9)	109 (12.8)
	金曜日	47 (15.7)	120 (21.6)	167 (19.5)
	土曜日	39 (13.0)	108 (19.4)	147 (17.2)
	日曜日	64 (21.4)	133 (23.9)	197 (23.0)
時間	9-12時	74 (24.7)	117 (21.0)	191 (22.3)
	13-16時	106 (35.3)	210 (37.8)	316 (36.9)
	17時以降	120 (40.0)	229 (41.2)	349 (40.8)

※単位: 回線 (%)

とした調査では、その電話回線の使用世帯に地域による違いはないと考えられる。

4 電話調査の曜日・時間

電話調査においても、他の調査手法と同様、対象者につながるまで何度かアクセスを図る必要がある。曜日や時間といった電話をかけるタイミングは、回答の回収の有無と関連していると考えられる。パラデータを用いて、電話をかけるタイミングを検討することができる。

2002年調査における調査回答と拒否でタイミングの状況を比較した(表5)。アクセス回数では、調査参加拒否となった群のほうがアクセス回数も少ない。また、アクセスした曜日では、拒否は週末に多くなっており、回答が得られた回線が日曜日にやや多いものの全体に分布しているのと異なっている。しかしながら、アクセスした時間帯は両者に違いは認められなかった。

2002年調査の調査回答者の調査状況を検討した(表6)。回答を最終的に得られた曜日は、日曜日が最も多く、次いで金曜日、木曜日、と続いた。また、時間帯では時間が遅いほど回答が得られている。1999年に私たちが東京都特別区に居住する成人を対象とした別の調査結果(城

川, 2003:115) とは曜日別の分布が異なり、週末に向かって回答割合が増えているようである。但し、土曜日での回答割合が少なくなっており、対象者のライフスタイルの変化が影響している可能性がある。

調査状況を回答者の属性で検討すると(表7-9)、分布状況は1999年調査のものと多少分布が異なるものの、回答者の性別や年齢は回答を得られた曜日によって分布が異なっており、また世帯人員数の違いは、回答を得られた曜日だけでなく、時間帯も関連している。このように、異なる回答者の属性を広く収集しようと考えた場合、様々な時間帯や曜日でアクセスを図る必要があると考えられる(城川, 2003:117)。

5 電話調査回答者の追跡調査における脱落

日本の場合、電話番号はその世帯が地区外に移動、あるいは回線使用を解約しない限り、同一の世帯が所有している。従って、先述したように電話番号を世帯番号として使用することが可能である。初回の調査で対象者が特定できるので、その回答者を追跡することが可能と考えられる。私たちの調査研究では、2002年調査の参加者、および2004年新規調査の参加者を



表6 曜日・時間帯別回答者数

	午前	昼間	夜間	計
月曜日	2.7	3.3	5.7	11.7
火曜日	1.7	4.3	6.3	12.3
水曜日	5.7	5.3	0.7	11.7
木曜日	3.0	2.3	9.0	14.3
金曜日	0.0	6.3	9.3	15.6
土曜日	3.7	5.0	4.3	13.0
日曜日	8.0	8.3	5.0	21.3
計	24.7	35.0	40.3	100 (300)

※単位：％（人）

表7 性別曜日別。時間帯別回答者数

		男性 (n=94)	女性 (n=206)	計 (n=300)
時間帯	午前	26.6	23.8	24.7
	昼間	36.2	34.9	35.3
	夜間	37.2	41.3	40.0
曜日	月曜日	10.6	11.7	11.4
	火曜日	12.8	12.2	12.4
	水曜日	7.4	13.7	11.7
	木曜日	13.8	14.6	14.4
	金曜日	11.7	17.6	15.7
	土曜日	19.1	10.2	13.0
	日曜日	24.5	20.0	21.4

※単位：％（人）

表8 世帯人数別曜日別。時間帯別回答者数

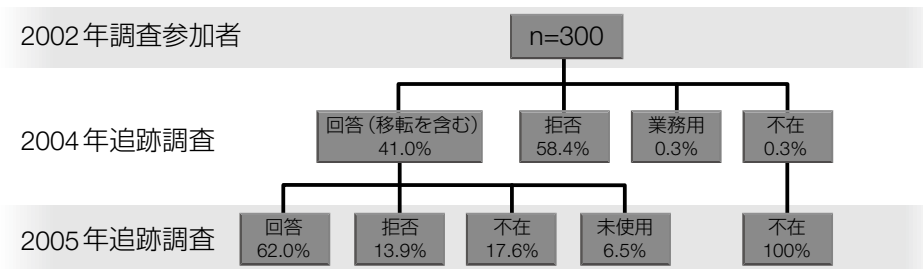
		1人 (n=74)	2人 (n=83)	3人以上 (n=144)	計 (n=300)
時間帯	午前	31.1	28.9	18.7	24.7
	昼間	32.4	31.3	38.9	35.3
	夜間	36.5	39.8	42.4	40.0
曜日	月曜日	16.4	8.4	11.1	11.4
	火曜日	10.9	10.8	13.9	12.4
	水曜日	9.6	16.9	9.7	11.7
	木曜日	9.6	13.2	17.4	14.4
	金曜日	9.6	14.5	19.4	15.7
	土曜日	15.1	14.5	11.1	13.0
	日曜日	28.8	21.7	17.4	21.4

※単位：％（人）

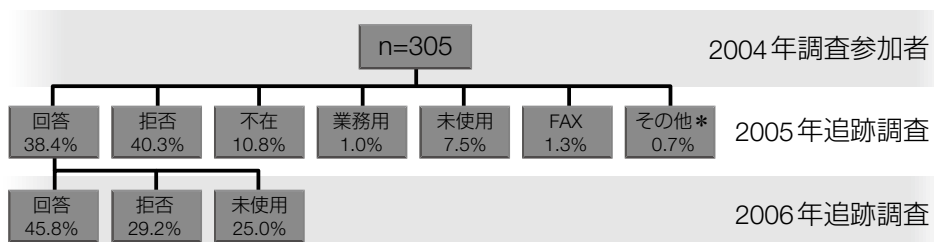
表9 年齢群別曜日別。時間帯別回答者数

		20・30歳代 (n=76)	40・50歳代 (n=112)	60歳以上 (n=112)	計 (n=300)
時間帯	午前	26.3	21.4	26.8	24.7
	昼間	27.6	33.9	42.0	35.3
	夜間	46.1	44.7	31.2	40.0
曜日	月曜日	13.2	14.3	8.1	11.4
	火曜日	9.2	9.8	17.1	12.4
	水曜日	7.9	10.7	15.3	11.7
	木曜日	14.5	14.3	13.5	14.4
	金曜日	18.4	20.5	9.0	15.7
	土曜日	17.1	12.5	10.8	13.0
	日曜日	19.7	17.9	26.2	21.4

※単位：％（人）



a) 2002年に初回参加した群



b) 2004年に初回参加した群

* その他: 死亡0.35%, 対象地区外に移転0.35%

図2 調査参加者の追跡調査

表10 追跡調査での調査アクセス時間

		2004年調査		
		午前	昼間	夜間
2002年調査	午前	54.6%	24.2%	21.2%
	昼間	38.1%	38.1%	23.8%
	夜間	14.0%	34.0%	52.0%

2006年まで追跡調査している(図2)。2002年調査参加者のうち、2004年の追跡では1回線が業務用となっており、また1回線は電話番号が新しい電話番号に転送された(対象地域内の移動)。さらに2005年の追跡調査では、新たに未使用に変わった回線が発生した。2004年に初回調査を実施した群では、2005年に業務用、未使用への変更が発生しており、対象地区外への移転、対象者の死亡も発生した。2006年の追跡では、未使用への変更が発生していた。また、追跡調査への参加割合は、1回目の追跡で約4割に減少し、2回目の追跡ではさらに半減した。同一対象の追跡調査に有線の電話番号を使用すること

は可能と考えられた。ただし、電話番号自体は世帯を示す番号であり、対象者の状況が変化すること、また追跡までの期間内に電話番号の使用形態が変更される可能性を考慮すべきである。

なお、2002年調査と2004年調査の両方に参加した者において、同様のアクセス時間で回答が得られる様子が見られた(表10)。追跡調査を実施する際には、初回調査時に取得したパラデータ(アクセス曜日や時間)を参考にできる可能性が高い。

6 さいごに

電話調査は、調査員管理が容易であり、また対象者のライフスタイルに合わせて、調査実施の時間帯をある程度延長することが可能という利点がある。その利点がどのように調査実施状況に影響しているかを検討することは、今後の電話調査にとって有用な情報であると考えられる。



平成25年3月の時点では、世帯数5,557万(総務省, 2013a: 28) に対して固定電話の契約数は5,681万(総務省, 2013b: 353)と1世帯1回線は所有している計算となり、調査での電話利用は可能と考えられる。しかしながら固定電話の契約数は減少の一途を辿る一方で携帯電話の契約数は増加し、その契約数は固定電話の契約数を大きく上回っているのが実情である。特に若い1人世帯は携帯電話回線のみを利用している

様子が窺われる。電話調査も固定電話から携帯電話の利用の検討に移行しつつある。しかしながら、携帯電話の電話番号は、スマートフォン、PHS等の携帯電話種類、回線業者、契約時の店舗所在地の情報が含まれているものの、肝心な利用者の現在地を示す情報がないために、ワークスバーク変法とは異なった工夫が必要である。電話調査の利点を活かした手法の確立が望まれる。

文献

城川美佳・西川浩昭, 1998, 「電話調査回答者の属性と電話帳収載状況」『日本公衆衛生雑誌』45(4): 309-319。

城川美佳, 2003, 「医療・保健領域の電話調査, Waksberg変法」『行動計量学』30(1): 111-119。

総務省, 2013a, 『住民基本台帳に基づく人口、人口動態及び世帯数(平成25年3月31日現在)』(http://www.soumu.go.jp/main_content/000244523。

pdf, 2016年11月30日確認)。

——, 2013b, 『平成25年版情報通信白書(PDF版)』(<http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h25/pdf/index.html>, 2016年11月30日確認)。

Waksberg, J., 1978, “Sampling Methods for Random Digit Dialing”, *Journal of the American Statistical Association*, 73(361): 40-46.

6

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

大隅 昇

統計数理研究所 名誉教授

矢口博之

東京電機大学理工学部 准教授

林 文

東洋英和女学院大学人間科学部 名誉教授

簗原勝史

株式会社シー・エス・マーケティング・ジャパン 代表取締役社長

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調査の質の確保が非常にむずかしくなっている。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、検証すべき課題も無数にある。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・鳩^{にお}, 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、本特集の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ある。ここではまず、われわれ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。さらに、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回答所要時間（パラデータから得られる情報のごく一部）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、早くからオンライン調査に関する多様な研究が行われてきた（詳しくは以下を参照：Bethlehem and 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比べ、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）。その後、2003年までに、継続的に実験調査（第2次～第4次）を進めた（以上「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）他、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（おもなウェブ調査が52回）の調査を行った。

いずれの実験調査でも「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（詳細は大隅ほか、2000；大隅ほか、2004；大隅・前田、2007, 2008）。その中に「トラッキ

ング・データの記録分析」「公募型・非公募型パネルの比較」がある。

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、意味するところはほぼ同じである。「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。

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「公募型と非公募型」としたが、これは欧米研究でい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。Couper (2000) の8分類によると、「公募型」はボランティア・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）。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。とくに、パラデータの分析からみ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。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）から回答者側（調査対象者）を観察・測定し、そのデータを調査の質の改善、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でもない（Couper, 1998; Couper, 2017; Nicolaas, 2011）。

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, 2005; Kaczmirek, 2009; Heerwegh, 2004; Heerwegh, 2011; Kreuter and Casas-Cordero, 2010などを参照）。

- 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・個人単位で取得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、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測定記録することで、通常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、回答者別に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, 2005; 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)に送ることによって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として取得したデータのほうが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とるか、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(マウス・クリック、キーストローク、タップなど)が生起した時点を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(商用)ソフトやサービスが多数公開されている(Click-density, 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(2011)の作成したJavaScriptを用いたソフトや、Kaczmarek(2009)のUCSP(Universal Client Side Paradata)などがある。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で、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なども利用可能である。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は表2の〈調査1〉で得られた文字列情報の一部で、ある回答者について、8ページの調査票のうち、はじめの2ページの追跡状態を示してある。

文字列はASCIIテキストファイルで、“`"`”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で、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。統

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=", "Q
 8_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)

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

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)	マイクロ秒 (ms) 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	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	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	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	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

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。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。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。

5 パラデータの簡単な分析

分析例とする調査データを、過去の実験調査から2つ選びその特徴を表2に要約した。2つの調査は2節で述べた区分に従うと、[第Ⅲ期]の一部に相当する。パラデータの特徴がよくみえ



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）
調査名略称	基底意識調査	本と読書に関する調査
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）
調査対象地域	首都圏（埼玉県、千葉県、東京都、神奈川県）	
ウェブ・パネルの種類	いずれも非公募型	
全質問数とその内訳	41	27
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
②テキスト・フィールド数	20	15
③自由回答の質問数	4	1
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]
調査票のページ数 （ウェルカム・ページを含む）	8	16
質問項目のおもな特徴	●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あ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	
	●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	
回答制御の有無	〈調査1〉〈調査2〉のいずれも「部分的に回答制御」として以下を設定 ●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しページ間の回答行動の履歴をすべて記録	
回答回収の強制打ち切り	いずれも行わない（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」	

るように、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）、ページ数、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。

調査の質を測る指標

前述のパラデータ取得と集約度の各段階に応じて、さまざまな指標が得られる。代表的な指標として回答率があるが、これだけでは十分とはいえない。回答率は、調査依頼者数（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（なお、非確率的パネルの場合は、“参加率”の利用が勧められている。AAPOR, 2016; ISO26362; Callegaro and DiSogra, 2008）。回答率をパラデータに含めるかどうかについては異見がある。提唱者の

Couper (2017) も、これをパラデータと考えるかどうかには慎重な議論が必要であると指摘している。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。一方、回答を完了できない無回答や中断を「無回答率」「中断率」で測ることができる。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重要である。ウェブ調査では高い回答率や参加率は期待できないので、無回答情報の分析は重要である。回答率や参加率、完了率や無回答率、中断率などを総合的に調べることはきわめて重要だがここでは指摘にとどめる。

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)。回答者は調査依頼を確認し、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、文字列の書き込み)、そのたびに各動作とその発生時点、その所要時間(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(なお、スマートフォンやタブレット端末の場合、クリックをタップやスタイラス・ペンによるタッチと読み替える)。

“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、“時間変数”として、回答時間(クリック時点)、ページ時間(データ送信時間)、マウス移動時の運動量記録(位置、軌跡など)、思考時間(回答内容を考える時間)、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)と設定し、これらの総和を、ページを配信後、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)”とする。こうした時間変数にもとづき、回答者が調査票の開始ページから終了ページまで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)の分析が、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)、パラデータの種類の質の違いはあるのか、上述の時間変数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だが、ここではSSPから得られるもっとも簡単な“ページあたり「回

答所要時間」”を調べる。取り上げた2つの調査は、表2にある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かなり違いがある(詳細は略)。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。よって1ページあたりの質問数が多い。

ここで〈調査1〉について、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2、ページごとの記述的統計量の評価分析が重要だが略した)。図2の回答所要時間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。標準偏差の他、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する。
- 変動係数が大きく、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間の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、

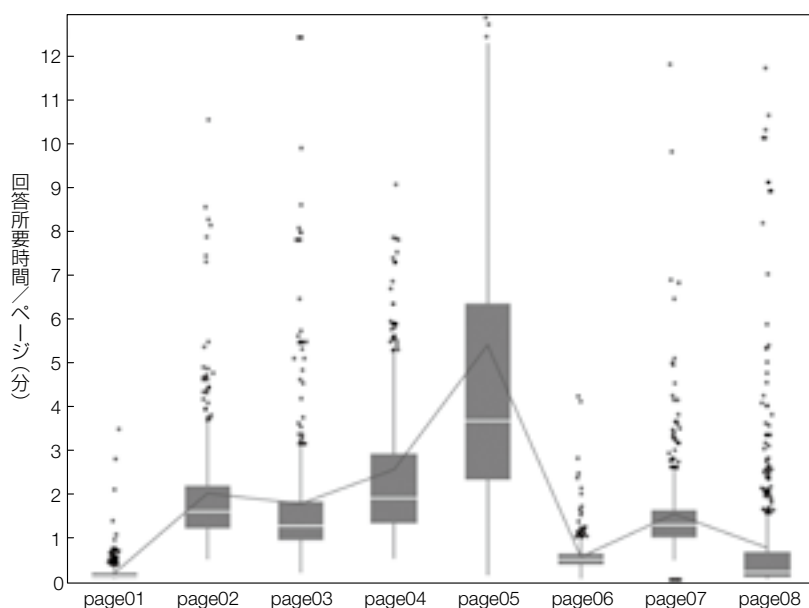


図2〈調査1〉のページ別の回答所要時間の分布

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

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。

さらに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と、反対に異常に長い回答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、つぎ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)。

〈調査2〉についても回答所要時間を調べよう(図3)。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)の数が多いこと、自由回答の書き込みなどで回答所要時間が増えて変動も大きい。ここでは、中央あたりと後半の自由回答、後半の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: ここに「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6つの選択肢からなる質問で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変動も大きい。

●page14~16: ここに人口統計学的項目があるが、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。

回答所要時間の回答推移パターンの観察

視点を変えて、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推移の傾向を「回答推移パターン」として調べた。例として〈調

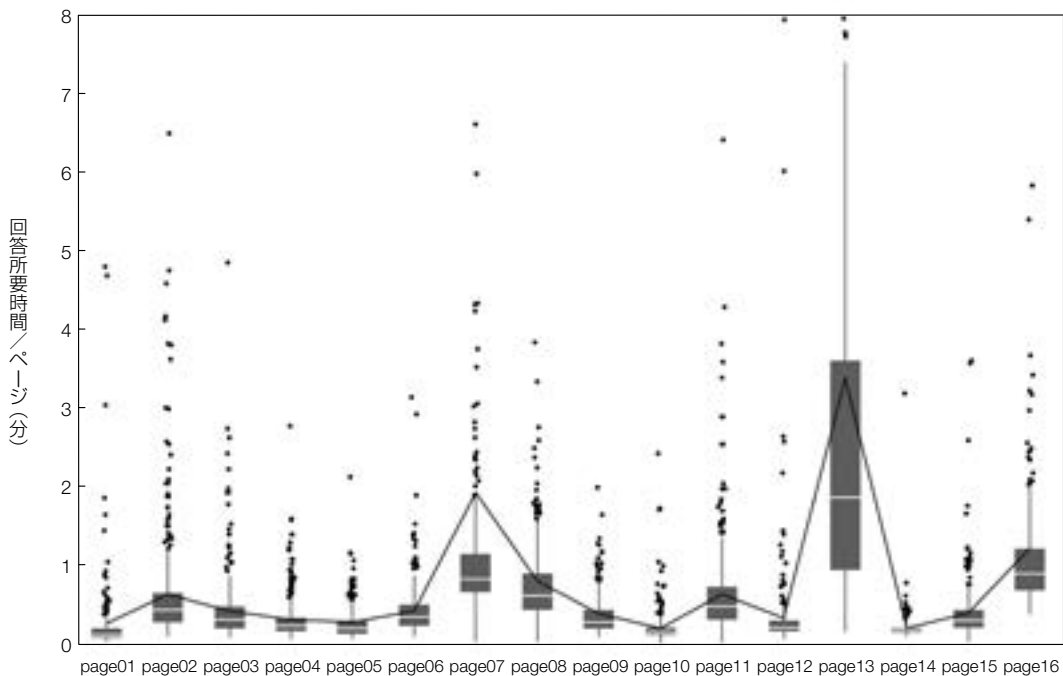


図3 〈調査2〉のページ別の回答所要時間の分布

査1〉の全回答者について、回答者ごとに各ページの回答所要時間の遷移を折線で結んだ回答推移パターンを描き全体の傾向を観察したうえ、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グループに分けてみた。

回答所要時間が短い「10分未満」のグループは、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り、各ページを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は、どこかのページで回答に戸惑うなど、さまざまなパターンが観察された。

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、回答所要時間を複数の層に分け(例: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必要がある(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、つぎのような場面でパラデータの利用が考えられる(こうした分析も多々あるが略す)。

- 回答の手抜き現象(労働最小化行動など)の可能性がある場合、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のためらいがあるか、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、よ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。

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るか、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。回答所要時間の区分化1つを考えても、実はむずかしい操作である。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」(スライダー形式、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そうか、その考え方、利用可能性、具体的な取得方法や適用場面について、例をまじえてごく簡単に紹介した。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スキルと回答内容や回答分布の傾向との関係、そして総調査誤差の低減にどうつながるかなど、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、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”を探すこと。
- 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(意外と面倒なこと)。とくに、諸指標の算出・分析を容易にするツールの開発。

情報通信技術の急速な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、回答環境が大きく変化している。このことで、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 and Couper, 2011；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ある報告。この他、Couper and 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で、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は説明していない。われわれの実験調査では、技術的制約もあるが、収集データの内容を限定し、おもに 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、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（回答率、中断率、無回答率や回答分布）。わずかな証拠しかないが、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がある）。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・欠点をよく知ったうえ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）の各氏には、貴重なお助言、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。



文献

- AAPOR, 2016, *AAPOR: Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys* (Revised 2016).
- Bethlehem, J. and Biffignandi, S., 2012, *Handbook of Web Surveys*, Hoboken: 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.
- 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 Key-stroke files to Assess respondent Difficulties with an Audio-CASI Instrument," *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 239-244.
- Couper, M. P., 2000, "Web surveys: A Review of Issues and Approaches" *Public Opinion Quarterly*, 64: 464-494.
- , 2005, "Technology Trends in Survey Data Collection," *Social Science Computer Review* 23 (4): 486-501.
- , 2008, *Designing Effective Web Surveys*, New York: Cambridge University Press.
- , 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).
- , 2014, *Introduction to Web Survey Paradata*, WebDataNet 2nd Training School: Paradata, Alexandroupolis, Greece, 1-3 October 2014.
- , 2017, 「パラデータ概念の誕生と普及」, 松本 渉訳『社会と調査』18: 14-26.
- Das, M., Ester, P. and Kaczmirek, L. (eds.), 2010, *Social and Behavioral Research and the Internet: Advances in Applied Methods and Research Strategies*. New York: Taylor & 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), Hoboken: John Wiley & Sons.
- Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E. and Tourangeau, R., 2004, 2009, *Survey Methodology*, Hoboken: John Wiley & 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).
- , 2011, The CSP Project Webpage. *Technical Report October 20, 2011*.
- Kreuter, F. and Casas-Cordero, C., 2010, Paradata, Working Paper No. 136, RatSWD, *Working Paper Series of the Council for Social and Economic Data*.
- Nicolaas, G., 2011, Survey Paradata: A Review, *ESRC National Centre for Research Methods Review Paper*, National Centre for Research Methods.
- Schonlau, M. and Toepoel, V., 2015, "Straightlining in Web Survey Panels Over Time," *Survey Research Methods*, 9(2): 125-137.
- Tourangeau, R., Conrad, F. G. and Couper, M. P., 2013, *The Science of Web Surveys*, New York: Oxford University Press. (『ウェブ調査の科学』として、翻訳版が朝倉書店から刊行予定)
- 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*, 4 (4).
- ※欧文献のみを挙げた。なお、本特集の Couper (2017) で引用の文献は除外した。本文もたくさんの紙幅をいただいたが、なお意を尽くせずわかりにくい部分もある。そこで、本稿の元となった原稿を以下のサイトに掲載した。ここでは、回答所要時間のほか、回答完了と中断、回答遷移パターン、回答率などについて述べた。なお、アップロード版には、本稿で引用の欧文・邦文文献および関連 URL のすべてを一覧としたので、ご利用いただければありがたい。
- URL: <http://wordminer.org/wp-content/uploads/2017/01/ParadataInWebSurveys.pdf>