

インターネット調査の諸特性と今後の展開のあり方

－第4次実験調査から見てきたもの－

○吉村宰*・大隅昇**・清水信夫**

*岡山大学教育学部・**統計数理研究所

1. はじめに

我々は、WWW環境を利用したデータ取得法の調査法としての性質、特徴を可能な限り明確にするという目的のもと、1997年から2000年に、計3回のWEB実験調査を行ってきた(Ohsumi and Yoshimura (1999), Yoshimura and Ohsumi (2000), 大隅 (2000,2001), 吉村 (2001))。第1次実験調査では、約6ヶ月計12回に及ぶ調査を実施し、WEB調査で起こりうる現象や検証すべき事項の把握に努めた。第2、第3次実験調査では第1次調査の結果をふまえ、WEB調査法の特性のさらなる明確化を試みた。その際に設定した実験調査の基本方針は次の通りであった。

- ・複数の調査機関の協力を得て複数のWEBサイトではほぼ同時期に複数回実施する
- ・原則として調査票は各機関とも同じ内容とする
- ・調査票へのアクセス方式は各機関の現状方式そのままとする
- ・各機関が保有する調査対象者集団の構築法及び管理・維持・運営方法を把握する
- ・従来型の調査方式との比較検証を行う。その際、調査実施時期、設問は可能な限りWEB調査に合わせる
- ・可能でな範囲で回答変動の検証を行う

表1：第4次実験調査全体の概要

WEB調査	電通リサーチ (DR)		東京サーベイ・リサーチ (TSR)	日本リサーチセンター (NRC)
	第1回：生活意識編			
リソース/パネル名	DENTSU_R-net	Hot Panel	e-HABIT	Cyber Panel
調査方式	リソース内 サンプリング方式	リソース内 サンプリング方式	パネル方式	リソース内 サンプリング方式
調査時期	3/28/2002-4/4/2002	6/13/2002-6/20/2002	3/28/2002-4/8/2002	4/11/2002-4/25/2002
調査票のページ分割	×	△ (半数に実施)	×	○
設問レイアウト実験	○	△ (半数に実施)	○	○
有効回収回答 (%)	939 (61.6)	3,392 (42.4)	931 (88.2)	725 (36.3)
第2回：インターネット編				
調査時期	5/16/2002_5/23/2002	6/20/2002-6/27/2002	5/16/2002-5/27/2002	5/22/2002-6/7/2002
調査票のページ分割	×	×	×	○
有効回収回答 (%)	894 (59.1)	2,587 (32.4)	896 (84.9)	691 (34.6)
オムニバス調査	DRPS	NOS	郵送調査	HABIT2001
	第1回：生活意識編			生活意識+暮らし向き編
調査期間	3/22/2002-3/31/2002	5/7/2002/-5/15/2002	調査期間	7/4/2002-7/25/2002
方法	面接法	訪問留置自記式	方法	郵送
対象地域	東京30km圏	日本全国	対象地域	首都圏40km・近畿20km
対象者	15～59歳男女	15～79歳男女	対象者	HABIT2001の一部 15～69歳男女
サンプリング法	住民基本台帳からの 二段無作為抽出	住民基本台帳からの 層化多段無作為抽出		
計画サンプル数	1,236	2,200	計画サンプル数	1,000
回収率 (%)	630 (51.0)	1,336 (60.7)	回収率 (%)	946 (94.6) *
不能票数 (%)	606 (49.0)	864 (39.3)	不能票数 (%)	54 (5.4) *
第2回：暮らし向き編				
調査期間	5/17/2002-5/26/2002	5/7/2002/-5/15/2002	*2002/07/25現在	
計画サンプル数	1,001	2,200		
回収率 (%)	630 (62.9)	1,389 (63.1)		
不能票数 (%)	371 (37.1)	811 (36.9)		

第4次実験調査では、第2、3次実験調査の基本方針を踏襲しつつ以下を新たな検証事項として加えた。実験調査の全体像は表1に示した通りである（なお表中に示した「有効回収回答」は、回収されたユニークな回答数として定義されるものであり、各調査機関がそれぞれ定める有効回答数とは異なる場合がある。またTSR以外のWEB調査では2回の調査における計画サンプルはそれぞれ独立にリソースから無作為に抽出したものである）。

- 1) 可能な範囲で回答者行動を追跡し関連データを収集することにより、WEB調査の特性をより詳細に把握するとともに、WEB調査の質的側面の評価を試みる
- 2) リソースもしくはパネル（調査対象候補となる登録者集団）の構築方法の差異と種々の調査結果との関連を検証する。特にself-selection型の構築方法と住民台帳からの無作為抽出による構築方法との比較に焦点をあてる
- 3) 調査票のページ分割：調査票全体をいくつかのページに分割する場合としない場合で、回答者行動その他にどのような影響が生じるかの検証を行う。
- 4) 設問レイアウト実験：設問の回答選択肢の形式の効果を検証するためのレイアウト実験を行う。なおこれは海外の研究者グループとの共同実験である。
- 5) 従来型調査として、2つのオムニバス調査及び1つの郵送調査を行う

本稿では、特に上記1) 2) から見えてくることとして、WEB調査の質的側面の評価とその標準的なあり方を中心に報告する。

2. WEB調査の質的側面の評価

従来型の調査では、調査概要として、少なくとも計画サンプルとその抽出方法、有効回収数、不能票数とその内訳等が報告される。こうした情報を通してその調査がどのようなものであるか（調査の質）が評価され、それをふまえた上での調査結果の解釈が行われる。ところが、WEB調査では、こうした調査の質的側面の評価に関する情報として、何をどこまで記述すべきかに関する共通の認識はまだない。

表2、3は、実験調査を通して得た情報に基づき整理された調査の質的側面に関する事項である。WEB調査においても最低限この程度の情報を調査概要として記述する必要があるという共通認識を形成することが急務であると考えられる。

2-1. 無回答の内訳から見るWEB調査の質 現行の多くのWEB調査では無回答についての情報が記述されることはまずない。しかし、調査の質的側面を評価する上で、なぜ回答されなかったかの内訳を明確にし検証することは調査の質

表2-1：電通リサーチDENTSU_R-net調査概要

調査回	第1回調査	第2回調査
調査コード	01000-077,129,159~162	01000-078
調査期間	3/28/2002-4/4/2002	5/16/2002_5/23/2002
調査テーマ	生活意識編	インターネット編
謝礼	回答者全員に500アルネ	回答者全員に500アルネ
調査方式	リソース内サンプリング方式	リソース内サンプリング方式
リソース	DENTSU_R-net	DENTSU_R-net
登録者数	24,053	24,053
計画標本数	1,542	1,542
依頼発信前除外	17	30
調査依頼発信数	1,525	1,512
有効回収回答 (%)	939 (61.6)	894 (59.1)
無回答 (%)	586 (38.4)	618 (40.9)
未達 (%)	63 (4.1)	34 (2.2)
無接触 (%)	420 (27.5)	414 (27.4)
アクセスのみ (%)	97 (6.4)	163 (10.8)
回答送受信異常 (%)	6 (0.4)	7 (0.5)
アクセス数 (延べ数)	1,270	1,403
アクセスのみ (延べ数) (%)	298 (23.5)	498 (35.5)
回答送信 (延べ数) (%)	972 (76.5)	905 (64.5)
回答送受信異常 (延べ数) (%)	27 (2.1)	11 (0.8)
回収回答 (延べ数) (%)	945 (74.4)	894 (63.7)
重複回収回答 (%)	6 (0.5)	0 (0)
非登録者回答	生起しない	生起しない
ID不明回答	0	0
調査依頼未達 (延べ数)	65	36
登録者数確定日時	3/14/2002	3/14/2002
リマインダー	1回 (E-mail, 4/3)	1回 (E-mail, 5/22)
発信数	720	723
未達数	—	—

注1) アクセスのみ：指定URLへアクセスしたが、回答送信ボタンを押していないもの

注2) 回答送信異常：回答送信ボタンを押したが、回答が正常に完了していないもの

注3) 重複回収回答数：延べの回収回答のうち非採用とした重複する回答数

注4) (%)の母数は調査依頼発信数

注5) 未達：調査依頼メールが戻ってくる（住所不明に相当）

注6) 無接触：調査票にアクセスしない（不在・回答拒否に相当）

注7) アクセスのみ：調査票にアクセスするが回答は完了せず（回答意思の有無は特定できない）

注8) 回答送受信異常：回答完了し送信操作を行ったが受信に失敗（回答の意思あり）

注9) (%)の母数は総アクセス数（延べ数）

表2-2：電通リサーチHot Panel調査概要

調査回	第1回調査	第2回調査
調査コード	02000-099~105	02000-107
調査期間	6/13/2002-6/20/2002	6/20/2002-6/27/2002
調査テーマ	生活意識編	インターネット編
謝礼	抽選で1,200名に500円の図書券	抽選で1,200名に500円の図書券
調査方式	リソース内サンプリング方式	リソース内サンプリング方式
リソース	Hot Panel 5 & 6	Hot Panel 5 & 6
登録者数	64,629	64,600
計画標本数	7,996	8,000
発信前除外	3	8
調査依頼発信数	7,993	7,992
有効回収回答 (%)	3,392 (42.4)	2,587 (32.4)
無回答 (%)	4,601 (57.6)	5,405 (67.6)
未達数 (%)	560 (7.0)	577 (7.2)
無接触 (%)	3,216 (40.2)	3,494 (43.7)
アクセスのみ (%)	778 (9.7)	1,302 (16.3)
回答送受信異常 (%)	47 (0.6)	32 (0.4)
アクセス数 (延べ数)	5,126	5,226
アクセスのみ (延べ数) (%)	1628 (31.8)	2581 (49.4)
回答送信 (延べ数) (%)	3498 (68.2)	2645 (50.6)
回答送受信異常 (延べ数) (%)	102 (2.0)	56 (1.1)
回収回答 (延べ数) (%)	3396 (66.3)	2589 (49.5)
重複回収回答 (%)	4 (0.1)	2 (0.0)
非登録者回答数	生起しない	生起しない
ID不明回答数	0	0
調査依頼未達 (延べ数)	781	579
登録者数確定日時	6/7/2002	6/10/2002
リマインダー	1回 (E-mail, 6/19)	1回 (E-mail, 6/26)
発信数	5,270	6,072
未達数	—	—

表3：東京サーベイ・リサーチe-HABIT調査概要

調査回	第1回調査	第2回調査
調査コード	200201	200202
調査期間	3/28/2002-4/8/2002	5/16/2002-5/27/2002
調査テーマ	生活意識編	インターネット編
謝礼	回答者全員に5ポイント	回答者全員に5ポイント
調査方式	パネル方式	パネル方式
パネル	e-HABIT	e-HABIT
登録者数	1,055	1,055
計画標本数	1,055	1,055
依頼発信前除外	0	0
調査依頼発信数	1,055	1,055
有効回収回答 (%)	931 (88.2)	896 (84.9)
無回答 (%)	124 (11.8)	159 (15.1)
未達数 (%)	0	3 (0.3)
無接触 (%)	113 (10.7)	125 (11.8)
アクセスのみ (%)	11 (1.0)	31 (2.9)
回答送受信異常 (%)	—	—
アクセス数 (延べ数)	1,136	1,153
アクセスのみ (延べ数) (%)	149 (13.1)	219 (19.0)
回答送信 (延べ数) (%)	987 (86.9)	934 (81.0)
回答送受信異常 (延べ数) (%)	—	—
回収回答 (延べ数) (%)	987 (86.9)	934 (81.0)
重複回収回答 (%)	56 (4.9)	38 (3.3)
非登録者回答数	生起しない	生起しない
ID不明回答数	7	6
調査依頼未達 (延べ数)	—	—
登録者数確定日時	3/25/2002	3/25/2002
リマインダー	1回 (E-mail, 4/4)	1回 (E-mail, 5/23)
発信数	356	394
未達数	2	4

注1) 未達数：調査依頼メールが戻ってくる（住所不明に相当）。なお、第1回調査では初め2ケースの未達が生じたが電話でアドレスを再確認の後、再発送、2ケースとも回答回収に至っている。また第2回調査では初め6ケースの未達が生じたが、電話確認の後、うち3ケースが回答回収に至っている
注2) 総アクセス数 (延べ数)、アクセスのみ (延べ数)、回答送信数 (延べ数)、回収回答数 (延べ数) にはID不明者による回答は含まれない

と信頼性を高める上で必要不可欠であると考える。

「未達」は発信された調査依頼が発信先に届かずに戻ってきた数である。登録者集団の維持・管理の状態に強く左右される指標と言える。

「無接触」は調査依頼に対する何の反応もなかった数である。従来調査での不在や回答拒否に相当すると言える。表2-1～表3を見比べると、リソース（パネル）ごとにその値は異なっており、各リソース内では2回の調査間で値が類似している。このことから「無接触」は主に登録者集団の質を反映する指標と考えられる。登録者集団構築法やその維持・管理の状態による影響を受けるのだろう。

「アクセスのみ」は調査票ページにアクセスするなど、調査依頼に対し何らかの行動を起こしたものの回答送信にまで至らなかった回答者の数である。表2-1～表3から、リソース間での違いは確かにあるものの、共通して第2回調査でこの値がほぼ倍増している。第2回調査の調査票ボリュームは第1回調査に比べかなり大きいものであった。このことから「アクセスのみ」の値は調査票のボリュームが適正なものであったかを判断する1つの指標となりうると言える。

「回答送受信異常」は、回答者が回答送信の操作を行ったことの記録が残っているにも関わらず回答データが取得されていなかった数である。あえて言うなら郵送調査における郵便事故に相当する。必ずしも調査実施機関だけの問題ではなく、インターネットという仕組みに起因する側面も大いにあるだろうが、WEB調査における技術面での質を評価する1つの手がかりとなることは間違いない。可能な限り情報の取得に努めることが望ましい。

2-2. アクセス数から見る WEB 調査の質 上で述べた無回答の内訳はいずれも回答者数に基準をおいたものである。ここではアクセス件数を基準とした考察を行う。表2-1～表3に示したアクセス数(延べ数)とは、回答者を区別せずに全体でどれだけの調査票への接触があったかを表している。

どのリソースにおいても、第2回調査でのアクセス数が第1回に比べて多くなっていることが分かる。第1回と第2回調査の大きな違いは調査票のボリュームである。調査票のボリュームが大きくなるとアクセス数は増える。また、その内訳としてアクセスのみの増加、回答送信数及び回収回答数の減少を共通する特徴として指摘できる。このことから、回答者は1度調査票を見てボリュームの大きさに直ちに回答を諦めるのではなく、何度も挑戦するが回答送信には至らない、という様子が読みとれる。

表4は回答者のアクセス回数である。ここから分かるように1度のアクセスで回答を完了している回答者は回答者全体の85%前後である。残りは1度のアクセスで回答を完了していない。もちろん回答者側の都合等によるものも多く含まれるだろうが、何らかのトラブルが生じている可能性もある。こうした情報も調査の質的側面の評価のために利用したい。

表4：回答者のアクセス回数

調査回	第1回：生活意識編			第2回：インターネット編			
	リソース	DENTSU_R-net	HotPanel	e-HABIT	DENTSU_R-net	HotPanel	e-HABIT
アクセス回数	N=939	N=3,392	N=931	N=894	N=2,587	N=896	
1	83.9	86.2	83.1	79.3	78.8	81.7	
2	12.4	10.7	14.2	15.8	16.0	14.7	
3	3.2	2.1	2.1	3.5	3.7	2.6	
4	0.2	0.5	0.4	0.8	0.9	0.7	
5	0.1	0.2	0.0	0.2	0.3	0.2	
6	0.1	0.1	0.0	0.2	0.2	—	
7	0.1	0.1	0.0	—	—	—	
8	—	—	0.0	0.1	0.1	—	
9	—	—	0.1	—	—	—	
13	—	—	—	0.1	—	—	
15	—	—	—	—	—	—	
16	—	—	—	—	0.0	—	
20	—	0.0	—	—	—	—	
22	—	—	—	—	—	0.1	
1回で回答完了	83.9	86.2	83.1	79.3	78.8	81.7	
2回までの累計	96.3	96.9	97.3	95.1	94.7	96.4	

4. おわりに

WEB 調査を有効に利用・活用するためには、その調査の質がどうであるかの評価が不可欠である。本稿では、回答者行動を追跡するデータの取得しそれを整理することによって、リソースの維持・管理状態、調査票ボリュームの適切さ、回答取得システムの適切さ等の調査の質的側面の評価のための指標を得ることができることを示した。そしてこれらの情報は調査結果の解釈のためにも欠かせない情報である。さらに、ここでは触れなかったが、リソースがどのような手続きで構築されたかも調査の結果を解釈する上での重要な情報である。併せて示し参考とすることが望ましい。

現在も WEB 上で様々な「調査」が行われているが、どの調査でも今回ここで示したような情報を調査の「顔」として必ず示すという共通の土壌ができれば、他調査との比較もこれまで以上に意味のあるものとなるだろう。WEB 調査の今後の展開を考える上で必要不可欠な事項であると考え。

引用文献

Ohsumi, N. and Yoshimura, O. (1999) : The online survey in Japan: An evaluation of emerging methodologies, Bulletin of the International Statistical Institute 52nd Session, Book2, 171-174.

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

大隅 昇 (2001) : 電子調査, その周辺の話題—電子的数据取得法の現状と問題点—, 統計数理, 49(1), 201-213.

Yoshimura, O. and Ohsumi, N. (2000) : Some experimental surveys on the WWW environments in Japan, Data Analysis, Classification, and Related Methods (eds. H. Kiers, J. Rasson, P. Groenen and M. Schader), 353-358, Springer, Berlin.

吉村 宰 (2001) : インターネット調査にみられる回答者像, その特性, 統計数理, 49(1), 223-229.