

電話調査における調査所要時間と 回答傾向について

土屋 隆裕[†]

(受付 2004年9月28日; 改訂 2004年12月21日)

要 旨

本稿の目的は、電話調査における調査所要時間と回答傾向との関係を調べ、データ分析に際しての調査所要時間の活用方法を検討することである。そこで、電話調査の回収サンプルを調査所要時間に基づいて2群に分類し、短時間群と長時間群の間で調査への回答分布を比較した。まず、ロジスティック回帰分析の結果、高齢者ほど調査所要時間は長くなることを示した。回答者が調査に協力した理由が、調査所要時間の長短に関係しているのではないかと考えられたが、実際には関係は認められなかった。次に、短時間群に比べ、長時間群の方が、その他・わからない(D.K.)の回答比率が大きくなることを示した。さらに、split-ballot デザインを用いた調査結果から、選択肢の数や表現、提示順序の違いが回答に与える影響について述べた。特に選択肢の数と提示順序の違いについては、長時間群に比べ、短時間群の方が、その影響を大きく受ける傾向にあった。最後にまとめとして、これらの結果に基づき、調査データ分析のために、調査所要時間を活用する方法について論じた。

キーワード：電話調査法，日本人の国民性調査，調査所要時間，反応潜時，調査協力理由。

1. はじめに

1.1 本稿の目的

本稿の目的は、電話調査において、調査所要時間が異なる回答者の間では、回答にどのような異なる特徴が見られるのかを明らかにし、データ分析に際しての調査所要時間の活用方法を検討することである。なお、ここで言う調査所要時間とは、最初の質問項目の質問文が読み上げられ始めてから、最後の質問項目の回答が得られるまでの、回答者ごとの所要時間を指す。したがって、質問に入るまでの、回答者に対する調査の趣旨説明や協力依頼などの時間は含まれない。

調査所要時間の特性を検討することには、次のような意義がある。調査データの分析にあたっては、人口統計学的属性など枠の情報はもちろん、調査協力の同意が得られるまでの過程や調査実施の状況など、質問項目への回答以外の補助情報を用いることがある。例えば、面接調査や電話調査では、全回答者を、一度は協力を拒否した回答者やなかなか接触できなかった回答者と、容易に回収できた回答者とに分類することがある。そして、前者の、回収が困難であった回答者を、調査不能者の意見分布の推定に用いようというのである(例えば、Lin and Schaeffer

[†] 統計数理研究所：〒106-8569 東京都港区南麻布 4-6-7

(1995); Lynn et al. (2002) など)。また、Web 調査では、調査所要時間が極端に短い回答者は、インセンティブのみを目的としている可能性があるとして、集計から除外することもある。

このように、調査実施時の補助情報をデータ分析のために活用したり、補助情報を用いてデータの質を高めたりするためには、その補助情報がどのような特性を持ち、回答者のどのような特徴を反映しているのか調べておく必要がある。林(2004)は、郵送調査における調査票の早期返送者と後期返送者の特性を比較した研究をレビューし、“調査内容に対する調査対象者の関心の程度や親近性”が返送時期を左右しているのではないかと述べている(p.91)。また土屋・前田(2003)は、調査方法が異なる二つの電話調査の回答分布を用いて、各調査方法の回収サンプルを、もう一方の調査方法でも回収できるであろう共通クラスと、その調査方法でのみ回収できるであろう独自クラスとに分類することを試みている。仮に、両クラスの判別に有効な補助情報を見つけることができれば、単独の調査においても、回答者を両クラスに分類することができるであろう。

そこで本稿では、調査実施時の補助情報の一つとして、電話調査における調査所要時間を取り上げる。現在では、多くの電話調査が、コンピュータを利用した CATI (Computer Assisted Telephone Interviewing) であると考えられる。CATI では、調査員(オペレータ)がコンピュータ画面に表示された質問文を読み上げ、回答をコンピュータに直接入力する。そのため、あらかじめ CATI システム内にプログラムさえしておけば、回答者やオペレータに負担をかけることなく、コンピュータによる調査所要時間の測定が可能である。さらに、次の 1.2.1 節で述べる反応潜時ほどの精密な測定(具体的な方法は Bassili (1996a) などを参照)も必要としないため、多くの電話調査で容易に利用可能な補助情報と言える。

1.2 従来の研究結果と予想される結果

1.2.1 反応潜時

繰り返しになるが、本稿で用いる調査所要時間とは、最初の質問が始まってから最後の質問が終わるまでの、調査全体にかかった時間である。これに対し、Bassili らの一連の研究では、個々の質問項目の反応潜時を用いている。反応潜時とは、回答者が質問文を聞き終わってから回答を発するまでの所要時間のことである。回答者が回答を発するまでには、質問内容を理解し、情報を心的に引き出し、判断し、回答を選択する、といったプロセス(Tourangeau et al. (2000))をたどっており、反応潜時の長さは、各プロセスに要した時間の影響を受けると考えられる。例えば、Bassili and Scott (1996) は、二重否定を含む質問、double-barreled 質問、定義があいまいな単語を含む質問といった、内容理解のプロセスに時間を要すると考えられる不適切な質問項目と、それらの問題点を適切に修正した質問項目との間で反応潜時を比較した。そして、二重否定を含む質問と double-barreled 質問の反応潜時は、修正した質問項目のそれよりも有意に長いことを示している。また、Bassili (1996a) は、わからない(Don't Know; D.K.) という回答の反応潜時は、提示された選択肢の中から回答が選ばれた場合の反応潜時の 2~3 倍長い、と述べている。これは、判断あるいは回答選択のプロセスに時間を要したためであろう。

反応潜時を用いた調査研究では、反応潜時の短さを、回答者の態度の強さ・安定性を示す指標と考えることが多い(Tyebjee (1979); Bassili (1996a))。例えば、Bassili and Fletcher (1991) は、CATI を用いて、「大企業では、従業員に占める女性の割合を定めておくべきか、それとも女性を特別扱いすべきではないか」という質問への反応潜時を調べた。続けて、回答者が選んだ方の回答の難点を説明し、それでも回答は変わらないかを質問したところ、回答が変化した回答者の反応潜時に比べて、変化しなかった回答者のそれは有意に短い、という結果を示している。また、アメリカ大統領選挙などの選挙前に行われた実験や調査では、候補者に対する態度項目への反応潜時が短いほど、その回答と実際の投票行動が一致していることが繰り返し見

出されている(Fazio and Williams (1986); Bassili (1995); Stein and Johnson (2001)). さらに Bassili (1993) は、投票意向と実際の投票行動とが一致するかどうかを判別するための説明変数としては、投票意向についての回答が今後変わる可能性があるか、という質問への回答者自身の回答(Bassili (1996b) はこれを meta-judgmental 指標と呼んでいる)よりも、投票意向に関する質問の反応潜時(Bassili (1996b) はこれを operative 指標と呼んでいる)の方が有効であるとしている。ただし、Bassili and Krosnick (2000) は、意見の確実さ・強さ・重要さ・変化可能性・極端さ(これらの具体的な内容は、Krosnick and Schuman (1988); Krosnick et al. (1993) などを参照)といった meta-judgmental 指標と、operative 指標である反応潜時を用いて、質問の順序効果や中間的な選択肢の提示効果の大きさなどを説明しようとしているが、反応時間が短いほど効果が小さいという結果は得られていない。

1.2.2 予想される結果

以上の結果などを踏まえ、本稿では、主に以下の三つの観点から、調査所要時間との関係を調べることとする。

観点1：調査所要時間と調査への協力理由(3.1 節)

最終的には調査への協力に同意した場合であっても、調査に対する回答者の態度は、どの回答者も同様というわけではない。はじめから調査内容に興味を持ち、積極的に調査に参加する回答者もいれば、オペレータに繰り返し懇願され、消極的な態度で調査に応じる回答者もいる。もし、消極的な回答者ほど、調査を早く切り上げたがるとすれば、以下の傾向が見出されると考えられる。

予想される結果1：調査に対して積極的な回答者ほど調査所要時間は長くなり、消極的な回答者ほど調査所要時間は短くなる。

観点2：調査所要時間と‘その他・D.K.’の回答比率(3.2 節)

本稿で分析する電話調査では、提示した選択肢の中から回答者が回答を選べない場合には、‘その他・D.K.’として処理することとしている。調査を早くに終わらせたいと思っている回答者が、輕易に「わからない」と答えたり、気を利かせすぎってしまったオペレータが、早々に「わからない」へ誘導してしまえば、以下の傾向が見出されると考えられる。

予想される結果2a：調査所要時間が短い回答者ほど；‘その他・D.K.’の回答比率が高くなる。

その一方で、Bassili (1996a) の結果では、‘D.K.’の反応潜時はそれ以外の選択肢の反応潜時に比べて長いことが示されている。長考の末、回答者が結局どの選択肢も選べず；‘その他・D.K.’となるのであれば、上記2aとは逆に、以下の傾向が見出されると考えられる。

予想される結果2b：調査所要時間が長い回答者ほど；‘その他・D.K.’の回答比率が高くなる。

観点3：調査所要時間と選択肢の提示効果(4.2 節)

調査における回答が、質問文や選択肢の表現・順序などの影響を受けることがあることは、古くからよく知られている(Payne (1951))。例えば西平(1956)は、四つの面接調査の間で、回答方法の違いや、質問文・選択肢のワーディングの違いなどが、回答分布に与える影響を調べている。また、鈴木(1968)や鈴木 他(1981)は、面接調査においてパネルなどを利用し、選択肢の数や表現を変えることや、中間的な選択肢を用いることの影響などを調べている。さらに、坂元 他(2004)と土屋 他(2004)にも、選択肢の表現あるいは順序を変えて実施した調査の結果がまとめられている。Bassili らの一連の研究結果に基づけば、反応潜時の短さは、回答者の意見や態度の強さ・安定性を示す。そのため、以下の傾向が見出されると考えられる。

予想される結果3a：調査所要時間が短い回答者ほど、選択肢の表現や選択肢の提示順序などの影響を受けにくい。

その一方で、鈴木(2003)が指摘するように、“電話調査では調査時間の制約が強”く、回答者に“早く回答しなければ”という心理的圧力が生じる可能性がある(p. 84)。そのような心理的圧力が原因となるなどして、選ぶ選択肢の適切さを時間をかけて吟味しなかった回答者ほど、選択肢の提示方法の影響を受けやすいとすれば、上記 3a とは逆に、以下の傾向が見出されると考えられる。

予想される結果 3b：調査所要時間が短い回答者ほど、選択肢の表現や選択肢の提示順序などの影響を受けやすい。

以下、第 2 章では、分析に用いる電話調査データと調査所要時間の概要を述べる。第 3 章では、調査所要時間が短い群と長い群との間で、調査への協力理由や‘その他・D.K.’を含む回答分布を比較する。第 4 章では、調査所要時間と選択肢の提示効果との関係について示し、第 5 章では、本稿における分析で得られた結果をまとめる。

2. 調査の概要

2.1 電話調査の方法

本稿では、2002 年度と 2003 年度に実施した二つの電話調査データを取り上げる^{注1)}。この二つの電話調査の質問項目や、性別・年齢層別集計結果は、土屋 他(2004)にまとめられている^{注2)}。調査の概要は、表 1 に示すとおりである。なお、性別や年齢など、2002 年度調査では 7 問、2003 年度調査では 8 問の回答者属性に関する調査項目があるが、これらは表中の質問項目数には含まれていない。また共通項目とは、後述する 2 種類の調査票の間で、質問文や選択肢が共通の項目である。

いずれの調査も、Random Digit Dialing (RDD) による電話聴取法 (CATI) を用いて実施した。具体的には、前田・土屋(2001)で用いた電話調査法に準ずる以下のような方法である。すなわち、全国で使用されている可能性がある固定電話番号を単純無作為抽出法により抽出し、あらかじめ機械的なスクリーニングを行った。各調査実施機関は、調査期間中に、用意された電話番号に対して発信を行い、その電話番号が、調査対象が居住する世帯用番号であることが判明した場合には、調査主体名(統計数理研究所) 調査実施機関名・オペレータ名・調査趣旨を述べて、調査への協力を依頼した。協力への同意が得られた場合には、世帯に居住する調査対象の人数を尋ね、その回答と電話番号下 2 桁から成る乱数表に基づき、世帯内での回答者を指定した。オペレータは、コンピュータ画面上に表示される質問文と選択肢を読み上げ、回答者の回答をコンピュータに直接入力した。回答者が質問文や選択肢を聞き漏らしたり、内容を一度

表 1. 調査の概要。

	2002 年度調査	2003 年度調査
調査対象	: 全国の 20 歳以上日本人	全国の 20 歳以上 80 歳未満日本人
調査実施機関	: 株式会社 アダムスコミュニケーション	株式会社 もしもしホットライン
調査日	: 2003 年 1 月 10 日 (金) ~ 23 日 (木)	2003 年 9 月 26 日 (金) ~ 10 月 9 日 (木)
調査時刻	: 10 時 ~ 21 時 30 分	9 時 ~ 22 時
質問項目数	: 14 問 (うち共通項目 4 問)	12 問 (うち共通項目 2 問)
オペレータ数	: 66 名	87 名
回収数 (回収率)	: C 型調査: 747 (32.9%) D 型調査: 715 (31.6%)	K 型調査: 1,191 (28.3%) M 型調査: 1,189 (28.4%)

注) 回収率は、世帯用であると判明した電話番号数に対する回収数の割合

で理解できなかった場合には、省略せずに、再度同じ質問文と選択肢を読み上げた。回答者が、あらかじめ用意した選択肢の中から回答を選べなかった場合には、'その他・D.K.'を選んだものとした。コンピュータは、最初の質問が画面上に表示された時刻と最後の質問に対する回答が入力された時刻を記録した。発信に対して無応答であった場合や、指定した回答者が不在の場合には、日時を改めて再発信を行った。

2002年度・2003年度調査とも、それぞれ2種類の調査票を用意した。2002年度調査では、2種類の調査票(これらをC型調査票とD型調査票と表記する)の間で、質問文は同一であるが、選択肢の数あるいは副詞表現、並び順が異なる質問項目を10問用いた。他にC型調査票とD型調査票で共通の質問項目を4問用いた。2003年度調査では、2003年秋に個別面接聴取法で実施した「日本人の国民性第11次全国調査(坂元 他(2004))の中から主に質問項目を選び出し、2種類の調査票(これらをK型調査票とM型調査票と表記する)の間で、基本的に内容を違えた質問項目を10問ずつ用いた。他にK型調査票とM型調査票で共通の質問項目を2問用いた。2003年度調査では、共通項目のうちの1問は、最後に、今回調査にご協力いただいた主な理由はどういったことでしょうか'という質問文と、表4に示す選択肢を用いて調査への協力理由を尋ねるものであり、調査の一番最後に質問した。なお、2002年度調査では、調査協力理由は質問していない。

用意した電話番号には、いずれかの調査票をあらかじめ無作為に割り当てた。また、各オペレータは、調査時間帯ごとにいずれかの調査票だけを担当するようにした。

2.2 調査所要時間の概要

当然、調査所要時間に影響する要因は、回答者の性格や調査に対する態度などといった、回答者に帰属させられるものばかりとは限らない。オペレータが話すスピードやコンピュータ操作能力、回答者に接する態度などオペレータ側の要因や、CATIシステムのトラブル・障害など、調査実施上の要因が考えられる。そのため、調査所要時間を分析対象とする方法としては、なまの時間をそのまま一つの変数として扱い、精緻なモデルを考えるよりも、調査票ごとに回答者を調査所要時間が短い群(短時間群と略記)と長い群(長時間群と略記)とに分割し、両群の間の比較を行う、というロバスタな方法の方が適当であろう。

そこで、まず、調査中に一時中断があった回答者(D型調査で1名、K型調査で6名、M型調査で6名)を除き、回答者ごとに、最後の質問の回答が入力された時刻から最初の質問がコンピュータ画面上に表示された時刻を引いて、全質問項目に対する調査所要時間を求めた。調査所要時間の分布は表2のとおりである。ただし、2002年度調査では、分単位での記録しか残されていない。短時間群と長時間群がなるべく同サイズとなるよう考慮し、短時間群への分類

表 2. 調査所要時間の分布。

調査 所要時間	C型調査		D型調査		調査 所要時間	K型調査		M型調査	
	頻度(累積%)	頻度(累積%)	頻度(累積%)	頻度(累積%)		頻度(累積%)	頻度(累積%)	頻度(累積%)	頻度(累積%)
3分以下	22 (2.9)	21 (2.9)	~269秒	143 (12.1)	189 (16.0)				
4分	280 (40.4)	235 (35.9)	270~299秒	283 (35.9)	172 (30.5)				
5分	301 (80.7)	286 (75.9)	300~329秒	270 (58.7)	393 (63.7)				
6分	102 (94.4)	127 (93.7)	330~359秒	180 (73.9)	70 (69.7)				
7分以上	42 (100.0)	45 (100.0)	360秒~	309 (100.0)	359 (100.0)				
短時間群	302 (40.4)	256 (35.9)	短時間群	609 (51.4)	679 (57.4)				
長時間群	445 (100.0)	458 (100.0)	長時間群	576 (100.0)	504 (100.0)				
合計	747 (100.0)	714 (100.0)	合計	1,185 (100.0)	1,183 (100.0)				

基準は、C 型調査と D 型調査では 4 分以下、K 型調査では 320 秒未満、M 型調査では 310 秒未満とした。この結果、各群に分類された回答者数を表 2 に示す。

3. 調査所要時間と回答分布

3.1 調査所要時間と回答者属性

この小節では、まず、短時間群と長時間群との間で、人口統計学的属性を含む回答者属性の違いについて検討する。そのため、調査所要時間に基づく 2 群を基準変数、表 3 に示す 8 変数 (2003 年度調査では「協力理由」を加えた 9 変数) を説明変数としたロジスティック回帰分析を行った。なお、いずれの調査においても、各回答者の抽出確率を考慮し、

$$(3.1) \quad w^{(1)} = \frac{\text{その世帯に居住する調査対象人数}}{\text{その世帯が保有する固定電話番号数}}$$

を集計用の重みとして用いた。また、計算には SUDAAN9 (Research Triangle Institute (2004)) を用いた。表 3 には、Wald F 検定の結果 5% 未満で有意であった説明変数について各選択肢の係数も示してある。なお、ランク落ちを避けるため各説明変数の最後の選択肢の係数は 0 としてあり、係数が正に大きいほど短時間群に属する確率が高いことを意味する。

表 3. 短時間群・長時間群のロジスティック回帰分析結果。

説明変数	自由度	C 型調査		D 型調査		K 型調査		M 型調査	
		F 値	係数	F 値	係数	F 値	係数	F 値	係数
性別	1	0.44		0.13		0.18		0.44	
年齢層	5	5.79***		3.13***		2.55**		3.96***	
20 歳代			1.2		1.1		0.8		1.2
30 歳代			1.7		1.2		0.9		1.0
40 歳代			1.5		1.0		0.7		0.8
50 歳代			1.4		0.9		0.9		0.9
60 歳代			0.5		0.2		0.6		0.5
区市郡	2	0.46		2.18		0.36		0.05	
住居形態	1	0.02		5.62		1.68		1.31	
就業状態	2	1.06		0.63		0.19		0.41	
コール数	2	1.08		3.09**		3.73**		9.43***	
1 回					0.6		-0.4		-0.6
2 回					0.0		0.1		0.0
世帯人員数	2	4.60**		0.83		0.18		1.06	
1 人			0.8						
2 人			0.5						
時間帯	2	0.97		0.01		2.05		0.32	
協力理由	2	—		—		1.20		0.34	
R^2		.090		.069		.037		.051	

*** $p < .01$ ** $p < .05$ * $p < .10$

注) 区市郡は、区部・市部・郡部の 3 区分である。

住居形態は、一戸建・一戸建以外の 2 区分である。

就業状態は、常勤で働いている・非常勤で働いている・働いていないの 3 区分である。

コール数は、その電話番号に対する発信回数であり、1 回・2 回・3 回以上の 3 区分である。

世帯人員数は、その世帯に居住する調査対象人数であり、1 人・2 人・3 人以上の 3 区分である。

時間帯は、質問を開始した時刻であり、12 時台まで・17 時台まで・18 時以降の 3 区分である。

協力理由は、表 4 の回答を 選択肢 1 と 2・選択肢 3・選択肢 4 から 7 の 3 区分にまとめたものである。

表 4. 短時間群と長時間群の調査への協力理由(% ; K 型・M 型).

選 択 肢	短時間群	長時間群
1 役に立ちたいと思ったから	15.6 (1.1)	16.2 (1.3)
2 内容がおもしろそうだったから	3.2 (0.5)	2.8 (0.6)
3 文部科学省の研究所が行っている調査だから	20.2 (1.2)	23.0 (1.4)
4 特に断る理由もなかったから	50.4 (1.5)	45.3 (1.7)
5 繰り返し協力を依頼されたから	9.8 (0.9)	9.4 (1.0)
6 その他	0.7 (0.3)	2.7 (0.5)
7 D.K.	0.1 (0.1)	0.5 (0.2)

四つの調査のいずれにおいても、年齢層は有意となっており、若年層ほど短時間群に属する確率が高いことが分かる。世帯人員数は C 型調査でのみ有意であり、逆にコール数は C 型調査を除く三つの調査で有意となっている。ただし、コール数の係数の符号の向きは D 型調査と K 型・M 型調査とは異なり、1 回目の(初めての)発信で回答が得られた回答者は、D 型調査では調査所要時間が短いのにに対し、K 型と M 型調査では長くなっている。年齢層と調査所要時間との関係については、十分納得できる結果であるものの、コール数や世帯人員数については、その理由は判然としない。四つの調査の間で結果が異なることもあり、コール数や世帯人員数と調査所要時間との関係については、他の調査データによる検証がさらに必要であろう。

性別、区市郡、住居形態、就業状態、時間帯、協力理由は、いずれの調査においても有意ではなく、調査所要時間とは関係が認められなかった。特に、1.2.2 節の観点 1 で「予想される結果 1」として述べた、調査に対して積極的な回答者ほど調査所要時間が長い、という関係も認められなかったことになる。ちなみに、表 4 は、K 型調査と M 型調査を合わせて、短時間群と長時間群の回答分布(%)とその標準誤差の推定値(カッコ内)を示したものである。長時間群では、短時間群に比べて、「文部科学省の研究所が行っている調査だから」という回答が 2.8 ポイント多く($\chi^2 = 2.34, p = .13$)、逆に短時間群では、長時間群に比べて、「特に断る理由もなかったから」という受動的な回答が 5.1 ポイント多い($\chi^2 = 4.92, p = .03$)。短時間群の方が消極的であるという可能性は否定はできないが、表 4 の選択肢を用いたのでは、2 群の間の実質的な違いはとらえられないと言ってよいであろう。

3.2 その他・D.K.

3.1 節では、短時間群と長時間群との間で、回答者の年齢層の分布に違いがあることが示された。そこで本来であれば、年齢層ごとに分析を行うのが理想的であろう。しかし、年齢層でサンプルを分割するとそのサイズが小さくなり、傾向が見えにくくなってしまう。そのため、以降の分析では、各調査所要時間群ごとに年齢層で事後層化し、年齢層の分布を群の間で一致させた集計結果を用いることとする。すなわち、ある年度の調査回答者全体の年齢層(6 区分)の分布を P_1, \dots, P_6 、ある調査所要時間群の年齢層の分布を p_1, \dots, p_6 とすると、年齢層のある区分 h に属する回答者の重みを

$$(3.2) \quad w^{(2)} = w^{(1)} \times P_h/p_h$$

とする。

図 1 は、短時間群と長時間群との間で「その他・D.K.」のパーセントを比較したものである。プロットされた点は質問項目を表し、四つの調査票の質問項目全てを図 1 に示している。対角線よりも下に位置し、長時間群に比べ短時間群の方が「その他・D.K.」のパーセントが大きい項目がないわけではない。しかし、明らかに、対角線よりも上にプロットされる点が多く、短時

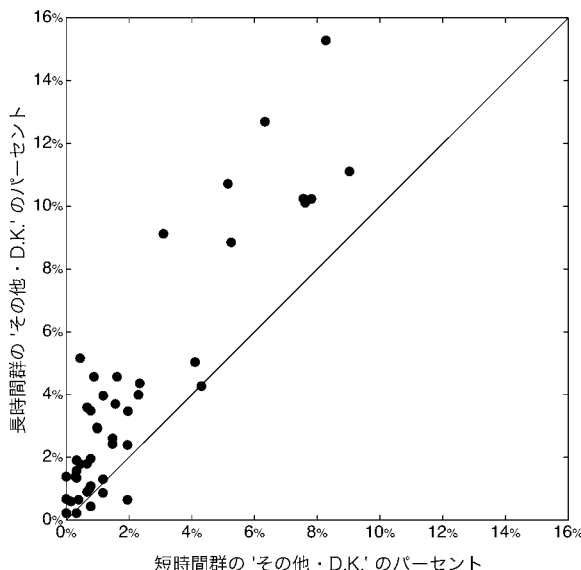


図 1. 短時間群と長時間群の「その他・D.K.」のパーセント比較(C型・D型・K型・M型).

間群よりも長時間群の「その他・D.K.」のパーセントの方が大きい傾向が認められる。つまり、1.2.2 節の観点 2 については、「予想される結果 2a」ではなく「予想される結果 2b」が得られたと見てよいであろう。

3.3 回答分布

1.2.2 節で述べた、本稿で検討しようとする観点からは外れるが、参考のため、K 型調査票と M 型調査票の項目の中で、短時間群と長時間群との間でほぼ 5 ポイント以上の差が見られた選択肢を含む質問項目の結果(カッコ内は標準誤差)を表 5 に示す。2002 年度調査(C 型調査と D 型調査)の結果は、次章で検討するため、ここでは触れない。また、3.2 節では、短時間群と長時間群とでは「その他・D.K.」の回答比率が大きく異なることを示した。そこで 2 群間の比較のため、「その他・D.K.」は除き、それ以外の選択肢の合計が 100% となるよう再調整を行っている。さらに、2 群の間で、「その他・D.K.」を除いた選択肢の回答分布の Wald F 検定の結果も示した。

長時間群に比べ、短時間群の方が回答比率が大きい選択肢は、「家庭に満足(#2.3c)」、「自分の好きなことかどうかはともかく、人のためになることをしたい(#2.11)」、「他人の役にたとうとしている(#2.12)」、「その人の現在の能力を重視して決めるべきだ(#5.23)」、「気の合った人たちと働くこと(#7.24)」、「日本人は西洋人とくらべてすぐれている(#9.6)」、「外国で暮らしてみたいと思わない(#9.19*)」などである。

4. 調査所要時間と選択肢の提示効果

4.1 電話調査における選択肢の提示効果

2002 年度調査では、C 型調査と D 型調査との間で、選択肢の提示方法が異なる質問項目をいくつか用いている。この 4.1 節では、ひとまず調査所要時間の検討から離れ、そのような選択肢の提示方法が回答に与える影響を見てみることにする。なお、この 4.1 節では、集計用の

表 5. 短時間群と長時間群の回答分布(‘その他・D.K.’を除く%)。

項目	短時間群	長時間群	項目	短時間群	長時間群
#2.3c 家庭に満足か(M型)*			#9.6 日本人・西洋人の優劣(K型)***		
満足	47.1 (2.1)	39.5 (2.5)	すぐれている	56.0 (2.3)	41.9 (2.4)
やや満足	40.9 (2.1)	44.1 (2.6)	劣っている	12.3 (1.4)	12.2 (1.6)
やや不満	9.5 (1.3)	13.2 (1.9)	同じだ	13.8 (1.6)	18.8 (1.9)
不満	2.5 (0.7)	3.2 (0.8)	ひとくちではいえない	17.9 (1.7)	27.1 (2.2)
#2.11 好きなくらし方か人のためか(M型)			#9.19* 外国暮らし(K型+M型)***		
自分の好きなこと	38.3 (2.1)	43.0 (2.7)	思う	30.2 (1.4)	36.7 (1.7)
人のためになること	61.7 (2.1)	57.0 (2.7)	思わない	69.8 (1.4)	63.3 (1.7)
#2.12 他人のためか自分のためか(K型)**			#9.20* 日本人の考え方(K型)		
他人の役に	44.4 (2.3)	37.1 (2.4)	そう思う	25.2 (2.0)	22.2 (2.0)
自分のことだけ	55.6 (2.3)	62.9 (2.4)	ややそう思う	28.4 (2.0)	35.5 (2.3)
#5.23 能力か功労か(M型)***			あまり思わない	25.0 (2.0)	23.5 (2.0)
能力重視	72.7 (2.0)	64.1 (2.6)	そう思わない	21.4 (1.8)	18.8 (1.8)
功労重視	27.3 (2.0)	35.9 (2.6)			
#7.24 就職の第1の条件(K型)					
よい給料	3.1 (0.7)	3.3 (0.8)			
失業の恐れがない	11.9 (1.5)	15.9 (1.8)			
気の合った人たち	32.7 (2.1)	26.2 (2.1)			
やりとげたという感じ	52.2 (2.2)	54.7 (2.4)			

*** $p < .01$ ** $p < .05$ * $p < .10$

重みとして 3.1 節(3.1)式による $w^{(1)}$ を用いている。

4.1.1 選択肢の数の違い

表 6 に示す問 2 と問 4 は、2 種類の調査票の間で、選択肢の数が異なっている。例えば問 2 は、D 型調査票では‘その他・D.K.’を除き四つの選択肢を用いて賛否を尋ねているのに対し、C 型調査票では二つの選択肢を用いて尋ねている。

選択肢の語尾に着目して、対応する選択肢の回答比率を調査票間で比較すると、選択肢数が二つの場合の選択肢 1 のパーセントに比べ、選択肢数が四つの場合の選択肢 1 と選択肢 2 を合わせたパーセントの方が、問 2 では 12 ポイント($\chi_1^2 = 17.4, p < .01$)、問 4 では 14.3 ポイント($\chi_1^2 = 26.4, p < .01$)大きい。また、選択肢数 2 の‘その他・D.K.’のパーセントは、選択肢数 4 のそれよりも、問 2 では 9.9 ポイント($\chi_1^2 = 40.1, p < .01$)、問 4 では 1.5 ポイント($\chi_1^2 = 2.85, p = .09$)大きい。

調査票間で‘その他・D.K.’に差が生じた原因としては、選択肢が二つの場合には、中間的なあるいは場合によるといった意見を表すには選択肢の数が少なすぎるため、選択肢が四つの場合に比べ‘その他・D.K.’に逃れる回答者が多くなったことが考えられる。また、選択肢が四つの場合に、はじめの二つの選択肢を合わせた回答比率が高くなった理由の一つとしては、選択肢の数が多くなることで、はじめの方に提示された選択肢ほど選ばれやすくなる初頭効果 (Schuman and Presser (1981); Sudman et al. (1996)) がより強く現れた、ということが挙げられる。さらに、例えば問 4 の場合、D 型調査では二つの選択肢を改めて読み上げていないのに対し、C 型調査では四つの選択肢を全て読み上げている。回答者にとっては、C 型調査の方が、自身の経験を振り返る時間的余裕があり、その分、「ついていけない、と感じること」をより多く思い出したということもあり得る。

選択肢の数が少なくなると‘D.K.’の比率が高くなる現象は、面接調査においても見られる。

表 6. 選択肢の数の違い(%).

問 2 「日本人のものの考え方は外国人には理解できない」という意見があります。あなたはそう思いますか、それともそう思いませんか。[以下、D 型調査票では選択肢を読み上げ]					
D 型調査			C 型調査		
1 そう思う	18.0 (1.6)	} 55.6 (2.0)	1 そう思う	43.6 (2.0)	
2 ややそう思う	37.5 (2.0)		2 そう思わない	43.4 (2.0)	
3 あまりそう思わない	25.6 (1.8)	} 41.4 (2.0)	3 その他・D.K.	13.0 (1.4)	
4 そう思わない	15.7 (1.5)				
5 その他・D.K.	3.1 (0.7)				
問 4 あなたは、最近の世の中の流れは速すぎてついていけない、と感じることがあります。それともありませんか。[以下、C 型調査票では選択肢を読み上げ]					
C 型調査			D 型調査		
1 よくある	14.7 (1.4)	} 70.5 (1.9)	1 ある	56.2 (2.0)	
2 ときどきある	55.7 (2.0)		2 ない	41.1 (2.0)	
3 あまりない	24.0 (1.8)	} 28.3 (1.8)	3 その他・D.K.	2.7 (0.7)	
4 全くない	4.3 (0.7)				
5 その他・D.K.	1.2 (0.5)				

例えば、鈴木(1968)は、パネル調査において、社会問題に関する「関心度」を三つの選択肢で尋ねた場合と、四つの選択肢で尋ねた場合の結果を示している。論文中で言及されていないが、付録の第 30 表から第 36 表によれば、四選択肢に比べ、三選択肢の方が‘D.K.’のパーセントは大きい。また、Converse(1976)は、Gallup と Harris の世論調査の結果を利用し、‘D.K.’の比率が高くなる条件を探っている。選択肢の数に関しては、Harris のデータで、二つの選択肢だけを用いた質問項目の‘D.K.’比率が、そうでない質問項目の‘D.K.’比率よりも高い傾向が見られている。ただし、Gallup のデータではそういった傾向は見られず、原因についてはよく分からないとしている。

4.1.2 選択肢の副詞の違い

表 7 に示す問 3 と問 11 は、選択肢の数は全て四つであるものの、2 種類の調査票の間で、各選択肢に冠する副詞が異なる。そのため、四つの選択肢の間の心理的な距離が異なっている。

2 種類の調査票の間で回答分布を比較すると、期待されるとおり、非常に‘や’まったく‘など極端な副詞を用いた場合の方が、用いない場合に比べパーセントが小さくなっている。四つの選択肢から成る尺度の中心は、表 7 の選択肢ではいずれの場合も、選択肢 2 と選択肢 3 の中間にあると考えられるので、2 種類の調査票の間で、選択肢 1 と選択肢 2 の合計どうし、また選択肢 3 と選択肢 4 の合計どうしを比較する。まず、問 3 では 2 種類の調査票の間に差は認められず、副詞表現が変わっても尺度の中心は動かないと言ってよいであろう。問 11 では、C 型調査の‘満足’+‘やや満足’に比べ、D 型調査の‘非常に満足’+‘満足’の方が 4.9 ポイント大きくなっている($\chi^2_1 = 4.42, p = .04$)。この理由はよく分からないが、あえて挙げるとすれば、D 型調査票では‘満足’と‘不満’の間の距離が大きすぎたため、‘不満’とまではいえずに‘満足’に留まった回答者がいたのかもしれない。

4.1.3 選択肢の提示順序の違い

表 8 に示す問 12、問 13、問 14 は、選択肢の数や副詞表現は、2 種類の調査票の間で変わらないものの、選択肢の提示順序が逆転している。つまり、C 型調査票では、表 8 の一番上にある

表 7. 選択肢の副詞の違い(%) .

問3 あなたは、日常、世間でつかわれている言葉づかいなどからみて、今の国語は乱れていると思いますか。[以下、選択肢を読み上げ]					
C型調査			D型調査		
1 非常に乱れていると思う	27.4 (1.8)	} 86.3 (1.4)	1 乱れていると思う	42.2 (2.0)	} 87.4 (1.4)
2 少し乱れていると思う	58.9 (2.0)		2 少し乱れていると思う	45.2 (2.1)	
3 あまり乱れているとは思わない	12.0 (1.3)	} 12.7 (1.4)	3 あまり乱れているとは思わない	7.2 (1.1)	} 11.3 (1.3)
4 まったく乱れているとは思わない	0.7 (0.4)		4 乱れているとは思わない	4.1 (0.9)	
5 その他・D.K.	0.9 (0.4)		5 その他・D.K.	1.3 (0.5)	
問11 あなたは、「自分の暮し向き」について満足していますか、それとも、不満がありますか。[以下、選択肢を読み上げ]					
D型調査			C型調査		
1 非常に満足	10.0 (1.3)	} 80.9 (1.6)	1 満足	30.3 (1.9)	} 76.0 (1.7)
2 満足	70.9 (1.9)		2 やや満足	45.7 (2.0)	
3 不満	15.9 (1.5)	} 18.4 (1.6)	3 やや不満	19.1 (1.6)	} 23.7 (1.7)
4 非常に不満	2.5 (0.6)		4 不満	4.5 (0.8)	
5 その他・D.K.	0.7 (0.3)		5 その他・D.K.	0.3 (0.3)	

選択肢から下にある選択肢へと順に読み上げているのに対し、D型調査票では、'その他・D.K.'を除いた一番下にある選択肢から順に上にある選択肢へと読み上げている。

問12では、'離婚したいならさっさとしたらよい'や'場合によっては離婚はやむを得ない'という離婚肯定の選択肢は、それらが四つの選択肢の後半に提示されたD型調査よりも、前半に提示されたC型調査の方がパーセントが大きくなっている(それぞれ $\chi^2_1 = 5.66, p = .02$ と $\chi^2_1 = 2.80, p = .09$)。また、問14の'生命や健康や自分'は、8番目に提示されたD型調査よりも、1番目に提示されたC型調査の方が12.1ポイント大きき ($\chi^2_1 = 24.9, p < .01$)、'愛情や精神'は6番目に提示されたC型調査よりも、3番目に提示されたD型調査の方が9.6ポイント大きくなっている ($\chi^2_1 = 25.0, p < .01$)。これらの結果は、選択肢がリストのはじめの方で提示されることによって選択されやすくなる初頭効果によるものと考えてよいであろう。

一方、問13は、C型調査とD型調査とでは回答分布が異なるようであるが、その理由は明白ではない。考えうる理由の一つとしては、D型調査では、'困ったことであり、危険なことでもある'という選択肢が最初に提示されたことで、極端過ぎる意見という印象を与えてしまい、これを嫌った回答者が、2番目に提示された'望ましいことではないが、避けられないことである'へ流れた、ということがあるかもしれない。

4.2 調査所要時間と選択肢の提示効果

4.1節では、選択肢の数や提示順序などが回答に与える影響を見た。この節では、短時間群と長時間群との間で、その影響の大きさを比較する。

表9は、調査所要時間の群ごとに、D型調査の回答分布の、C型調査の回答分布からの Kullback-Leibler (KL) 距離を示したものである。つまり、 $P_{a,C}$, 短時間群 を、C型調査におけるある選択肢 a に対する短時間群の回答比率とし、 $P_{a,D}$, 短時間群、 $P_{a,C}$, 長時間群、 $P_{a,D}$, 長時間群 をそれ

表 8. 選択肢の提示順序の違い(%).

問		C型調査 (1)	D型調査 (1)
問 12	「離婚」について、あなたの考えに最も近いものを、次の4つの中から1つだけあげてください。[以下、選択肢を読み上げ]		
	離婚したいならさっさとしたらよい	7.4 (1.0)	4.3 (0.8)
	場合によっては離婚はやむを得ない	44.7 (2.0)	40.0 (2.0)
	努力して離婚はなるべく避けた方がよい	38.9 (2.0)	45.6 (2.1)
	どんな場合にも離婚は避けた方がよい	8.2 (1.2)	9.0 (1.3)
5	その他・D.K.	0.7 (0.2)	1.1 (0.4)
問 13	コンピュータがいろいろなところに使われるようになり、情報化社会などということがいわれています。このような傾向が進むにつれて、日常生活の上で変わっていく面があると思われまます。あなたは、このような変化を、どう思いますか。[以下、選択肢を読み上げ]		
	望ましいことである	18.2 (1.5)	17.0 (1.5)
	望ましいことではないが、避けられないことである	66.8 (1.9)	73.3 (1.8)
	困ったことであり、危険なことでもある	12.4 (1.4)	8.7 (1.2)
4	その他・D.K.	2.5 (0.7)	1.1 (0.4)
問 14	あなたにとって一番大切と思うものは何ですか。[以下、選択肢を読み上げ]		
	生命や健康や自分	31.1 (1.9)	19.0 (1.6)
	子ども	7.8 (1.1)	5.4 (0.9)
	家族	42.4 (2.0)	44.1 (2.0)
	先祖や家	1.4 (0.5)	2.3 (0.6)
	金や財産	1.3 (0.5)	1.8 (0.5)
	愛情や精神	7.8 (1.0)	17.4 (1.6)
	仕事や信用	4.0 (0.8)	5.1 (0.8)
	国家や社会	2.4 (0.6)	2.3 (0.6)
9	その他・D.K.	1.7 (0.5)	2.6 (0.7)

それ同様に定義すると、

$$KL_{\text{短時間群}} = \sum_a \ln \left(\frac{P_{a,C, \text{短時間群}}}{P_{a,D, \text{短時間群}}} \right) P_{a,C, \text{短時間群}}$$

$$KL_{\text{長時間群}} = \sum_a \ln \left(\frac{P_{a,C, \text{長時間群}}}{P_{a,D, \text{長時間群}}} \right) P_{a,C, \text{長時間群}}$$

を示したものである。ただし、C型調査とD型調査の回答分布を比較できるよう、以下のような処理を行った。まず、選択肢の数が異なる問2と問4では、表6にあるように四つの選択肢は二つにまとめた。また、選択肢の副詞が異なる問3と問11では、表7にあるように四つの選択肢を二つにまとめた。集計のために用いた重みは(3.2)式の $w^{(2)}$ であり、'その他・D.K.'を含む場合と、それを除いた上で合計が100%となるよう調整した場合とを示している。さらに参考のため、C型調査とD型調査とで同じ質問項目を用いた問1・問5・問9・問10のKL距離も示した^{注3)}。

‘その他・D.K.’を除いてC型調査とD型調査の間の比較を行うと、選択肢の数が異なる場合と選択肢の提示順序が異なる場合に、長時間群は短時間群に比べてKL距離の値が小さい。つまり、長時間群はC型調査とD型調査の間で回答分布の差が小さく、短時間群に比べて選

表 9. 回答分布間の Kullback-Leibler 距離(C 型・D 型).

質問項目	‘その他・D.K.’を含む		‘その他・D.K.’を除く	
	短時間群	長時間群	短時間群	長時間群
選択肢の数の違い				
問 2 日本人の考え方	11,143	11,596	3,662	63
問 4 世の中の流れ	6,691	4,147	6,669	3,588
選択肢の副詞の違い				
問 3 国語の乱れ	169	283	179	174
問 11 暮し向き満足か	1,172	789	790	794
選択肢の提示順序の違い				
問 12 離婚観	4,885	1,968	4,008	1,958
問 13 コンピュータ社会は好ましいか	4,904	3,215	4,050	1,296
問 14 一番大切なもの	14,867	5,571	14,471	5,607
同一項目				
問 1 外国暮らし	106	76	0	66
問 5 金か人間のつながりか	48	468	49	370
問 9 就職のとき	2,324	1,312	2,000	1,364
問 10 子供	1,367	306	1,108	265

注) 10,000 倍した値を示してある。

択肢の提示方法の影響が小さい、と言える。選択肢の副詞が異なる場合には、長時間群と短時間群の差は認められない。これは表 7 で見たとおり、副詞が変わっても、問 3 では尺度の中心は動かず、問 11 では動くとしてもその大きさがわずかであるためであろう。

‘その他・D.K.’を含めて比較を行うと、長時間群は、‘その他・D.K.’を除いた場合に比べ、特に問 2 の KL 距離の値が、63 から 11,596 へと大きくなっている。これは、問 2 では選択肢の数を四つから二つへと減らすと‘その他・D.K.’の回答比率が大きくなりやすいことと、長時間群では‘その他・D.K.’のパーセントが大きくなりがちであることが影響しているためであろう。

表 10 は、‘その他・D.K.’を除いた回答分布を C 型調査と D 型調査の間で比較したとき、パーセントの差の絶対値が短時間群と長時間群との間で 3 ポイント以上異なっていた択肢を具体的に示したものである。つまり、

$$100 \times (|P_{a,C, \text{短時間群}} - P_{a,D, \text{短時間群}}| - |P_{a,C, \text{長時間群}} - P_{a,D, \text{長時間群}}|) \geq 3 \text{ ポイント}$$

を満たす択肢 a を挙げたものであり、表中の数値は、 $100 \times |P_{a,C, \text{短時間群}} - P_{a,D, \text{短時間群}}|$ と $100 \times |P_{a,C, \text{長時間群}} - P_{a,D, \text{長時間群}}|$ である。問 2 と問 4 は択肢が二つであるため、例えば問 2 の‘そう思わない(あまりそう思わない + そう思わない)’は、二つの調査間で、短時間群では 13.4 ポイント、長時間群では 1.8 ポイント差があることになるが、冗長であるため表には示していない。

χ^2 検定の結果、有意水準 10% 未満で有意な差が見られた択肢は三つだけであるが、問 12 の‘努力して避ける’を除き、表 10 に挙げた全ての択肢で、調査票の違いによるパーセントの変動は、長時間群に比べ短時間群の方が大きい。つまり電話調査では、択肢の数や提示順序など、択肢の提示方法が回答に影響を与えることがあるが、その影響は、表 9 や表 10 に示す質問項目では、長時間群よりも短時間群の方が大きいようである。1.2.2 節の観点 3 については、本稿の調査データでは、「予想される結果 3a」ではなく「予想される結果 3b」が得られたことになる。

表 10. 短時間群と長時間群との間で、2 種類の調査問の差が 3 ポイント以上異なっていた選択肢(C 型・D 型).

問番号	選択肢	2 種類の調査問の%差	
		短時間群	長時間群
問 2	そう思う (そう思う + ややそう思う)	13.4	1.8
問 4	よくある + とときどきある (ある)	17.8	12.9
問 10	必要はない	4.9	1.2
問 12	努力して避ける	4.0	9.3
問 12	どんな場合にも避ける	4.6	0.7
問 13	望ましいこと	8.6	3.2
問 13	避けられないこと	12.4	1.2
問 14	生命や健康や自分	15.7	10.0
問 14	愛情や精神	14.4	7.5

≫: $p < .01$ ≧: $p < .05$ >: $p < .10$

5. おわりに

5.1 結果のまとめ

本稿における分析の結果を、以下に箇条書きの形でまとめる。

- ・若年層に比べ、高齢層の方が調査所要時間は長くなる傾向が見られる。
- ・長時間群と短時間群との間で、調査協力理由に関する質問への回答に実質的な違いは認められない。つまり、1.2.2 節の観点 1 で述べた「予想される結果 1」は得られなかった。
- ・短時間群に比べ、長時間群の方が、「その他・D.K.」の回答比率が高い。つまり、1.2.2 節の観点 2 で述べた「予想される結果 2a」ではなく、「予想される結果 2b」が得られた。
- ・長時間群に比べ、短時間群の方が、選択肢の提示方法による回答への影響が大きい。つまり、1.2.2 節の観点 3 で述べた「予想される結果 3a」ではなく、「予想される結果 3b」が得られた。

以上の結果から判断すると、電話調査における調査所要時間の長さは、回答にあたっての回答者の吟味の深さを表すものと考えられる。つまり、本稿で得られた短時間群と長時間群の特徴を、やや乱暴ではあるが、簡単にまとめると次のようになる。短時間群は、与えられた選択肢の中から直ちに回答を選ぶことができるため、調査所要時間は短くなるが、時間をかけて選択の適切さを吟味しないため、質問文や選択肢の表面的な印象の影響を受けやすい。一方、長時間群は、与えられた選択肢の中から回答をなかなか選べず悩んでしまうために、調査所要時間が長くなり、「その他・D.K.」は多くなる。しかし、与えられた選択肢の中から回答を選んだ場合には、時間をかけて選択の適切さを吟味しているため、質問文や選択肢の表面的な印象の影響は受けにくい、と考えられる。

もちろん、以上の結果は、特定の質問項目についての 1, 2 回の調査だけから得られたものである。特に選択肢の順序効果などは、“sometimes you see them, sometimes you don't” (Sudman et al. (1996), p. 160) であるため、他の調査データによって繰り返し検討していく必要がある。また、高齢層の方が調査所要時間が長くなるからといって、必ずしも高齢層の「その他・D.K.」が多くなるとは限らない(ただし西平(1958)は高年齢に「わからない」が多いと述べている)、高齢者の方が選択肢の提示方法の影響を受けにくいわけではないことにも注意する必要がある。

5.2 今後の課題

1.2.1 節で紹介した Bassili らの研究結果の多くでは、反応潜時の短さが、態度の強さ・安定

性を示していた。これに対し、調査所要時間の場合には、時間が長い方が、選択肢の提示効果の影響が小さいという意味で、回答は安定していた。回答に要する時間という意味では反応潜時と調査所要時間も同じであるにもかかわらず、両者の間で一見矛盾したような結果が得られた理由としては、以下のようなことが考えられる。

第一に、1.2.1 節で紹介した調査の多くは、選挙に関連した項目を用いているのに対し、本稿の調査は「日本人の国民性調査」の項目を多く用いている。選挙に関して態度が一貫している回答者は、調査に回答する以前から、調査内容のような事柄については十分思案し、態度を決めているのであろう。そのため、質問をされても、短い反応潜時で回答することができる。一方「日本人の国民性調査」の質問項目は、普段はあまり意識して考えず、調査時にはじめて態度を決めなければならないような内容が多い。そのため、時間をかけて回答を選んだ方が、選択肢の提示方法といった表面的な印象の影響は小さくなると考えられる。実際、統計的には有意ではなかったが、Bassili and Krosnick (2000)の結果では、妊娠中絶の是非に関する質問の順序効果や、民主主義に反対する演説の可否を問う質問文でワーディングを変えた効果が小さいのは、反応潜時が長い回答者であった。したがって、質問項目の内容・分野によって、反応潜時や調査所要時間の意味が異なる可能性がある。

第二に、調査所要時間には、オペレータが質問文や選択肢を読み上げる時間も含まれている。例えば、オペレータが再度、質問文と選択肢を読み上げた場合は、選択肢の提示順序などの効果は薄まると考えられる。その上、オペレータが再度読み上げている最中にも、回答者はどの選択肢を選ぶべきか熟考を重ねることができる。したがって、調査所要時間が長いからといって、個々の質問項目の反応潜時も長いとは限らない。

Krosnick and Alwin (1987)は、回答する際に選択の適切さを心的に十分吟味しないことを *satisfice* と呼び、*satisfice* な回答者は順序効果などの影響を受けやすいとしている。これは、まさに上記の短時間群の特徴に合致しており、調査所要時間の短さは *satisfice* の程度を示すものと考えられる。ただし、Krosnick and Alwin (1987)は、回答者が *satisfice* になるのを避けるためには、回答者の動機づけを高めるのがよいかもしれないと述べている(p.216)。しかし本稿の結果では、短時間群と長時間群との間で、調査への協力理由に大きな違いは認められなかった。したがって、この点も含め、以下のような疑問が、今後の課題として残されている。

- ・ 調査協力理由に関する質問項目や、そこで用いられた選択肢が、はたして調査協力の理由あるいは動機づけを調べるものとして適切であるのか。

- ・ 各質問項目の反応潜時と調査所要時間との関係はどうなっているのか。短時間群は本当に *satisfice* なのか。逆に長時間群は、短時間群に比べ、調査所要時間のうちのどのプロセス(オペレータが質問文を読み上げる段階、回答者が回答を考える段階、オペレータが回答を入力する段階、など)に時間がかかっているのか。

- ・ 選挙調査のような質問項目でも、調査所要時間が短いほど回答は不安定なのか。質問項目の内容・分野によって、調査所要時間の意味が異なってくるのか。

- ・ ‘その他’と‘D.K.’を別の選択肢として分けて扱くと、それぞれ調査所要時間との関係はどうなっているのか。

- ・ 学歴や職業など、本稿の調査では調べなかった回答者属性(表3に示された以外の回答者属性)と調査所要時間との関係はどうなっているのか。

5.3 調査所要時間の活用方法

5.1 節でまとめた短時間群と長時間群の特徴はあくまでも仮説にすぎない。しかし、仮にそれらの特徴が多くの調査で言えるものであれば、本稿で検討した選択肢の数やその提示順序に

限らず、さらに一步進めて、短時間群は、直前の質問項目や調査モードなど広い意味での文脈効果の影響を受けやすいのではないかと、という想像は容易に成り立つ。つまり、短時間群に属する人は、同じ質問項目であっても、電話調査の場合の回答と面接調査の場合の回答とが異なる可能性が大いにある。本稿の結果は、調査モード間の比較可能性を考える際には、調査所要時間も考慮するとよいのではないかと、という示唆を与えるものと言える。

調査所要時間の活用方法としては、他にも次のようなことが考えられる。普通、質問項目の表現方法に起因する非標本誤差の大きさを評価するためには、本稿の C 型調査票と D 型調査票のような 2 種類の調査票を用いた split-ballot デザインを利用するのがよい。しかし多くの調査では、予備調査でない限り、複数の調査票を用いることは困難である。そこで、1 種類の調査票であっても、調査所要時間によって回収サンプルを二分し、短時間群と長時間群の間で回答分布を比較するのである。もし、2 群の間で回答分布が大きく異なっていたとすると、その原因の一つとして、短時間群が質問の意味内容や自身の回答の適切さをよく吟味せず、質問項目の第一印象にとらわれて回答してしまったことが考えられる。つまり、2 群間で回答分布が大きく異なる質問項目は、表面的な印象によって回答が選ばれてしまいがちな、非標本誤差の大きな質問項目である可能性がある。もし質問項目の表現方法が変われば、その影響を受けて短時間群の回答分布が変動し、結果として全体の回答分布も変動するおそれがあるということである。

例えば、表 5 の '#5.23 能力か功労か' では、長時間群に比べ、短時間群では 1 番目の選択肢である '能力重視' という回答が 8.6 ポイント多い。'能力重視' が最初に提示された選択肢であることと、その回答比率は短時間群の方が大きいことを勘案すると、この差は、短時間群で初頭効果が強くあらわれたためである可能性がある。そこで、仮に選択肢の提示順序を逆転すれば、短時間群では、1 番目に提示されることになる '功労重視' のパーセントが、表 5 の 27.3% よりも大きくなり、全体としても '功労重視' の回答比率が上がる、と予想される。選択肢の順序だけでなく、選択肢の表現も異なっているため、直接の比較はできないが、実際に K 型調査では選択肢の順序を変えて同じような内容を質問した結果、'功労重視' は全体で 47.6% となり、M 型調査の 28.6% に比べ増加している^{注 4)}。

電話調査に限らず、面接調査においても CAPI (Computer Assisted Personal Interviewing) など CASIC (Computer Assisted Survey Information Collection) への動きは盛んである (Couper et al. (1998))。CASIC では、コンピュータを利用して、調査所要時間を自動的に記録することは容易である。電話調査を含め、さまざまな調査モードにおいて、調査所要時間をどのように活用できるのか、検討していく価値はあると思われる。

注 1) 2002 年度調査の実施にあたっては、平成 14 年度・15 年度文部科学省科学研究費補助金(若手研究 A)課題番号: 14701019 研究代表者: 土屋隆裕 研究課題名: 統計的調査法としての RDD 法の比較実験調査)の助成を得た。

注 2) ただし、土屋 他(2004)で集計対象としたサンプルは「日本人の国民性第 11 次全国調査」のそれに合わせるため、20 歳以上 80 歳未満に限定している。

注 3) 問 1 は、質問文が若干異なり、C 型調査票では「あなたは、機会があれば外国で暮らしてみたいと思いますか」、D 型調査票では「あなたは、機会があれば外国で暮らしてみたいと思いますか、それとも暮らしてみたいとは思いませんか」である。結果は、「思う」が 30.5%(C 型)と 30.6%(D 型)、「思わない」が 68.9%(C 型)と 68.4%(D 型)であり、両調査票の間に差は見られない。

注 4) K 型調査の質問文は「会社で給料や賃金を決めるときに、その人のこれまでの会社に対する功労を重視して決めるべきだと思いますか、それとも、その人の現在の能力だ

けで決めるべきだと思いますか」であり；功労重視'は47.6%；能力重視'は43.5%，
 'その他・D.K.'は8.9%。M型調査の質問文は「会社で給料や賃金を決めるときに，そ
 の人の現在の能力を重視して決めるべきだと思いますか，それとも，その人のこれま
 での会社に対する功労を重視して決めるべきだと思いますか」であり；能力重視'は
 64.0%；功労重視'は28.6%；その他・D.K.'は7.4%。

謝 辞

分析にあたりましては，統計数理研究所 藤澤洋徳助教授から貴重なご助言をいただきました。また，2名の匿名の査読者の方と担当編集委員からは，論文の構成・表現などについて建設的なご提案をいただきました。この場を借りて感謝いたします。

参 考 文 献

- Bassili, J. N.(1993) Response latency versus certainty as indexes of the strength of voting intentions in a CATI survey, *The Public Opinion Quarterly*, **57**, 54-61.
- Bassili, J. N.(1995) Response latency and the accessibility of voting intentions: What contributes to accessibility and how it affects vote choice, *Personality and Social Psychology Bulletin*, **21**, 686-695.
- Bassili, J. N.(1996a) The how and why of response latency measurement in telephone surveys, *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research* (eds. N. Schwarz and S. Sudman), 319-346, Jossey-Bass Publishers, San Francisco, California.
- Bassili, J. N.(1996b) Meta-judgmental versus operative indexes of psychological attributes: The case of measures of attitude strength, *Journal of Personality and Social Psychology*, **71**, 637-653.
- Bassili, J. N. and Fletcher, J. F.(1991) Response-time measurement in survey research. A method for CATI and a new look at nonattitudes, *The Public Opinion Quarterly*, **55**, 331-346.
- Bassili, J. N. and Krosnick, J. A.(2000) Does attitude strength moderate susceptibility to response effects? New evidence using response latency, attitude extremity, aggregate indices, and continuous measures, *Political Psychology*, **21**, 107-132.
- Bassili, J. N. and Scott, B. S.(1996) Response latency as a signal to question problems in survey research, *The Public Opinion Quarterly*, **60**, 390-399.
- Converse, J. M.(1976) Predicting no opinion in the polls, *The Public Opinion Quarterly*, **40**, 515-530.
- Couper, M. P., Baker, R. P., Bethlehem, J., Clark, C. Z. F., Martin, J., Nicholls II, W. L. and O'Reilly, J. M.(eds.)(1998) *Computer Assisted Survey Information Collection*, John Wiley & Sons, New York.
- Fazio, R. H. and Williams, C. J.(1986) Attitude accessibility as a moderator of the attitude-perception and attitude-behavior relations: An investigation of the 1984 presidential election, *Journal of Personality and Social Psychology*, **51**, 505-514.
- 林 英夫(2004). 『郵送調査法』, 関西大学出版部, 大阪.
- Krosnick, J. A. and Alwin, D. F.(1987) An evaluation of a cognitive theory of response-order effects in survey measurement, *The Public Opinion Quarterly*, **51**, 201-219.
- Krosnick, J. A. and Schuman, H.(1988) Attitude intensity, importance, and certainty and susceptibility to response effects, *Journal of Personality and Social Psychology*, **54**, 940-952.
- Krosnick, J. A., Boninger, D. S., Chuang, Y. C., Berent, M. K. and Carnot, C. G.(1993) Attitude

- strength: One construct or many related constructs?, *Journal of Personality and Social Psychology*, **65**, 1132–1151.
- Lin, I. F. and Schaeffer, N. C. (1995). Using survey participants to estimate the impact of nonparticipation, *The Public Opinion Quarterly*, **59**, 236–258.
- Lynn, P., Clarke, P., Martin, J. and Sturgis, P. (2002). The effects of extended interviewer efforts on nonresponse bias, *Survey Nonresponse* (eds. R. M. Groves, D. A. Dillman, J. L. Eltinge and R. J. A. Little), 135–147, John Wiley & Sons, New York.
- 前田忠彦, 土屋隆裕 (2001). 日本人の国民性 2000 年度吟味調査報告～電話・郵送・面接調査の比較～, 統計数理研究所研究レポート, No. 87.
- 西平重喜 (1956). 面接調査法の諸問題その 2, 統計数理研究所彙報, **4**(2), 1–7.
- 西平重喜 (1958). 面接調査法の諸問題その 4, 統計数理研究所彙報, **6**(1), 73–85.
- Payne, S. L. (1951). *The Art of Asking Questions*, Princeton University Press, Princeton, New Jersey.
- Research Triangle Institute (2004). *SUDAAN Language Manual, Release 9.0*, Research Triangle Institute, Research Triangle Park, North Carolina.
- 坂元慶行, 中村 隆, 前田忠彦, 土屋隆裕 (2004). 国民性の研究第 11 次全国調査—2003 年全国調査—, 統計数理研究所研究レポート, No. 92.
- Schuman, H. and Presser, S. (1981). *Questions and Answers in Attitude Surveys*, Academic Press, New York.
- Stein, R. M. and Johnson, M. (2001). Who will vote? The accessibility of intention to vote and validated behavior at the ballot box, Paper presented at the annual meeting of the American Political Science Association, San Francisco, California.
- Sudman, S., Bradbur, N. M. and Schwarz, N. (1996). *Thinking about Answers*, John Wiley & Sons, New York.
- 鈴木達三 (1968). 面接調査における回答の安定性について, 統計数理研究所彙報, **16**(1), 47–102.
- 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.
- 鈴木達三, 水野欽司, 大隅 昇, 中村 隆, 長坂和子, 田中さえ子 (1981). 社会調査の実施過程における調査誤差の研究, 統計数理研究所研究レポート, No. 52.
- Tourangeau, R., Rips, L. J. and Rasinski, K. (2000). *The Psychology of Survey Response*, Cambridge University Press, Cambridge.
- 土屋隆裕, 前田忠彦 (2003). 二種類の電話調査法の比較実験調査, 行動計量学, **30**(1), 93–109.
- 土屋隆裕, 前田忠彦, 中村 隆, 坂元慶行 (2004). 調査モード間の比較研究～2002 年度・2003 年度調査～, 統計数理研究所研究レポート, No. 93.
- Tyebjee, T. T. (1979). Response latency: A new measure for scaling brand preference, *Journal of Marketing Research*, **16**, 96–101.

Length of Interview and Response Effects in Telephone Surveys

Takahiro Tsuchiya

The Institute of Statistical Mathematics

This paper investigates the relationship between the lengths of telephone interviews and response effects, and explores the practical use of the lengths of interviews. Respondents to two telephone surveys are classified into two groups based on the length of their interviews, and the characteristics of the responses between the short time and the long time groups are compared. Firstly, logistic regression analyses reveal that interviews with older respondents tend to be longer. The respondents' reasons for cooperation with the survey are irrelevant to the interview length, although the existence of such relevance was hypothesized. Secondly, more respondents in the long time group tend to select the option "Other or Don't Know" than those in the short time group. Thirdly, the response effects of changing options in terms of their number, wording, or ordering, are reported based on the results of a split-ballot survey. The magnitude of the response effects appears to be more significant in the short time group than in the long time group. Finally, some potential uses of the time length of the interview for data analyses are discussed on the basis of the results obtained in this paper.