

面接調査における回答誤差

鈴木 達 三

(1964年11月受付)

A Study of Interviewing in Social Research

Tatsuzo Suzuki

In this paper we studied some of the response errors arising from the interviewing.

Our data are based on analyses of the study carried out by H. Aoyama and others [1].

Institute of Statistical Mathematics

1. はじめに

この報告は、「面接調査における偏りの統計的研究—I」[1]につづくものである。[1]では面接調査における偏りのうち、主として a) 無回答によるもの、b) 調査員のゴマカシ、被調査者の虚偽、c) 面接の場における雰囲気、調査員の熟練の程度、指示の履行状態などに起因するものについて分析している。ここでは実際の調査のとき、面接調査における回答の模様を記録する場合の誤差、すなわち、被調査者の回答したものが正しく調査票に記入されているかどうかを、この目的のため調査に協力を依頼したサンプルから得られたデータを使って分析する。

ふつう行なわれている社会調査の結果の分析についてもそうであるが、とくに面接調査における誤差の分析についても多くの場合は、面接調査の結果得られた調査票に記入してあるものについての分析が主になっている。したがって、面接調査をしたとき、被調査者の回答したものを、調査員が正確に調査票に記入しているかどうかを検討することは、いろいろな意味で必要なことである。もちろん、調査指示を守らないような場合には、それだけで誤りや、偏りが生じてしまうわけであるが、たとえ実際の場面において正確に調査指示を守って調査をおこなっても、調査員が調査票に記入する段階で、被調査者の回答を正確に記入しなければ、調査全体として好ましくないことになる。

ここでおこなうものは、このような点について検討するための実験データの分析である。すなわち、被調査者のうちの何人かに、あらかじめ研究の目的を話して調査に協力するよう依頼し、各自に自分が回答したものを書いてもらい、それと調査員の記入したものとを付け合せて、どのような場合にどのような誤りが生じるかをみようとしたのである。

これと同趣旨のことは、調査の準備段階で屢々検討されている。たとえば、[2]においては調査員間の記入の喰い違いによる影響を、二、三の誤差モデルを用いて検討している。この結果によれば、多くの場合、回答記入の誤りに一定の傾向がみられないので、誤差モデルにより、回答誤差を処理しうることが示されている。しかし、特定の調査員による予備調査の資料であるから問題は残っていると思われる。

われわれの実験調査で得られたことから、一般の調査における、この種の調査誤差を類推す

るには、調査誤差の起り方、その大きさ、およびそれが推定結果に与える影響を考察する必要がある。

調査員による記入誤差が、質問の形式、調査の形式等々により特定の傾向をもつかどうか分れば、それにより、適当な誤差モデルを作成して検討することが可能であると思われる。

2. 実験の方法

われわれは[1]にくわしく述べられているように、多面的な実験調査を計画したが、調査票記入のさいの誤りを検討するためつぎのような方法をとった。

a) 調査の計画

調査員は一般のランダムサンプル(Aサンプルと呼ぶ)12人と、調査協力を依頼したサンプル(Bサンプルと呼ぶ)2人、転出してから調査時期までが1年以内のサンプル(Cサンプルと呼ぶ)1人の合計15人を受持ち面接調査するように計画された。調査員にはA、B、Cサンプルの区別は知らされていないので、Bサンプル、Cサンプルについて行なわれた調査の結果から、通常の調査における調査票記入の際の誤りなどを検討することができる。

b) 調査票

調査に用いた質問項目は第1表にしめすように、われわれが毎年定期的実施しているマス・コミュニケーションの効果に関する研究における質問項目と同様のものが大部分をしめているが、その中でも、いろいろな型式の質問をとり上げ、それぞれの場合に生じる誤りを検討できるようにした。

c) 調査に協力を依頼したサンプル(Bサンプル)への指示

Bサンプルは研究所職員の家族、知人等に研究の目的を説明し調査協力を依頼した。調査にあたり、注意した点はつぎのようなことである。

- i) あらかじめ配布してあるBサンプルの調査票に面接調査員の質問に対して回答したことをそのまま記入してもらうようにした。
- ii) 質問文を省略したり、一部分を強調したり、あるいはいわなかったりした場合、調査を指示通しにおこなわなかった場合は、その旨を記入してもらう。
- iii) そのほか、調査員の質問時における態度、動作、服装などを一定の基準で判定してもらった。

d) このように、面接調査時における各種のデータをとるために調査協力を依頼した人々の数はそれほど多くなかったため、これらの人々はそれぞれ同一の人が2人の調査員から面接調査を受けるように計画した。

調査が実験的なものになるのをできるだけ避けるため、調査の順序、調査員の配置などを適当に組合せ、同一のBサンプルには、調査の前半に1人の調査員が面接調査し、調査の後半に別のグループの調査員が訪問するようにした。

3. 調査協力を依頼したサンプルから得られたデータによる実際場面におこる誤差の検討

3.1 調査協力を依頼したサンプル(Bサンプル)のデータ

調査のとき、調査協力を依頼したサンプル(Bサンプル)は、他の一般のサンプルと同じ状態で調査員の訪問調査を受けるように計画されていたので、調査員がとくにBサンプルと一般のサンプルとの間で面接調査のやり方を差別したとは考えられないから、Bサンプルによるデータ(Bサンプル自身が調査票に自分が面接調査のとき回答したものと同じことを記入したものと)、調査員の記入したものととの関係をしらべれば実際場面におこる誤差がどのようなものであるかみることができる。

第1表 B-サンプルの記入と調査員の記入の一致度(質問別)

質問形式*	問 項 目**	コード違い (記入違い)	質 問 省 略	記入 不足	記入 過多	一致	計(B-サン プル延数)	備 考	
F	1a とっている新聞	2.4	0.8	2.4	—	94.4	100.0(125)	記入不足の() はコード違いも 含む	
M	1b いつもよむか	6.1	4.3	2.6	—	87.0	100.0(116)		
F	2 特によむ記事	15.2	2.5	7.6 (5.1)	8.9	60.7	100.0(79)		
F	5b 雑誌名	6.3	—	2.5	1.2	90.0	100.0(80)		
F	7 犬	33.8	1.2	3.8	—	61.2	100.0(80)		
F	9b 憲法改正理由	13.6	1.2	3.7	—	81.5	100.0(81)		
F	12c 歳暮理由	25.3	6.3	5.2	—	53.2	100.0(79)		
F	15 趣味	7.4	—	12.3 (2.5)	—	77.8	100.0(81)		同 上
M	3 重点おく政策	9.9	—	—	—	90.1	100.0(81)		
F+af	4a ラジオ番組 [1番]	15.4	3.9	—	—	80.7	100.0(78)		
"	" " [2番]	18.2	9.1	—	—	72.7	100.0(77)		
"	" " [3番]	29.9	9.1	—	—	61.0	100.0(77)		
F	4b 昨夜の番組	6.3	5.1	—	—	88.6	100.0(79)		
M	5a 特集記事	44.4	—	—	—	55.6	100.0(81)		
D	6a 人工衛星(現在)	2.5	1.3	—	—	96.2	100.0(81)		
M(F+af)	6b " (将来)	11.1	—	—	—	88.9	100.0(81)		
D	8 自衛隊	7.4	2.5	—	—	90.1	100.0(81)		
D(M)	9a 憲法改正	28.4	1.2	—	—	70.4	100.0(81)	完全な逆転は 5%	
R(L)	10 予算	12.5	2.5	—	—	85.0	100.0(80)	コード違いには 2ヶ所違うもの 2.5%を含む	
M(L)	11 科学者	23.3	10.0	—	—	66.7	100.0(80)	質問省略には全 部5%一部5% がある	
D	12a 歳暮	8.8	—	—	—	91.2	100.0(80)	完全な逆転は 5%	
M(F+af)	12b 歳暮一般	44.3	2.5	—	—	53.2	100.0(79)		
M(P)	13a 広告	1.3	10.1	—	—	88.6	100.0(79)		
M(P)	13b 広告	6.4	12.8	—	—	80.8	100.0(78)		
M(L)	16 悩み	8.7	6.2	—	—	85.1	100.0(81)		
D(L)	14 持物	2.0	5.0	—	—	93.0	100.0(801)		
F+af	17a 職業(本人)	15.2	3.8	—	—	81.0	100.0(79)		
F+af	17b 職業(中心人物)	9.2	5.3	—	—	85.5	100.0(76)		
F	18 有権者数	3.8	—	—	—	96.2	100.0(80)		
F	19 年齢	7.6	—	—	—	92.4	100.0(81)		
F+af	20 学歴	3.7	1.3	—	—	95.0	100.0(80)		
F+af	21 支持政党	7.6	1.3	—	—	91.1	100.0(79)		

- *) F : 自由回答法
M : 多項選択技法
D : 二項選択技法
+af : 回答をそのままとり、その後で coding をおこなう
L : リスト(回答票使用)
P : 写真
R : rating

**) 質問文の全文は [1] をみよ。

しかし、ここで問題になるのは、Bサンプル自身の面接調査時における調査員への回答と、調査票記入時における回答との不一致である。とくに質問の形式が、多項選択法の場合、あるいは自由回答法の場合などは、力点のおき方(強調する場所)により、Bサンプル自身が調査員に回答したと信じている答と調査員の受け取り方との間には不一致がみられる場合も考えられる。それが調査員の個人的な誤りでなく、誰でもが不一致になるような種類のものであれば、Bサンプルのデータの方がおかしいことになるので、まずこの点を検討しておく必要がある。

そのため、数人のBサンプルには、テープレコーダーにより、実際場面の録音を依頼し、その回答状況を多くの人が聴取して、それぞれ調査票に記入して、Bサンプル自身の記入したものと比較検討したところ、とくに複雑な回答場面で、わずかの喰い違いがみられたが、全体的にはBサンプル自身の記入と回答状況とはよく合っていることが認められたので、以下の分析では、Bサンプルの記入した回答は、実際場面における正しい回答と考え、Bサンプルの記入したものと、調査員の記入したものとが不一致の場合は、すべて調査員の誤りと考えることにした。したがって、賛成、反対などの程度が問題になるような種類の質問では、Bサンプルの記入したものと調査員の記入したものととの不一致(コード違い)が多くなる場合も実際にみられるが、これは、質問の作成方法を改善すれば避けられることが多いと思われる。

3.2 Bサンプルの記入と調査員の記入との一致度

a) 質問別にみた一致度

調査に使用した質問文について、全文はここでは述べないが、質問別にBサンプル記入と調査員記入の一致度をみると第1表ようになる。

一致率が90%以上のものは、「持物(耐久消費財)」、「家族中の有権者数」、「学歴」、「年齢」、「支持政党」、「月ぎめ購読新聞名」、などのいわゆるフェースシート、あるいは事実をきく質問群のほか、「人工衛星(現在)」、「自衛隊」、「歳暮をやるか」などの質問である。また一致率の低いものは、「特によむ記事」、「人工衛星の特集記事」、「人工衛星にのった犬についての感想」、「お歳暮についてどう思うか」および「その理由」などの自由回答法の質問が多くなっており、このほか、「憲法改正」、「科学者と政治」についての質問も一致率が低かった。とくに「科学者と政治」の質問は、わざと質問のやり方を複雑に作ってあったので一致率が低いのは仕方がないが、「お歳暮についてどう思うか」および「憲法改正」の質問のように、賛成あるいは反対の程度をみるような質問では、賛成、反対が逆になってしまうものは、両者とも5%ほどであるが、強い賛成から弱い賛成へ、あるいは、弱い反対から強い反対へというように、程度の差をつけるところでコード違いが多くみられた(第2表参照)のは、いろいろの場合に注意しなければならないところであろう。すなわち、質問の結果を解釈するとき、賛否の程度を問題にするならば、調査員によるコード違いが存在する可能性が多いわけであるが、これを賛成の方、反対の方と、ただ傾向的に区分するだけならば、調査員のコード違いをそれほど問題にできなくともよいのであるから、質問作成のときに考慮に入れる必要もあろうと思われる。

つぎに、質問を省略された割合についてみると、「広告(写真をみせる)」、「科学者と政治」、「ラジオの番組の2番目と3番目」が10%前後でやや多くなっている。これは、それぞれ写真をみせる必要がある、質問のやり方が複雑である、被調査者の答が得にくいなどの理由によるものであろうか。その他の質問でもいくらか「質問省略」がみられる。

通常おこなわれている郵便調査による調査状況のチェックでは、指定されたサンプルについて訪問調査が正確におこなわれたかどうかをみるだけで、このような細かい点まで調べることはできないが、実際場面においては、郵便調査で発見できる以外に、このような「質問省略」などがあるので、調査が指示通りに正確におこなわれる割合はいくらか減じるものと思われる。

第2表 B-サンプルと調査員の記入のコード違い

a) 問9a 憲法改正						b) 問12b 一般の歳暮について								
調査員	Bサンプル	賛成	今は反対	反対	その他無答	計	調査員	Bサンプル	やめるべき	やめた方がよい	どちらともいえない	やった方がよい	やるべきだ	計
	賛成							やめるべき						
賛成			①	①		2	やめるべき			9			②	11
今は反対				5		5	やめた方がよい	8			3	①		12
反対		②	6			8	どちらともいえない			1		3		4
その他無答		3	1	2	2	8	やった方がよい			①	4		2	7
計		5	8	8	2	23	やるべきだ						4	34
							計	8	11	7	4	4	4	

注) ○印が逆転したもの、表中の数字はコード違いのみの数(実数)

る。

b) 質問の形式よりみた一致度

すでに前項でみた通り、記入の一致率の高いものは、フェースシート的なもの、事実をありのままに回答してもらい、それを記入する型の質問、あるいは「人工衛星では、米、ソどちらが進歩しているか」というような二項選択肢法の質問などである。一方、一致率の低いものは、自由回答法、多項選択肢法のものである。とくに自由回答法の場合に一致率を低くした原因をみると、意見、理由をきく場合には調査員の記入不足、記入過多ということよりも、要点の受けとり方の違いによるもの(コード違い)の方が多くみられ、「特によむ記事の種類(名前)」とか、「趣味」などの、どちらかといえば、事実(実態)についての回答を求めた場合には、記入不足が目立っている。

質問を形式別にまとめて、延質問数に対する「コード違い」などの割合などをみると、第3表のようになる。(質問のまとめ方は表の備考欄をみよ)。

第3表 質問の形式別にみた記入の不一致

質問の形式	コード違い(%)	記入過不足(%)	延質問数	備考
多項選択	20.4	—	396	問 3, 5a, 6b, 11, 12b, 16
二項選択	12.3	—	309	問 6a, 8, 9a, 12a
自由回答	15.1	7.9	595	問 1a, 2, 5b, 7, 9b, 12c, 15
基本項目 (フェース・シート)	4.1	—	1,146	問 14, 17a, 17b, 18, 20, 21

c) 調査票リストを使用した質問

質問によっては、回答の各選択肢をあらかじめカードに印刷しておき、そのリストを被調査者に呈示して、その中から被調査者の意見なり実態なりに一番適切と思われるものを選んでもらう方法をとる場合がある。これはいわば強制選択法ともいえるが、調査票リストを使用することにより、どのような影響があるかを「質問省略」と「コード違い」の二つの面から考えてみよう。

「質問省略」の割合は第4表のようになり、リストを使用した質問とそうでない質問とではやや差がみられる。

つぎに「コード違い」の割合をみると、リストを使用した質問と、そうでない質問とではい

第4表 「質問省略」の割合

質 問	質問省略(%)	計
リスト使用	5.1	1,042
リスト使用せず	2.5	1,271

くらか差がみられる(第5表参照)。もちろん、ここでの比較は同一の質問について、一方ではリストを使用し、他方でリストを使用しなかった場合を比較したのではなく、リストを使用した質問と、リストを使用することもできる別の質問との間の比較であるから、問題はありますが、リストを使用することにより「コード違い」が減るだろうということは考えられる。

第5表 コード違いの割合

質 問	コード違い(%)	延質問数
リスト使用	5.5	989
そ の 他	16.5	1,239

しかし、元のデータを考えてみると、Bサンプルは自分で調査票の選択肢に記入したのであるから、全部の質問をいわばリストを使用した状態で回答を記入したわけであり、リストを使用しない質問では、実際の調査場面における回答の受け取り方が、Bサンプルと調査員との間で喰い違っていたため、「コード違い」が多くなったとも考えられる。その上、調査票リストを使用することにより、被調査者の意見を拘束することになるから、質問のねらいによっては、好ましくない場合もあり、調査票リストを使用する方がよいかどうかは、にわかにはきめることはできない。

d) 質問の配列、訪問の前後による影響

調査票を質問の形式などには関係なく、前半分と後半分にかけて、質問を省略された割合をみると第6表のようになり、傾向的にみれば、後半では質問省略が多くなっているが、あまり配列順には関係はない。また、コード違いの方をみると、前半に多い傾向はみられるが、これも質問の配列順ということより、個々の質問別にみた「コード違い」の影響が強いように思われる。

第6表 質問の配列の前後による「質問省略」と「コード違い」の割合

質問の配列	質問省略(%)	計	コード違い(%)	延質問数
前(問1~問10)	2.7	1,350	16.7	1,313
後(問11~問21)	4.8	1,833	8.6	1,746

つぎに、Bサンプルを2人宛受け持った各調査員が、それぞれのBサンプルを自分の受け持った調査サンプルのうちの前半分の間に調査したか、後の方で調査したのかにより、前半、後半にかけて、訪問の前後による「質問省略」の割合をみると、第7表のようになる。

第7表 訪問の前後による「質問省略」の割合

訪問の順が	質問省略(%)	計
前 半	1.7	1,037
後 半	4.6	1,688

また、訪問の前後による質問形式ごとの「コード違い」の割合は第8表のようになる。

第8表 質問の形式と訪問の前後による「コード違い」の割合

質問の形式	訪問の	コード違い(%)	延質問数
多項選択	前	17.5	183
	後	22.9	292
二項選択	前	12.4	122
	後	11.7	197
自由回答	前	23.9	180
	後	28.0	193
基本項目 (フェース・シート)	前	3.6	449
	後	4.4	696

途中の質問を省略して調査をするという割合は、いくらか後半に多くなっているようであるが、「コード違い」の方は訪問の順には関係ないようにみえる。

e) 調査票単位にみたとき

これまででは、Bサンプルの記入したものと調査員の記入したものと的一致、不一致を質問を単位としてみてきたが、こんどは調査票を単位としてみることにする。まず、記入不一致の場合を、「コード違い」、「記入不足」、「フェース・シートの記入違い」、「質問省略」などの種類にわけてみると、1調査票当りの平均の誤りとその分散は第9表のようになる。

第9表 1調査票当りの誤りの数

誤りの種類	1調査票当りの平均	分散	延調査票数
コード違い	4.07	6.66	81
自由回答の記入過不足	0.52	0.52	81
基本項目記入違い	0.32	0.44	81
質問省略	1.49	12.41	81

これによれば、一般の質問項目における「コード違い」は1調査票（1サンプル）当り平均4ヶ所で、その他の記入過不足、記入違いなどはせいぜい1ヶ所である。また質問を省略される割合は平均すると1.5項目弱であるが、これは調査員（サンプル）の間で非常にばらつきが大きく2、3の調査員は、大半の質問を省略してしまっている。

つぎに、Bサンプルの方からみてみよう。調査は、同一のBサンプルのところへ、別の調査員が2人訪問するように企画されていたが、同じBサンプルに、異なる調査員が訪問したときに「コード違い」の数が非常に違った組合せが、Bサンプル31人中5人の場合にみられた。また、Bサンプルを性別、年齢別、職業別にわけて「コード違い」の数をみると、ほとんど差はみられない。これらのことからみると「コード違い」は調査員の方に、より一層その原因があるように思われる。

f) 調査員の方からみた記入の不一致

以下の節で調査員ごとの分析は詳しくのべる。ここでは、調査経験回数、不正行為の有無との関係のみをみておこう。調査の経験回数と1サンプル（1調査票）当りの「コード違い」、「質問省略」の平均数との関係は第10表のようになり、ほとんど関係はみられない。また、不正行為(cheat)との関係をみると、1調査員当りの不正行為の延数(cheatしたサンプル数)と記入不一致数との間には関係はみられないが、不正行為のあり、なしでは逆の関係がみられる。す

第10表 調査経験回数と「コード違い」

経験回数	コード違い	質問省略	Bサンプル数
0回	3.58	1.29	24
1	3.92	0.50	24
2	4.89	1.56	9
3	3.75	3.38	8
4	4.20	0.40	5
5	4.86	4.28	7
6回以上	5.25	1.25	4

第11表 不正行為と質問省略

不正行為	質問省略		計
	なし	あり	
なし	9	17	26
あり	14	7	21
計	23	24	47

(χ^2 5%有意)

なわち、不正行為の発見されなかった調査員26人中、Bサンプルについて「質問省略」のなかったのは9人にすぎない(第11表参照)。

3.3 記入不一致の種類による検討

a) 記入不一致の種類わけ

今までは、記入不一致をおもに「コード違い」、「質問省略」の点からみてきたが、ここで「コード違い」を細分し、つぎの4つの種類にわけて考察する。その種類はつぎの通りである。

イ) … 主に意見をきく質問のコード違い

問 6b, 8, 9a, 11, 12b, 16

ロ) … 一般の質問のコード違い

問 1b, 3, 4a(1番目), 4b, 5a, 10, 12a, 13a

ハ) … 自由回答のコード違い(記入過不足)

問 2, 5b, 7, 9b, 12c, 15

ニ) … 基本項目の記入違い

問 14, 17a, 17b, 18, 20, 21

ホ) … 各問の質問省略

イ) に入る質問群は意見をきく質問である。実際場面においては、回答をとるとき受け取り方が問題になるもの、あるいは調査員の判断が必要になる種類のもので、単純な事務的な記入違いとはみられないものであり、ロ)にくらべてやや複雑な性質の質問からなっている。これに対し、ロ)に入る質問群は、どちらかといえば事実、実態的なものをきく質問で、実際の回答時には、すぐどれかにきまるものであり、調査員の判断はあまり必要でないと思われる。ハ)は自由回答の質問群である。これは要点の取り方の相異によるもの、あるいは質問の突っこみ方が不足していることなどにより誤りが生じてくるものであるから、イ)とはやや異なる誤差と考えられよう。ニ)はロ)とほぼ同様なものであるが、ただ質問が被調査者の基本項目(フェース・シート)に関係しているのでロ)よりは回答がとり易いものと思われる。したがって、これらのイ), ロ), ハ), ニ)のコード違いは、ごく粗いい方をすれば、

イ) … 調査員の判断による誤り

ロ) … 一般の意見項目における事務的な誤り

ハ) … 調査員の質問追及(つっこみ)の仕方による誤り

ニ) … 事務的な誤り

とも考えられる。

b) 調査員別にみた「コード違い」

前項で述べた「コード違い」の種類別にみて、調査員ごとの平均を求めてみると第12表のよ

第12表 調査員当りの「コード違い」

コード違いの種類	調査員当りの平均	分散
イ	2.84	4.41
ロ	2.26	3.36
ハ	3.87	3.90
ニ	1.30	2.01
計	10.2	16.72

第13表 「コード違い」の種類間の相関

コード違いの種類	ロ	ハ	ニ
イ	0.113	0.062	0.002
ロ	—	0.165	0.212
ハ	—	—	-0.127

*) 調査員毎の数を算出するとき、Bサンプルを1人しか調査しないものは2倍してある、また質問を省略した場合は、その質問のコード違いを平均で置きかえて修正してある。

うになる。

これらの「コード違い」の種類相互の間の関係をみると、第13表のようになり、調査員ごとにみて、各種類の誤りの間の相関はほとんどない。これからみると、調査員の良否を検討するには「コード違い」の数だけを問題にするよりは、種類まで考えにいて検討の方が効果的だろうと思われる。

c) 「コード違い」と「質問省略」による調査員の検討

i) 調査員の評点

これまでに検討してきたことからみると、「コード違い」とか「質問省略」とかは、個々の調査員の質にも関係するが、どちらかといえば質問の形式とか、訪問時の環境とかに影響される面があるようにもみえた。そして、調査員ごとにみた場合、全部が偶然的なものともいえないであろう。そこで、各調査員の良否をみるため、つぎのように考えた。

各々の調査員について、前記のイ)、ロ)、ハ)、ニ)の4種類の「コード違い」およびホ)「質問省略」の5種類の誤りを考え、ある調査員が、それぞれの種類の誤りについて、その平均以上の個数の誤りをおかしていれば、その種類の誤りについては「悪い」とみることにした。すなわち、5種類の誤りで平均以上誤っていれば、その調査員の評点は5とすると、この評点の分布は第14表のようになる。

第14表 誤りの種類による評点

評点	0	1	2	3	4	5	計
調査員	5	13	12	12	2	3	47

また、この評点についてスケール・アナリシスをおこなうと、再現性係数は0.86になり、この評点と5種類の誤りとの関係はほぼ「自由回答法における質問追及の仕方による誤り」、「意見項目における調査員の判断による誤り」、「一般の項目における事務的誤り」、「基本項目における事務的誤り」、「質問省略」の順になる。従って、評点の悪い調査員は「調査員判断による誤り」、「質問の追及過不足による誤り」の外に各種の事務的誤り、質問省略がつけ加わることになるともいえよう。

ii) 評点による調査員の組分けと調査員資料との関係

調査員を前記の評点により、誤りの少いグループ(評点0, 1, 2)と、誤りの多いグループ(評点3, 4, 5)にわけ調査員資料のうち、事前にわかる資料との関係をみると、調査経験回数、説明会集合時刻ではやや差がみられるが、趣味(読書のあり、なし)、アルバイトの必要度などではあまり差はみられない。

第15表 評点と調査員資料の関係

評点	趣 味		経 験 回 数		集 合 時 刻		計
	読書あり	そ の 他	0, 1 回	2 回以上	定刻10分前	それ以後	
0, 1	9	9	15	3	13	5	18
2	8	4	6	6	7	5	12
3, 4, 5	7	10	6	11	6	11	17

ところで、調査員を使う立場からすれば、なるべく誤りの少い調査員のグループを選んだ方が実際にはよいから、事前に分る資料を用いて、調査員の選別が効果的にできれば都合がよい。

前記の調査員資料を要因として調査員の判別を試みると、判断成功率は約80%である。これからみると、事前の資料で調査員を選別することも可能であろうが、これは調査員の募集方法とも関連するので検討を要する問題である。

d) 調査員の良否と調査結果

前節では、Bサンプルと調査員の記入の不一致から得られた資料により、調査員の質をある意味で良否に判別しようと試みた。調査員には、Bサンプルと一般サンプルとの区別がないから、Bサンプルの記入と不一致が多かった調査員は、一般のサンプルに対しても不一致が多いということになれば問題であろう。

しかし、調査員を前節のように誤りの少ないグループと多いグループの2組にわけて、この両グループの調査員が調査した一般のサンプルの結果を比較してみると、通常おこなう比率の差の有意差検定では差はみられなかった。

調査員をこのような二つのグループに分ける方法にも関係するので問題はあるが、上に述べたことは「コード違い」の方向が、ある特定の方向には片寄っていないということを予想させる。

これまでにも述べたことであるが、われわれの調査結果によれば、調査員による記入誤差は、ごく特殊な場合をのぞいては一定の傾向をもつことはないように見える。したがって、調査員の記入誤差は、ごく一般的には、[2]における誤差モデルから検討できるものと思われる。

4. 要 約

面接調査における調査誤差については、種々の実験的研究がおこなわれているし、実際の調査の場面においての研究もいくつかみられる [3], [4]。ここでは、「面接調査における偏りの研究」[1]において得られた資料のうち、実際の調査場面において生じる誤りについて二、三検討した。

面接調査が指示通り正確におこなわれたかどうかをみるためには、通例簡単な郵便調査による面接状況のチェック、あるいは信頼できる調査員による再調査などがおこなわれているが、あらかじめ調査協力を依頼したサンプルに、面接状況および回答の様態を報告してもらうのも一つの方法である。このようにして、依頼したサンプルが記入した回答と、調査員の記入した調査票との関係をいくつかの面から検討した。要約すると、

1) 両者の記入の一致度をみると、質問の種類によって変動するが、年齢、学歴、持物などの事実項目や実態をきく項目での一致率は90%以上で高く、賛否の程度をきく項目あるいは自由回答法の質問では一致率は低かった。

2) 調査票リストを使用する質問と、そうでない質問とでは前者の一致率が高い。

3) 調査の途中で質問を省略したり、誤ってとばしてしまったものは、質問の種類や形式によって異なるが、複雑な質問および自由回答の質問などに多くみられた。また、調査票リストを使用するように指定された質問の方が質問を省略される割合が多い。

4) 記入の不一致をいくつかの種類にわけて、これらと調査員との関係をみると、一番誤りの多いのは自由回答の質問における要点の受け取り方の違い、質問追及の過不足によるものである。ついで調査員の判断による記入違いになっていて、単なる事務的な誤りはごく少い。

5) 前記の誤りの種類のそれぞれについて誤りの数が平均以上の調査員のグループと、そうでないものに分け、調査員に誤りの評点を与えた。これは調査員選別の一つの指標とみられ、「自由回答法における誤り」、「調査員判断による誤り」、「一般の事務的誤り」および「質問省略」の順に累加される傾向がみられる。

6) 調査員に関する資料のうち、調査経験回数、調査説明会への集合時刻など、調査員募集時にわかる調査員資料のみを用いて調査員の選別をすると、判断成功率は約80%になる。

7) 調査員は一般的にいて特定の方向にのみ誤るということはないと考えられるが、質問の形式や回答のとり方、調査員の訓練の仕方などにより特定の調査誤差を生じる場合もある。しかし、これらのうちの多くは適当な方法により改善できると思われる。

8) 以上のことから、たとえば質問票作成については、a) 回答が複雑なもの、あるいは、調査員の判断を必要とするようなものは、調査票リストを使用し、調査員の判断によるくい違いを少なくするようにした方がよいと考えられるが、一方では質問を省略される可能性がふえ、また回答を限定してしまうという不利が考えられるので、質問の目的により、リスト使用の得失を考えなければならないこと、b) 賛成、あるいは反対の程度が問題になるような場合は段階の判断にかなりの誤差が考えられるので、その対策を考える必要があること、などが分った。

調査員については、誤って質問をとばしてしまうもの、事務的誤りの多いものは、他の面でも誤りが多い傾向がみられるので、適当な方法により調査員の選別をするのがよいと思われるが、選別方法の検討は今後の問題である。

9) 調査員誤差の問題を誤差モデルで検討する場合、ごく一般的にはランダムな回答誤差を仮定できると思われるが、質問の形式等によっては特定の傾向をもつことも考えられるので、十分な検討が必要である。

統計数理研究所

参 考 文 献

- [1] 青山博次郎:「面接調査における偏りの統計的研究—I」統計数理研究所彙報, 第6巻, 第2号, 1959.
- [2] 林 知己夫:「回答誤差等を考慮に入れた標本調査計画」統計数理研究所彙報, 第5巻, 第1号, 1957.
- [3] 西平重喜:「面接調査法の諸問題」その1, 2, 3, 統計数理研究所彙報, 第3巻, 第1号, 1955, 第4巻, 第2号, 1956, 第5巻, 第2号, 1958.
- [4] H. H. Heyman and others: *Interviewing in Social Research*, 1954.