# …… いんふぉるむ（第9回）……

## ＜調査員の問題について＞

調査員に帰因するバイアスについては、ハイマンをはじめ多くの諸研究があるが、ここでは、比較的論及のすくない練習調査、イントネーション、エスニシティや調査員の自信などの面から掘下げた研究を外国文献（パブリック・オピニヨン・クォータリー）から紹介する。

### （調査員の選別に効果的な練習調査）

若し、調査員としての良否が簡単なテストで判別できるのであれば、調査会社は大助かりである。国立科学財団の助成でヒューストン大学マーケティング学部助教授ブレイアーがおこなった1979年の研究(*1)によると、本調査の前にあらかじめ調査員に実施させた練習調査での面接行動の誤りが本調査においても生じており、誤りの相関が高かったという。

*NORC*（*National Opinion Research Center*）の婦人調査員の中から調査能力の良いもの４１人を選び、全員に練習調査を一票実施させた。対象は友人・家族など任意とし、面接場面を録音した。調査内容は余暇活動、人生の満足と幸福、ギャンブル、飲酒、麻薬飲用習慣、性行動などの項目であり、全般的に面接のむつかしい質問から成り立つ。各調査員の録音テープを４カテゴリーについて点検した。第１は「非流暢さ」、第２は「読み誤まり」、第３は「プローブ（つっこみ）」、第４は「フィードバック」である。「非流暢さ」とは、質問を読みまちがえて言い直すなど、対象者が気付いてしまうようなものとし、「読み誤まり」は言葉を追加してしまうような対象者が多分気付かぬ程度のミスとした。「フィードバック」とは、面接場面でうなずいたり、あいずちを打つなどして、対象者にバイアスのかからない元気づけをすることである。

調査は面接法による全国調査で、成人1,172人（標本数の記載なし）を面接した。調査場面の録音を承諾した対象1,049人から無作為に３７２人を抽出し、前記４カテゴリーにつき調査員と対象者との面接テープを点検し、誤りを調査員別に集計した。その結果は練習調査と同じ傾向を示し、両者の相関係数は「非流暢さ」・76、「読み誤まり」・76、「プローブ」・70、「フィードバック」・70と非常に高かった。このことは、練習調査が調査員のフィールド行動予測の非常に良い材料であり、フィールド・パフォーマンスの的確な予言者であることを示すものだとブレイアーは述べている。

この研究を読んで興味ある点は、調査員の練習効果があまりないということである。一般に、調査票を読む場合、何度も繰返すうちに熟達し、流暢に読めるようになる。私ども調査機関では、調査員の流暢さによる早口を警戒し、注意を喚起している。ところが、この論文では、熟達による改善の効果よりも、欠点の執着・固守の傾向を指摘している。つまり、調査員ひとりひとりの個性的な癖、欠点はなかなか改善されず、持続するとの立場に立つ。だから、１票の練習調査で発見した癖は、２０〜３０票の本調査でも継続して出現したのであろう。１票の練習調査の内容をみて、調査員をスクリーニングできるという指摘は、調査機関としては注目される結論である。ただし、調査機関ＮＯＲＣは歴史の古い立派な調査機関として定評があり、そこで選抜した優秀調査員の評価の問題であることに注意すべきである。当然、評価の基準はきびしいものと考えられる。粒よりの調査員であるため、練習調査の段階から、レベルの高い面接場面を展開したのであろう。そのため調査慣熟による面接改善の傾向はなく、むしろ個々の調査員の個性的な癖が評価を左右し、このような結果となったことも考えられる。

### （イントネーションの影響）

調査員の声の抑揚（イントネーション）が上がり調子（*voice - rising*）の場合、対象者の回答行動に強い影響をあたえるという研究(*1)とそのような影響はないという反論の研究(*1)とを紹介しよう。

「影響あり」と云うのは、アルバド・バラス（ミシガン大・サーベイリサーチセンター準教授）とチャールス・キャンネル（同、教授）の1976年発表の研究であり、「影響なし」と云うのは、エドワード・ブレイアー（ヒューストン大マーケティング学部助教授）の翌７７年発表の研究である。

前者は米国公衆衛生局助成金による研究で、声の抑揚を上げた場合（*voice - rising*）と下げた場合（*voice - dropping*）とのどちらに肯定反応の影響があるかという問題とどちらの抑揚が対象者の調査関心に影響を与えるかの問題とを調査した。この実験の背景には、下げ調子のイントネーションは上げ調子に比べて、調査員のインタビュー関心の低さを想像させ、そのため、対象者の協力意識を低下させてしまい、否定反応を助長するのではないかとの仮説がある。

調査はトレド市、デトロイト市の無作為抽出した18〜65歳女性228人（標本数の記載なし）を９人の熟練した女性調査員が面接にあたった。２２８人のサンプルを上がり調子（*voice-rising*）、下がり調子（*voice - dropping*）の２グループ（同数）に無作為に分け、９人の調査員が２グループを同数ずつ受持った。健康状態４０項目を記した"*Yes - No*"リストを呈示し、最近１カ年間での症状の有無を質問した。質問する前に、あらかじめ、４０項目のリストを対象者に読ませ、症状経験の有無を想起させてから質問に入った。

その結果は、肯定回答の平均は「下がり調子」グループ（10.8）よりも「上がり調子」グループ（12.4）が多く、有意差がみとめられた（$t = 2.00$、$P < .025$）。４０項目についての２グループの反応は表１に示したが、「4. 皮膚のかゆみ」「6. 下痢」「11. 耳鳴り」「21. 皮膚炎」「35. 流感」「36. 憂鬱症」「39. アレルギー障害」などの症状では、とくに声の抑揚を上げた場合の*Yes*反応が増大している。

症状経験の有無を聞いたのち、同一リストについて健康上の関心を３段階尺度（「非常にある」１点、「少しある」２点、「まったくない」３点）で質問した。関心は「下がり調子」グループ（平均2.05）よりも「上がり調子」グループ（平均1.81）が高く（$t = 1.96$、$P < .05$）、声の抑揚を上げた場合、対象者の調査関心を促進する効果のあることを示している。

声の抑揚と結果との関係はつぎの４つが考えられる。

(1) 声の調子を上げると、正確な答を促進する（実際よりも少な目に言うようなことが減る）
(2) 声の調子を上げると、回答バイアス効果が働く（肯定回答がふえるなど）
(3) 声の調子を下げると、幾分正確な答を抑制する（否定回答を助長するなど）
(4) 声の調子を下げても、正確度には変わりがない

以上の実験から明らかになった結論としては、

（表１）　２グループの*Yes-No*回答分布（実数）

| | | （N：114）下げ調子（*Voice-Dropping*） | | | （N：114）上げ調子（*Voice-Rising*） | | | |
|---|---|---|---|---|---|---|---|---|
| | | YES | NO | N.A. | YES | NO | N.A. | (*) |
| 1 | のどが痛い | 71 | 42 | 1 | 74 | 39 | 1 | ＋ |
| 2 | 食欲がない | 30 | 84 | | 29 | 85 | | － |
| 3 | 動悸が早くなる | 34 | 74 | 6 | 36 | 78 | | ＋ |
| 4 | 皮膚のかゆみ | 34 | 79 | 1 | 56 | 57 | 1 | ＋ |
| 5 | 息切れ | 43 | 70 | 1 | 48 | 66 | | ＋ |
| 6 | 下痢 | 28 | 86 | | 42 | 71 | 1 | ＋ |
| 7 | 吐く | 28 | 86 | | 37 | 76 | 1 | ＋ |
| 8 | 頭痛 | 82 | 32 | | 93 | 21 | | ＋ |
| 9 | 身体がだるい | 35 | 78 | 1 | 41 | 73 | | ＋ |
| 10 | 四肢のいたみ | 48 | 65 | 1 | 56 | 58 | | ＋ |
| 11 | 耳鳴り | 16 | 99 | | 29 | 85 | | ＋ |
| 12 | 疲労感 | 46 | 68 | | 44 | 70 | | － |
| 13 | ストレス | 75 | 39 | | 71 | 42 | 1 | － |
| 14 | 悪夢 | 24 | 89 | 1 | 32 | 82 | | ＋ |
| 15 | 鼻血 | 5 | 109 | | 7 | 107 | | ＋ |
| 16 | 朝方の咳 | 28 | 86 | | 33 | 81 | | ＋ |
| 17 | 目まい | 36 | 78 | | 39 | 75 | | ＋ |
| 18 | 高熱 | 13 | 100 | 1 | 23 | 91 | | ＋ |
| 19 | 消化不良 | 38 | 76 | | 34 | 80 | | － |
| 20 | 歯痛 | 19 | 95 | | 19 | 95 | | 0 |
| 21 | 皮膚炎 | 21 | 93 | | 34 | 80 | | ＋ |
| 22 | リューマチ | 21 | 91 | 2 | 22 | 91 | 1 | ＋ |
| 23 | 気管支炎 | 13 | 101 | | 17 | 97 | | ＋ |
| 24 | うおのめ | 28 | 86 | | 22 | 92 | | － |
| 25 | 胃炎 | 27 | 86 | 1 | 36 | 77 | 1 | ＋ |
| 26 | 静脈炎 | 55 | 59 | | 48 | 65 | 1 | － |
| 27 | 怪我 | 10 | 104 | | 11 | 103 | | ＋ |
| 28 | 眠れない | 48 | 65 | 1 | 54 | 60 | | ＋ |
| 29 | 糖尿病 | 7 | 107 | | 2 | 111 | 1 | － |
| 30 | 歯痛 | 23 | 91 | | 21 | 93 | | － |
| 31 | 目、鼻の障害 | 22 | 92 | | 27 | 87 | | ＋ |
| 32 | 貧血 | 12 | 99 | 3 | 12 | 101 | 1 | 0 |
| 33 | 腎臓病 | 7 | 106 | 1 | 14 | 100 | | ＋ |
| 34 | ぜんそく | 4 | 109 | 1 | 5 | 109 | | ＋ |
| 35 | 流感 | 78 | 36 | | 90 | 24 | | ＋ |
| 36 | 憂鬱症 | 50 | 62 | 2 | 61 | 53 | | ＋ |
| 37 | ぼうこう炎 | 16 | 97 | 1 | 20 | 94 | | ＋ |
| 38 | 心臓病 | 4 | 107 | 3 | 6 | 107 | 1 | ＋ |
| 39 | アレルギー障害 | 11 | 102 | 1 | 30 | 84 | | ＋ |
| 40 | 脊椎障害 | 42 | 71 | 9 | 33 | 81 | | － |

(*) 声の調子を上げたことによる"*Yes*"回答の増減　＋　増　／　－　減　／　0　変わらず



調査員の発声のイントネーションが、経験の想起率など比較的やさしい質問の場合に、回答行動に強い影響を与えるということであり、上記(1)と(2)の例に該当すると報告している。

アルバド・バラス等がこの実験を思い立った理由は、調査員の面接場面の録音テープを聴いて、"*Yes-No*"や"*Agree-Disagree*"を尋ねる長いリストを読む際に、普通の質問文のときよりもイントネーションが下がるという共通傾向を発見したことによる。この理由として、リストが一語または一フレーズの沢山の集まりであるため、調査員の注意が散慢になりやすいためと推測している。日本語とちがい、英語の文章は、構文、フレーズ、文意などによりイントネーションの上げ下げが規定されている。調査員の恣意的な変更・誇張は起こり得ない。そのため、何の制約も受けない"*Yes-No*"リストの場合に、調査員の調査姿勢のようなものが自然とにじみ出てきて、精神の高揚していない調査員は*voice-dropping*、張り切っている調査員の場合は*voice-rising*傾向となることが考えられる。そのイントネーションが回答者に強いインパクトを与えているという傾向がもし事実であるならば、もっとシステマティックな研究を継続し、インタビューテクニック構築に役立てたらどうかというのが論旨である。

ところが、翌年には早速反対論があらわれた。国立科学財団助成によりヒューストン大助教授ブレイアーは、調査員の日常の自然的なイントネーションを分析した。バラス等は、調査員を２グループに分け、イントネーションの上げ・下げを指示し、人為的な実験場面を設定したが、ブレイアーは事前には調査員にまったくイントネーションに関する注意をせず、通常の面接場面を録音させた。標本数の記載はないが、全国調査1,172人のうち1,048人がテープ録音を承諾した。その1/2を無作為抽出し、テープ半転で録音の一部が途切れたとか、録音不良のものなど４１人を除外し、最終的に４８３人を分析対象とした。調査主題は余暇活動、始めに余暇活動全般について１２項目、次いでスポーツ活動につき９項目の"*Yes-No*"リストを使用した。テープ分析は２組のコーダーが担当し、"*Yes-No*"リストの全項目について、語尾が上がって終っているかどうかを判定した。

結果は表２にみられるように、ミシガン大グループの場合と逆転している。上げ調子における"*Yes*"回答の増加例は２１項目のうち６項目にとどまり、その増加割合も僅かであり、明らかに積極的な影響がみとめられない。むしろ下げ調子の方に*Yes*回答が多くなっている。

回答者の属性で分析したが、関係はみられなかった。一般に、高齢対象に対して上げ調子イントネーションを使うことがある。これは調査員が、高齢者への元気づけに使うのであるが、もし、この傾向が事実とすれば、上げ調子は高齢対象に集中し、その結果、高齢者層の余暇活動の実態を反映して、"*Yes*"回答が減るという影響が考えられる。しかし、分析の結果では、調査員は高齢者に意図的に上げ調子イントネーションを採用していなかった。

ミシガン大グループの結果で考えられることは調査員の地方的特性ではないかとブレイヤーは述べている。ミシガン大では９人の調査員を使ったが、こちらは５９人の調査員を使っている。５９人の調査員（性別は記載なし）について、宗教別、都市規模別などで分析してみたが、なんら影響はなかった。だから考えられることはミシガン大で僅か９人の調査員を使用した偏りではないかと言う。そこで結論としてブレイアーは、調査員イントネーションの上げ調子は"*Yes-No*"質問に何の影響も与えないとし、ミシガン大で影響があっ

（表２）　下げ調子、上げ調子における“*Yes*”回答の比率

| | | 下げ調子 (*Voice-Dropping*) N | 下げ調子 (*Voice-Dropping*) YES(%) | 上げ調子 (*Voice-Rising*) N | 上げ調子 (*Voice-Rising*) YES(%) |
|---|---|---|---|---|---|
| 1 | 映画に行った …… | 350 | 37.7 | 132 | 38.6 |
| 2 | レストランに行った …… | 326 | 73.3 | 156 | 70.5 |
| 3 | ウィンドーショッピングをした …… | 301 | 63.8 | 181 | 57.5 |
| 4 | 演劇、演奏会に行った …… | 307 | 21.8 | 174 | 19.0 |
| 5 | ピクニックに行った …… | 286 | 47.0 | 195 | 36.9 |
| 6 | 狩り、魚つりに行った …… | 310 | 20.7 | 172 | 20.9 |
| 7 | 読書した …… | 306 | 81.4 | 177 | 83.0 |
| 8 | ドライブした …… | 321 | 70.9 | 161 | 69.6 |
| 9 | 庭仕事をした …… | 308 | 55.2 | 172 | 57.9 |
| 10 | 市民団体、宗教団体に参加 …… | 345 | 45.5 | 138 | 38.4 |
| 11 | 散歩やハイキング …… | 311 | 59.4 | 170 | 55.3 |
| 12 | スポーツをみに行った …… | 402 | 11.4 | 79 | 7.6 |
| 1 | バドミントンをした …… | 325 | 20.6 | 157 | 17.2 |
| 2 | バスケットボールをした …… | 191 | 26.2 | 290 | 20.3 |
| 3 | ボーリングをした …… | 190 | 31.1 | 290 | 26.2 |
| 4 | フットボールをした …… | 202 | 15.8 | 280 | 13.2 |
| 5 | ゴルフをした …… | 169 | 11.8 | 312 | 10.6 |
| 6 | ラケットボールなどをした …… | 286 | 10.5 | 192 | 10.9 |
| 7 | ソフトボール、野球をした …… | 269 | 36.4 | 210 | 25.7 |
| 8 | 泳いだ …… | 208 | 54.3 | 273 | 45.0 |
| 9 | テニスをした …… | 366 | 18.6 | 111 | 18.9 |

たのは、調査テーマが健康というプライベートな領域であったためか、調査員バイアスか、サンプリング・エラーによるものではないかと述べ、一歩譲っても、上げ調子イントネーションは特別の状況においてのみ肯定反応をひき起すことがあるかも知れぬと言えるだけだと述べている。

私は、ミシガン大の実験で調査員バイアスはなかったと思う。アメリカで超一流といわれるミシガン大サーベイ・リサーチ・センターの調査員の中からよりすぐった９人の女性調査員であり、実験条件はまったく問題がないとみられる。問題なのは、対象が女性であること（ヒューストン大のは記載がなく、恐らく男女個人）、テーマが他人に知られたくないような疾病体験であったため質問のイントネーションが心理的なストレスに作用したと考える。だから、この事実だけで一般化は無理としても、特定の場合には大いに起こり得る傾向と考えられる。

この研究から学ぶべきことは、日本でのこの種の研究が全くないという反省である。日本では文盲率の低いこともあって、視力のおとろえた高齢者は別として、リスト（回答票）を読み上げることがなく、リストを読み上げることによって生ずるバイアスは生じない。したがって、問題は質問文のイントネーションに影響があるかということである。質問はすべて疑問文であるため、語尾のイントネーションを上げる人がいる。一方、疑問文であっても平叙文のように淡々と読む人もいる。この場合、どっちが良いかの指導ができる実験データはないのである。



## （調査員の人種の影響
## ―― 黒人・白人の関係 ―― ）

人種問題を扱った調査質問において、回答者と調査員との人種が異なる場合（例えば黒人と白人）と同じ場合（例えば黒人と黒人、白人と白人）とでは回答結果が違うという傾向については、従来から指摘されている。回答者は、同じ人種の調査員に対してはフランクであるが、違う人種の調査員には特別の感情をいだくことがある。この傾向は、すべての調査質問にあてはまるのではなく、人種問題のテーマに限られる。サドマンとブラッドバーン（１９７４年）やキャンベル（１９８１年）らの研究報告があるが、いずれも個人面接調査での結果である。

この傾向が、果して電話調査の場合にも起こり得るかを問題としたのが、これから紹介するコーター[*4]らの研究である。これは、アラバマ大学政治科学学部の助教授２人（コーター、コーエン）と教授（クールター）の共同研究（１９８２年発表）である。面接法とはちがい電話調査では、回答者と調査員とは相互に皮膚の色が見えない。物理的距離ばかりでなく心理的距離も離れている。この調査場面において人種間の特別な感情は湧かないはずである。

ところが、意外にも電話調査の場面でも、調査員の人種に帰因する回答バイアスが発生しているようだ。調査の際に、回答者は調査員の人種を探るために、何か言語を使った合図（*verbal cues*）を発しているのかも知れぬという。調査は１８歳以上アラバマ市民を無作為抽出し（グローブやワクスブルグが１９７８年に電話調査用に開発した抽出方法）、アラバマ大学キャッブストーン調査所が政治・社会問題（人種問題を含む）について電話調査を実施し、５９０人の回答（標本数については記載なし）を得た。

女性の南部出身者を主とする地元調査員１２名（高学歴者）を使用し、無作為にサンプルを割当てた。回答者の人種をきく質問は調査票末尾におき、調査員が事前に人種を知ってしまうことによる影響を排除した。回答者５９０人のうち人種を答えた人は５４８人で、データ分析は白人４３９人、黒人１０３人の５４２人についておこなった（６人はスペイン人や東洋人のため除く）。白人調査員（９人）の担当したサンプル３６９人の内訳は白人３０２、黒人６７であり、黒人調査員（３人）のサンプル１７３人の内訳は白人１３７、黒人３６である。

調査項目は「人種問題」13問と「非人種問題」１０問とからなる。調査員の人種に帰因する影響は、「人種問題」の質問の際に生じており、とくに黒人調査員が白人回答者にインタビューする場合に影響が発生したが、「非人種問題」の質問では影響はなかった。黒人回答者よりも白人回答者に強いインパクトを与えたという傾向は、従来とはちがう傾向であるが、その理由として、コーターらは設問の構成を指摘している。それは、設問が黒人を直接問題とするものばかりで、白人を問題とするものがなかったためだと云う。答える際に、黒人対象者は白人調査員に対して気兼ねする必要がなかったためではないかと推定している。

なお、この結果を一般化する場合の注意として、この調査が人種問題の盛んな地域で実施されたことをあげ、別の地域で電話法による再実験の必要を指摘している。ともあれ、それまでのあいだは、電話調査をする場合でも、人種問題のテーマの場

合、気を使うべきだという。実際問題として、調査員と回答者との人種関係を事前に管理することは至難のわざに属する。厳密に、このバイアスを排除しようというのなら、調査員がインタビューする前に、あらかじめ回答者の人種をしらべなくてはならないと匙を投げている。

### （調査員の人種の影響
### ―― エスニック・マイノリティとの関係 ――）

アメリカは人種の坩堝と言われる。黒人・白人の２分法で分割される白人のなかでもＷＡＳＰと呼ばれるアングロサクソン系ピュリタンの本流とそれ以外の非ＷＡＳＰとがある。また、黒人・白人以外にアメリカ原住民や中国人など少数民族もいる。

調査員と回答者との人種関係を黒人・白人の関係でしらべた文献は前述の調査をはじめとして割合あるが、黒人以外のエスニック・マイノリティ（少数民族）との相互関係をしらべたものはあまりない。そこでミハエル・ウィークスとポール・モールが1975年に行った調査を1981年に発表(*5)しているのでそれを紹介しよう。

調査は４地域、１４７２人の小学生で、学区の教育委員会名簿より該当人種を抽出した。抽出人種はキューバ人、メキシコ系アメリカ人、アメリカ原住民、中国人の４種で地域はそれぞれマイアミ（キューバ人）、エルパソ（メキシコ系）、アリゾナ（原住民）、サンフランシスコ（中国人）である。抽出に際しては1/2を英語を話す子ども、1/2を英語を話さない子どもにしたが、その判定は教育委員会に従った。調査員（１０１人）は５０人の該当人種と５１人の非該当人種、後者は若干調査経歴に優るが、全員３日間の調査訓練を実施した。さらに、全サンプルの１５パーセントについては、インタビューに同行して観察、全サンプルの１０パーセントは調査終了後監査をおこない精度の均質に留意した。

調査内容は英語能力に関することである。調査手順は、まず対象児のいる世帯から、世帯の事情を知っている１４歳以上の世帯員を対象児の代理として選定してもらい、面接法により世帯のバックグランドを調査した。調査項目は世帯での使用言語や英語会話能力、アメリカ移民の時期など１０問からなり、その結果は英語能力（教育委員会が判定した評価）との相関が高く、英語能力の予測に有効とされたものを使用している。世帯調査（バックグランド調査）が終ってから、ひきつづき対象児にたいし英語能力のペーパーテストを実施した。ペーパーテストの質問は英語で印刷されている。テストの指示は英語だが、必要な場合は通訳した。ただし、質問文は通訳しない。

実験の狙いは、エスニック調査員とノン・エスニック調査員とで、バックグランド調査とペーパーテストとの結果がちがうのではないかという仮説である。バックグランド調査とペーパーテストによる英語能力スコアとは、従来、相関の高いことが分っており、調査員バイアスがなければ結果は一致するはずである。しかし、エスニック調査員が同じエスニック世帯を訪問する場合のフランクな調査場面とノン・エスニック調査員の調査場面とは微妙なちがいがありはしないかということであろう。ところが、結果は人種による調査員バイアスがまったくなかったのである。

アメリカ移民にとってアメリカ人になるということは、英語をおぼえることであった。英語をおぼえ、アメリカン ウエイ オブ ライフを身につけることであった。英語がよく話せないということはアメリカ人として失格することであった。このような重大な意味をもつ英語能力については、エスニック・マイノリティの人は何らかのコンプレックスがあるかも知れない。ということから出発したこの実験は、仮説が検証できなかった。最近のアメリカは、エスニックの実態に即して、*bi-linguistics* つまり英語と母国語の２カ国語を初等教育で教える方向になってきている。「人種の坩堝」の考え方は、今や*myth*（神話）になっているという考え方もある。*myth*は架空の話という意味もあり、多様な民族性を肯定する考えが強まっている。このような風潮が背景となって、この結果が生じたのかも知れないと思う。ともあれ、興味ある実験である。

### （調査完了に明るい見とおしをもつ
### 調査員の回収率は良い）

われわれの経験では、むつかしい調査の場合、質問票の中身をみて恐れをなし、出かけるのをためらうような気の弱い調査員は、多くは途中で放棄するか低い回収率で帰ってくる。一方、対象者の調査協力説得に自信のある、活発な調査員は、多くは高回収率で堂々と帰ってくる。このように、対象者の回答協力への調査員の期待感と実際の回収率とのあいだには、強い関係があるようだ。

この点に関して、電話調査で実験した例(*6)（1983年発表）がある。研究者はエレアノール・スインガー（社会科学センター主任研究員）、マーチン・フランケル（バルーチ大学統計学教授）、マルク・グラスマン（統計コンサルタント：ニューヨーク市在住）の３人である。調査機関はＮＯＲＣ（*National Opinion Research Center*）。調査員３５人の８割は１年以内の調査経験又は未経験者であり、また、調査員の８割は黒人であった。調査項目は一般的な余暇活動につづき、心情的幸福感、精神衛生、飲酒、マリファナ使用、性行動その他属性として収入をきいた。完了は1,１００（サンプル数の記載なし）。

調査に入る前に、調査員は調査の難易についての予想を評価した。対象世帯から対象者を選定する作業の難易度、調査協力説得の難易度、質問の各項目についての無回答率などを事前に見込んでもらった。

その結果判明した点は次のとおりである。

① 調査員の属性との関係では、高年齢調査員ほど対象の協力度が高く、スクリーニングや回収率の成績が良い。これは、加齢に伴なう特質として、自信のようなものが声の調子ににじみ出てきて、回答者を安心させるのであろう。

② 調査員の調査経験との関係はほとんどない。まったくの未経験者よりは１度でも２度でも経験のある人の方が成績が良いとは言えるが、経験が深まったからといって成績が良くなることはない。

③ 調査員への割当を多くすると、スクリーニングや回収率が悪くなる。仕事が増えたことによる調査員の士気低下が原因なのか、困難な例がとくに混入したためなのかは分らない。

④ スクリーニングや回収率で悲観的な予想をした調査員は、楽観的な評価をした人よりも成績が良くない。自信のない調査員の回収率は低く、自信のある調査員の回収率は高い。調査票に対する調査員の態度が、調査票そのものよりも、回収率に影響をあたえるということを示唆している。

⑤ 上記の諸点については、面接調査の事例で調査したサドマン（１９７７）やシンガー（１９７９、１９８２）らの先行研究があるが、傾向は大体一致している。しかし、各質問についての無回答率と調査員の事前予想との関係はなく、調査員の属性との関係もみられなかった。この点は先行研究と異る。

この研究で少し気になるのは、ＮＯＲＣのよう

な一流調査機関を使用しながら、何故３５人の電話調査員の構成を、８割が黒人、８割が１年以内の調査経験という新人編成にしたのかということである。個人面接の場合でしらべたサドマンやシンガーの先行研究の追試実験を同じ調査機関ＮＯＲＣを使ってやるのなら、調査員の条件を先行研究に合わせて、白人女性中心の経験者で揃えるべきではなかったか。敢て勘ぐるならば、*Face-to-Face*でなく、回答者の顔色を気にしなくてもよい電話インタビューの場面で調査員バイアスを発揮させるには、黒人の新人調査員の方が仮説データが出やすいと研究者が判断したのではないかなどと疑いたくなる。調査員の条件を大体同じにしなければ、電話調査でも同様の調査員バイアスがあるとは断言できないのではないか。一歩譲って、調査員構成は問題がなかったとしても、先行研究は７０人の調査員、今回は1/2の３５人であるのが気にかかる。一般に、電話調査員の問題点として指摘されていることは、動員調査員の少ないことによるバイアスの問題である。動員調査員が少なければ少ないほど１人当り受持サンプルが多くなり、調査員によるバイアスは何倍にも増幅される可能性がある。このような欠点はあるが、調査員のパーソナリティが調査の質に影響をあたえるという結果は注目すべき指摘であるように思う。


＊1 *Edward Blair*
1980 "*Using Practice Interviews to Predict Interviewer Beheviors*"
*Public Opinion Ouarterly* 1980 (*vol.* 44－2)

＊2 *Arpad Barath Charles F. Cannell*
1976 "*Effect of Interviewer's Voice Intonation*"
*Public Opinion Ouarterly* 1976 (*vol.* 40－3)

＊3 *Edward Blair*
1977 "*More on the Effects of Interviewer's Voice Intonation*"
*Public Opinion Ouarterly* 1977 (*vol.* 41－4)

＊4 *Patrick R. Cotter, Jeffrey Cohen, and Philip B. Coulter*
1982 "*Race-of-Interviewer Effects in Telephone Interviews*"
*Public Opinion Ouarterly* 1982 (*vol.* 46－2)

＊5 *Michael F. Weeks and R. Paul Moore*
1981 "*Ethnicity-of-Interviewer Effects on Ethnic Respondents*"
*Public Opinion Ouarterly* 1981 (*vol.* 45－2)

＊6 *Eleanor Singer, Martin R. Frankel, and Marc B. Glassman*
1983 "*The Effect of Interviewer Characteristics and Expectations on Response*"
*Public Opinion Ouarterly* 1983 (*vol.* 47－1)